Centrum voor Wiskunde en Informatica
Centre for Mathematics and Computer Science
P.O. Box 4079, 1009 AB Amsterdam, The Netherlands

The CWI is a research institute of the Stichting Mathematisch Centrum, which was founded on February 11, 1946, as a nonprofit institution aiming at the promotion of mathematics, computer science, and their applications. It is sponsored by the Dutch Government through the Netherlands Organization for the Advancement of Research (N.W.O).

CWI Tract 61

Central limit theorems for generalized multilinear forms

P. de Jong

Centrum voor Wiskunde en Informatica
Centre for Mathematics and Computer Science

1980 Mathematics Subject Classification: Primary: 60F05, 60G15. Secondary: 62E20, 15A69.
ISBN 90 6196 369 9
NUGI-code: 811

Preface

This monograph contains the main results of a research which resulted from my interest in graph theory (and its applications in social science research). In the course of this research the aspect of graph theory gradually disappeared. Having finished my mathematics studies I used the subject for my Ph. D. research.

I owe many thanks to dr A.A. Balkema. His ability to simplify seemingly complicated matters has led me to unify what often appeared to me a collection of curious but interesting results. I treasure our stimulating discussions on mathematical subjects.

Prof. dr J.Th. Runnenburg read the final draft with painstaking precision. Many errors and omissions were detected and avoided. For this, and his many helpful suggestions I am very grateful. Of course I bear responsibility for all deficiencies in this work.

With drs Bert van Es I had some interesting discussions on the more practical sides of my research.

The personal attention of Teyung Fu made it possible to produce this text on a modern text-processing device. The joint efforts of drs Edo Velema and drs Antje Melissen improved the readability of this monograph.

Contents

1 Introduction and summary 1

2 A central limit theorem for clean random variables 10
- 2.0 Introduction 10
- 2.1 A central limit theorem for components in the Hoeffding decomposition 10
- 2.2 Formulation of the main result 13
- 2.3 Proof of Theorem 2.2.3 18
- 2.4 Estimating the quantities τ and τ^* 26

3 Extensions and variations 32
- 3.0 Introduction 32
- 3.1 Miscellaneous results 33
- 3.2 Results involving singular values and truncation 37
- 3.3 Inhomogeneous sums 51

4 W(n) as a Gaussian process 64
- 4.0 Introduction 64
- 4.1 Simple operations on W(n) 67
- 4.2 Convergence to a Gaussian process 73
- 4.3 The dimension of a multilinear form 77

References 81

List of terms 83

List of symbols 84

1. Introduction and summary

In this monograph we present some central limit theorems for homogeneous multilinear forms and for generalizations of multilinear forms. We are concerned with one particular generalization of the homogeneous multilinear form: 'clean' random variables to be introduced in Sect. 2.2. The purpose of this introduction is to make the reader acquainted with the subject matter of this monograph rather than to present the results of the subsequent chapters in full generality. Some special cases may serve to illustrate some peculiarities of the subject matter and of the methods used in the proofs below.

We start with a sketch of the general setting. Consider a probability space $(\Omega,\mathcal{F},P)$ on which independent random variables $X_1,\ldots,X_n$ are defined. Define for a finite subset $I \subset \{1,\ldots,n\}$ the σ-algebra $\mathcal{F}_I = \sigma\{X_i : i \in I\}$ and let W_I be a $\mathcal{F}_I$-measurable random variable. We assume the random variables W_I to be centered, square integrable and uncorrelated:

$$E\,W_I = 0,\quad E\,W_I^2 = \sigma_I^2 < \infty,\quad E\,W_I W_J = 0 \text{ if } I \neq J.$$

Notice that the distribution of the underlying random variables X_i is immaterial. We can write

$$W_I = w_{nI}(X_{i_1},\ldots,X_{i_d}) \text{ for } I = \{i_1,\ldots,i_d\},$$

with $w_{nI}(\ldots)$ a Borel measurable function $\mathbb{R}^d \to \mathbb{R}$ which may depend on n. (We shall suppress the subscript n where possible.) The random variables W_I are *dissociated*, that is $W_{I_1},\ldots,W_{I_q}$ are independent if the sets $I_1,\ldots,I_q$ are mutually disjoint. (See McGinley and Sibson (1975).)

We shall mainly be concerned with conditions that ensure asymptotic normality for d-homogeneous sums,

$$Z(n) = \sum_{|I| = d} W_I,$$

where the summation extends over all $\binom{n}{d}$ subsets $I \subset \{1,\ldots,n\}$ of size $|I| = d$. It is convenient to assume the sum Z(n) to be normalized to have unit variance:

$$\sum_{|I| = d} \sigma_I^2 = 1.$$

The following condition will play a crucial role in the theory below:

$$E\,Z(n)^4 \to 3 \quad \text{for } n \to \infty,$$

with 3 being just the fourth moment of the standard normal distribution. In the proofs below we need a technical condition

$$E\,W_I^4 \le D\,\sigma_I^4, \text{ with D not depending on n.}$$

Under these assumptions we do not have a central limit theorem; Z(n) may converge (even in the simple case d = 1) to any centered random variable with unit variance and fourth moment equal 3. What is needed is a negligibility condition which forces the contribution of each individual random variable X_i to the total variance to be small with respect to the total variance:

$$\max_i \sum_{I \ni i} \sigma_I^2 \to 0 \quad \text{for } n \to \infty.$$

Do these three conditions above ensure asymptotic normality for the homogeneous sum Z(n)? The answer in general is no; more structure is needed. However, in the important special case of homogeneous multilinear forms in independent centered random variables,

$$Z(n) = \sum_{|I| = d} a_I \prod_{i \in I} X_i,$$

the above assumptions imply asymptotic normality for Z(n). (In fact, it will be shown that, given the negligibility condition and the uniform bound on the fourth moments of W_I / σ_I, the convergence of the fourth moment to 3 is also a necessary condition for asymptotic normality.) This result on multilinear forms follows from the results in the next two chapters (especially Th. 2.1.1 for the if part and Th. 3.2.5 for the only if part). These results are valid for more general random variables W_I.

Before introducing the more general case, we shall consider multilinear forms in some detail, especially the bilinear case. The above mentioned central limit theorem is not completely self evident. For the quadratic form in iid normal N(0,1) random variables

$$Z(n) = \sum_{1 \le i \ne j \le n} a_{ij} X_i X_j$$

there is a simple proof for the asymptotic normality of Z(n). However, this proof rests on a non-trivial result from linear algebra and on a special property of the normal distribution, as can be seen from the following sketch of the proof.

Without loss of generality we may assume the matrix (a_{ij}) to be symmetric with zero diagonal, $a_{ii} = 0$. There is an orthogonal transformation that brings (a_{ij}) into diagonal form, and we can rewrite Z(n):

$$Z(n) = \sum_{1 \le i \le n} \mu_i Y_i^2,$$

with μ_i the eigenvalues of the matrix (a_{ij}) and the random variables Y_i normal N(0,1), orthogonal and hence independent. Since the diagonal elements vanish, we have

$$\sum_{1 \le i \le n} \mu_i = \operatorname{trace}(a_{ij}) = \sum_{1 \le i \le n} a_{ii} = 0,$$

and Z(n) is a weighted sum of independent centered random variables

$$Z(n) = \sum_{1\le i\le n} \mu_i (Y_i^2 - 1), \text{ with } \operatorname{var} Z(n) = 2 \sum_{1\le i\le n} \mu_i^2.$$

Assume var Z(n) = 1. The above considerations imply that Z(n) has a normal limit distribution iff

$$\max_i \mu_i^2 \to 0 \quad \text{for } n \to \infty,$$

which is equivalent to

$$\sum_{1\le i\le n} \mu_i^4 \to 0 \quad \text{for } n \to \infty.$$

Straightforward calculation shows that the latter condition is equivalent to

$$E\, Z(n)^4 \to 3 \quad \text{for } n \to \infty.$$

The above proof combines two approaches, an algebraic one: the orthogonal decomposition of symmetric matrices, and a probabilistic one: the special properties of the normal distribution and a simple central limit theorem. The proof itself has a limited scope: If the random variables are not normally distributed, the orthogonal decomposition results in a weighted sum of squares of uncorrelated random variables. Moreover, if $d \ge 3$, then there is no orthogonal decomposition in the above sense.

Matrices with 'many' zero entries seem easy to handle by a probabilistic approach. Especially block diagonal matrices (i.e. matrices divided into blocks by partitioning the index set {1,...,n}, with only those blocks which meet the diagonal containing non-zero entries) allow a simple analysis: The quadratic form can be written as a sum of independent random variables V_r, with V_r the rth block around the diagonal.
Whittle (1964) gives an interesting example. The matrix (a_{ij}) is defined by $a_{ij}^2 = p_{j-i}$ if $j > i$ and $a_{ij} = 0$ else, with $p_1 + ... + p_n = 1$. Then var Z(n) / n → 1, since

$$\sum_{1\le i\le n} (i/n)\, p_i \le (\sqrt{n})/n + \sum_{i>\sqrt{n}} p_i \to 0, \quad n \to \infty.$$

Taking blocks of size k_n (with $k_n = [\sqrt{n}]$, the largest integer not exceeding $\sqrt{n}$), we have

$$Z(n) = V_1 + ... + V_{k_n} + R_n,$$

with $V_1, ..., V_{k_n}$ iid and var $R_n / n \to 0$, since the random variables $a_{ij} X_i X_j$ are orthogonal and since var $V_k / k_n \to 1$ for $n \to \infty$, by the same argument as for the total variance.

The above approach can be easily generalized to multilinear forms or to uncorrelated random variables W_I. This raises the important preliminary question: Is it possible to rewrite homogeneous sums Z(n) in a trivial way as a sum of independent random variables plus a vanishing remainder term, as in the example above? The answer to this question is provided by the Gaussian example above.

Consider the matrix (a_{ij}) with all off-diagonal entries equal (and positive) and $a_{ii} = 0$; then Z(n) is asymptotically chi-squared distributed:

$$Z(n) = a_{12} \sum_{1 \le i \ne j \le n} X_i X_j = a_{12} \Big(\sum_{1 \le i \le n} X_i \Big)^2 - a_{12} \sum_{1 \le i \le n} X_i^2 ,$$

with $a_{12} = (n(n-1))^{-1/2}$ (since var Z(n) = 1). The variance of the second term equals $2n\, a_{12}^2 = 2 / (n-1)$ and hence this term tends to 1 in L^2. The first term equals $(n / (n-1))^{1/2}\, Y^2$, with Y standard normal. Thus Z(n) has a non-normal limit distribution.

In Sect. 4.0 it is shown how to construct a matrix (a_{ij}) with all off-diagonal entries having equal absolute value (and diagonal elements equal 0), such that the eigenvalues vanish uniformly for $n \to \infty$. In this case Z(n) is asymptotically normal, as is shown above. This shows that Z(n) may have a normal limit distribution, while block diagonalization fails. More generally, this example shows that any condition for asymptotic normality which is phrased in terms of the absolute values $|W_I|$ is not sharp.

We have now touched upon the main themes of the next two chapters. We shall give a survey of these chapters.

Sect. 2.1 begins with an important generalization of multilinear forms in independent random variables. Again we start with a probability space $(\Omega, \mathcal{F}, P)$ on which independent random variables $X_1, \dots, X_n$ and the σ-algebras $\mathcal{F}_I = \sigma\{X_i : i \in I\}$ (with $\mathcal{F}_\varnothing$ the trivial σ-algebra) are defined. Then a square integrable $\mathcal{F}_{\{1,\dots,n\}}$-measurable random variable Z(n) can be approximated by a sum of independent random variables:

$$Z(n) = \sum_{1 \le i \le n} E\,(Z(n) - E\,Z(n) \mid X_i\,) + R(n),$$

with the remainder term R_n orthogonal to the independent random variables $E\,(Z(n) - E\,Z(n) \mid X_i)$. If the remainder term vanishes (e.g. in L^2) for $n \to \infty$, then Z(n) can be analysed as a sum of independent random variables.

Our main concern is the situation where the remainder term does not vanish. In many interesting problems the latter is the case. To analyse this situation we pursue the projection in the following way.

Any square integrable $\mathcal{F}_{\{1,\dots,n\}}$-measurable random variable Z(n) can be decomposed:

(1.1.1) $$Z(n) = \sum_{I \subset \{1,\dots,n\}} W_I ,$$

where the random variables W_I are uniquely determined by the following conditions:

a) W_I is $\mathcal{F}_I$-measurable,

b) $E(W_I \mid \mathcal{F}_J) = 0$ a.s. if $I \setminus J \neq \emptyset$.

Thus

$$W_\emptyset = E\, W_\emptyset = E\,(Z(n) - \sum_{J \neq \emptyset} W_J) = E\, Z(n)$$

and

$$W_I = E\,(Z(n) - \sum_{J \neq I} W_J \mid \mathcal{F}_I) = E\,(Z(n) - \sum_{J \subsetneq I} W_J \mid \mathcal{F}_I) \text{ a.s.}$$

The decomposition is orthogonal. If $I \neq J$ the symmetric difference $I \,\Delta\, J = (I \setminus J) \cup (J \setminus I) \neq \emptyset$. Suppose $J \setminus I \neq \emptyset$, then $E\, W_I W_J = E\, W_I\; E\,(W_J \mid \mathcal{F}_I) = 0$.

The above decomposition was used in Hoeffding (1948) to obtain central limit theorems for Z(n), Z(n) being approximately a sum of independent random variables. We shall refer to (1.1.1) as the *Hoeffding decomposition* (see Van Zwet (1984)).

For d-homogeneous sums in the Hoeffding decomposition satisfying the negligibility condition and with a uniform bound on the fourth moments $E\,(W_I / \sigma_I)^4 \leq D$ for all I, the fourth moment condition $E\, Z(n)^4 \rightarrow 3$ implies asymptotic normality (Th. 2.1.1). In Sect. 2.1 it is shown that the assumption of homogeneity in $|I|$ cannot be dropped. Th. 2.1.1 follows from a slightly more general theorem (Th. 2.2.3).

In Sect. 2.2 we drop the assumption of underlying independent random variables. Then we only have a family of random variables indexed by finite subsets of the integers, $\mathcal{W} = \{W_I : I \subset \{1,...,n\}\}$ and the σ-algebras generated by (subsets) of these random variables. Define $\mathcal{F}^{(i)} = \sigma\{W_I \in \mathcal{W} : i \notin I\}$. Random variables W_I are *clean* if

$$E\,(W_I \mid \mathcal{F}^{(i)}) = 0 \text{ a.s. for all } i \in I.$$

For d-homogeneous sums of clean random variables W(n) we have a central limit theorem under the conditions of Th. 2.1.1 if we add as extra condition that the sum of correlations between the squares vanishes:

$$\sum_{I,\,J} (E\, W_I^2 W_J^2 - \sigma_I^2 \sigma_J^2) \rightarrow 0, \quad n \rightarrow \infty.$$

The final two sections of Ch. 2 contain the proof of the Th. 2.2.3. We shall not go into the details of the proof. We make one remark on the sort of result that is obtained in these sections. We obtain for fixed n a bound on the distance

$$\sup_x \; | P\{W(n) \leq x\} - P\{Y \leq x\} |,$$

with Y a standard normal random variable. This bound can be expressed (but for one universal constant C_1) in the parameters in which the central limit theorem is formulated: D, $|E\, W(n)^4 - 3|$, $\max_i \sum_{I \ni i} \sigma_I$ and $|\sum_{I,\,J} (E\, W_I^2 W_J^2 - \sigma_I^2 \sigma_J^2)|$. Since this bound does not seem to have any practical value, we formulate the results in terms of convergence of distributions.

Ch. 3 starts with some preliminary results with a technical flavour (Sect. 3.1). From Sect. 3.2 on we restrict ourselves to homogeneous sums in the Hoeffding decomposition. The main aim of Sect. 3.2 can be formulated as follows. Recall the central limit theorem for quadratic forms in Gaussian random variables: A sharp criterion for asymptotic normality of the bilinear form is given in terms of the eigenvalues of the symmetric matrix. The bilinear form $\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is regarded as a mapping $\mathbb{R}^n \to \mathbb{R}^n$, the matrix (a_{ij}). (Here a short detour is needed: The bilinear form

$$Z(n) = \sum_{1 \le i \le n} \sum_{1 \le j \le m} a_{ij} X_i Y_j ,$$

with $X_1,\ldots,X_n$, $Y_1,\ldots,Y_m$ independent, is included in the present setup: Consider the $(n + m) \times (n + m)$ matrix (b_{ij}) with all entries zero except those in the upper rectangle $1 \le i \le n$, $n < j \le m + n$, and $b_{ij} = a_{ij-n}$. Thus the bilinear form reduces to the cases treated above.) In the same way as the bilinear form, the d-linear form $\mathbb{R}^n \times \ldots \times \mathbb{R}^n \to \mathbb{R}$ can be regarded as a mapping $\mathbb{R}^{n^e} \to \mathbb{R}^{n^{d-e}}$ for $e = 1,\ldots,[d/2]$. This gives $[d/2]$ (rectangular) matrices. Instead of the eigenvalue decomposition we now use the singular value decomposition. (For real symmetric matrices the singular values equal the absolute values of the eigenvalues.) We arrive at results analogous to those in the case $d = 2$: The d-linear form satisfying the usual conditions of negligibility and with uniformly bounded fourth moments $E\,(W_I / \sigma_I)^4$ has a normal limit distribution iff the maximal singular value vanishes. These results, involving singular values, cannot be extended in full generality to the general case of d-homogeneous sums in the Hoeffding decomposition. However, in De Jong (1987) some partial results (for the case $d = 2$) are obtained. These results are extended for general d in Sect. 3.2.

In Sect. 3.3 inhomogeneous sums are treated. All results formulated until here concern homogeneous sums (except the counter-example which shows that homogeneity is essential in Th. 2.1.1). However, inhomogeneous sums arise in many interesting situations (cf. Hall (1984)). Consider the finite sum of homogeneous sums

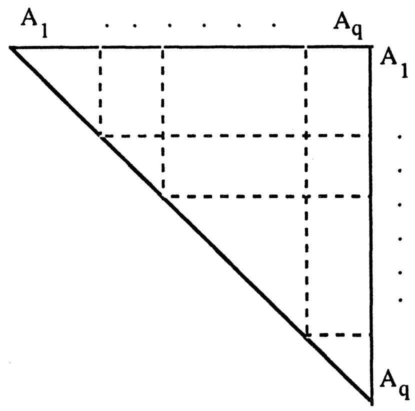
$$V(n) = W^{(1)}(n) + \ldots + W^{(d)}(n), \text{ with } \operatorname{var} W^{(e)}(n) \to \sigma^2(e) > 0, \; n \to \infty,$$

and $\operatorname{var} V(n) = 1$. If $W^{(e)}(n) / \operatorname{var}^{1/2} W^{(e)}(n)$ satisfies the conditions of Th. 2.1.1, then $V(n)$ has a normal limit distribution. Moreover, the joint distribution of $(W^{(1)}(n), \ldots, W^{(d)}(n))$ tends to a d-variate normal distribution with vanishing covariances.

The chapter ends with an elaborate example of a simple multilinear form in zero-one valued random variables, which is an inhomogeneous sum. This example is used to test the merits of some of the previously obtained results.

In Ch. 4 we start from the following observation (prompted by a question of A.A. Balkema): Consider the matrix (W_{ij}) of components in the Hoeffding decomposition $(d = 2)$. Let $A_1,\ldots,A_q$ be a partition of the integers $1,\ldots,n$ (q not depending on n). This partition induces a partition of (W_{ij}) with $\binom{q+1}{2}$ elements. This is illustrated in Figure 1. (Without loss of generality we may assume the elements of the partition to be consecutive blocks.)

Figure 1



If the homogeneous sums W(n) satisfy the conditions of Th. 2.1.1, then the joint distribution of the $\binom{q+1}{2}$ partial sums tends to a multivariate normal distribution with orthogonal components, provided the variance of each partial sum converges. This is the basic result of Ch. 4. By straightforward approximation it can be extended in the following way. Let $x_1, x_2, \ldots$ be points in R and embed the random variables W_{ij} as random point masses in $\mathbb{R}^2$ at (x_i, x_j), with $x_i < x_j$ if $i < j$. Define the discrete measures μ_n on $\mathbb{R}^2$ with mass σ_{ij}^2 at point (x_i, x_j). Suppose that the probability measures μ_n converge weakly to μ. Define the stochastic integral

$$\int f \, dW(n) = \sum_{1 \le i < j \le n} f(x_i, x_j) W_{ij}.$$

If the sums W(n) satisfy the conditions of Th. 2.1.1, then $\int f \, dW(n)$ has a normal limit distribution $N(0, \int f^2 d\mu)$, if the function f is bounded and μ-a.e. continuous. It is remarkable that the same result is obtained as would have been obtained under the (stronger) assumption that the random variables W_{ij} are independent. However there is an important difference: if the random variables W_{ij} are all independent they can be embedded in $\mathbb{R}$ (instead of $\mathbb{R}^2$) and the same result holds. In Ch. 4 it is shown that this is not the case for W_{ij} components in the Hoeffding decomposition. For these random variables the special (coordinatewise) embedding is important. Sect. 4.3 is concerned with this aspect of the Hoeffding decomposition. Rather, a criterion is given such that homogeneous sums in the Hoeffding decomposition satisfying this criterion can be embedded in $\mathbb{R}$ (instead of $\mathbb{R}^2$).

This text is meant to be selfcontained as far as it is concerned with generalized multilinear forms. Except for general results which can be found in textbooks like Chung (1974) and a central limit theorem for martingale differences, no results from probability theory are needed to read the text. (E.g. properties of martingales which are used without reference can be found in Chung (1974).)

We conclude this chapter with some references to related results. There are many papers on central limit theorems for generalized multilinear forms scattered throughout the literature. We shall not try to be exhaustive here. Instead, we shall give a rough classification according to methods of proof used to derive these results and provide a few references. We distinguish four approaches.

The first one applies to proper multilinear forms in independent random variables. In Rotar' (1973) it is shown that the limit distribution of a quadratic form in iid random variables with zero mean and unit variance does not depend on the actual distribution of the random variables. More generally, in this approach invariance classes of distributions are identified. For each invariance class the limit distribution is the same for any distribution in this class. Then the limit distribution can be determined with the help of one member in the class for which the limit can be computed. The limit behaviour of the quadratic form in independent N(0,1) random variables is treated exhaustively in Sevast'yanov (1961). In Rotar' (1979) invariance classes are given for multilinear forms.

The second approach we distinguish is the method of projection, given above. In this situation a central limit theorem can be obtained by methods from martingale theory. In Beran (1972) a central limit theorem for quadratic forms is proved with the help of a martingale method, a result which is related to that in Whittle (1964). A special class of generalized multilinear forms are U-statistics (d = 2):

$$Z(n) = \sum_{1\le i<j\le n} z_n(X_i, X_j),$$

with X_i iid and z_n a symmetric Borel function not depending on the indices i,j. Weber (1983) proves a central limit theorem using a technique based on backward martingales. If Z(n) is a homogeneous sum in the Hoeffding decomposition, the U-statistic is said to be degenerate. This case is treated in Hall (1984). The method used by Hall is a generalization of that in Beran (1972) and is essentially the same as the one used in De Jong (1987). In the author's Master's thesis (1982) this method was used to obtain central limit theorems for components in the Hoeffding decomposition. Backward martingales are also applied in case of weakly exchangeable arrays in Weber (1980).

The third method is based on a result in Stein (1970). It is used in Barbour and Eagleson (1985) to derive a central limit theorem for dissociated random variables. This concept was defined in McGinley and Sibson (1975). (A definition is given in Sect. 2.2; components in the Hoeffding decomposition are an example of dissociated random variables.) Earlier a central limit theorem for dissociated random variables was obtained in Noether (1970). Both Barbour and Eagleson (1985) and Noether (1970) treat cases of random variables with $d \geq 3$ indices.

The latter paper used the method of moments; the fourth approach. The method of moments can be applied in a great variety of situations. Although usually strong assumptions are imposed on the moments of the random variables, it is often possible to obtain results in situations where at first sight there is a very intractable dependence between the random variables. As an example may serve the monograph Bloemena (1964), where by means of the method of moments a central limit theorem is proved for quadratic forms in dependent random variables. (These quadratic forms are not sums of dissociated random variables.) When kth moments are calculated of sums of random variables indexed by pairs of indices, graph theory can give a heuristically useful description of complicated products. In particular it can be used to describe the way different random variables have indices in common. The description of higher moments by means of graph theory was already employed in Moran (1948). Later on it was used by several authors, e.g. Bloemena (1964), Kester (1975), Brown and Kildea (1978), and Jammalamadaka and Janson (1986). It was also used in De Jong (1982), where the method of moments was used to obtain central limit theorems for dissociated random variables. Contact with this approach seems to have long after-effects: In the next chapter several ways are introduced to describe higher moments, which are adaptations of the graph theory techniques, adapted for random variables indexed by more than two indices.

2. A central limit theorem for clean random variables

2.0. Introduction

In Ch. 1 the first two sections of this chapter have been introduced extensively. Here we shall make some general remarks on the proof of Th. 2.2.3 which is contained in the final two sections. The proof rests on a martingale central limit theorem (Heyde and Brown (1970)). By this theorem we have asymptotic normality for the sum of martingale differences $\sum_{1\le k\le n} U_k$, with $\sum_{1\le k\le n} E\, U_k^2 = 1$ and $\max_{1\le k\le n} E\, U_k^2 \to 0$, if

1) $\sum_{1\le k\le n} E\, U_k^4 \to 0, \quad n \to \infty,$

2) $\sum_{1\le k\le n} U_k^2 \xrightarrow{L^2} 1, \quad n \to \infty.$

These requirements can be relaxed. The fourth moment in 1) can be replaced by the $(2 + \varepsilon)$th moment; the L^2 convergence in 2) can be replaced by convergence in probability. However, the above formulation is very suitable in the present situation, since we are working with fourth moments.

If we write W(n) as a sum of martingale differences, we break the symmetry on W(n): All conditions on W(n) are invariant under permutation of the indices, whereas for a martingale the order of the index set plays an essential role. In Sect. 2.3 and 2.4 the fourth moment $E\, W(n)^4$ is split into partial sums over the quadruples (I,J,K,L) of d-point subsets of $\{1,\dots,n\}$. In Prop. 2.3.1 - 2.3.3 it is shown that the two requirements above are satisfied if certain partial sums vanish. Prop. 2.3.3. deals with the asymmetric character of the above requirements. In Prop. 2.3.4 these results are summarized in a technical central limit theorem, phrased in terms of these partial sums. In the remainder of this chapter it is shown that under the conditions of Th. 2.2.3 these partial sums vanish.

2.1. A central limit theorem for components in the Hoeffding decomposition

In the previous chapter the Hoeffding decomposition was introduced. On the probability space $(\Omega,\mathcal{F},P)$ a sequence of independent random variables $X_1, X_2, \dots$ is given. Define for finite sets of the integers I the σ-algebras $\mathcal{F}_I = \sigma\{X_i : i \in I\}$ and $\mathcal{F}_\varnothing$

$= \sigma\{\emptyset,\Omega\}$. Any square integrable $\mathcal{F}_{\{1,...,n\}}$-measurable random variable Z can be written $Z = \sum_{I \subset \{1,...,n\}} W_I$, where the components W_I are uniquely determined by

(2.1.1) a) W_I is $\mathcal{F}_I$-measurable,

b) $E(W_I \mid \mathcal{F}_J) = 0$ a.s. if $I \setminus J \neq \emptyset$.

For *d-homogeneous sums* in the Hoeffding decomposition $W(n) = \sum_{|I| = d} W_I$ - we shall reserve the notation W(n) for homogeneous sums - we have the following central limit theorem which follows from Th. 2.2.3 of Sect. 2.2.

Theorem 2.1.1. Let W(1),W(2),... be d-homogeneous sums in the Hoeffding decomposition, $W(n) = \sum_{|I| = d} W_I$, for fixed d with var W(n) = 1, for n = 1,2,... . Suppose

a) $\max_i \sum_{I \ni i} \sigma_I^2 \to 0$ for $n \to \infty$,

b) $\max_I E\, W_I^4 / \sigma_I^4 \leq D$, D not depending on n,

c) $E\, W(n)^4 \to 3$ for $n \to \infty$.

Then

$$W(n) \xrightarrow{d} N(0,1) \text{ for } n \to \infty.$$

This is Th.2.2 in De Jong (1987), for general d instead of d = 2. We give some comments on the conditions.

Condition a) excludes degenerating forms with the masses σ_I^2 concentrating on one or a few 'hyperplanes' $\{ I : i \in I\}$. These forms usually have a limit distribution depending on one or a few random variables X_i. By condition a) the following example is excluded. (If we consider random variables W_I indexed by sets I containing one or two elements, we employ the usual notation: W_i instead of $W_{\{i\}}$ and W_{ij} instead of $W_{\{i,j\}}$.)

Example 1. Let X_i be iid, $E\, X_i = 0$, $E\, X_i^2 = 1$ and

$$W(n) = (n-1)^{-1/2} (X_1X_2 + \quad + X_1X_n)$$

Then W(n) has a normal limit distribution iff $X_i = \pm 1$ with probability equaling1/2. The if part follows from the central limit theorem for sums of iid random variables with diverging total variance. The only if part can be deduced from the characteristic function of W(n). We shall return to this example below.

Condition b) is imposed to exclude random variables with heavy tails. In Ch. 3 we shall return to this issue and show how condition b) can be relaxed in two respects. In the first place in Th. 3.1.2 it is shown that we can allow D to diverge in a controlled way. In Sect. 3.2 it is shown how condition b) can be relaxed by means of truncation. If d = 2, then condition b) can be dropped (De Jong (1987)).

Condition b) can be replaced by the somewhat weaker condition

b') $\sum_{|I| = d} (E\, W_I^4)^{1/2} \leq C$, C not depending on n.

If condition b') holds we can apply Th. 2.1.1 to the partial sum

$$W'(n) = \sum_{\mathcal{A}} W_I,$$

with $\mathcal{A} = \{I : E\,(W_I / \sigma_I)^4 \leq D\}$ for some $D \geq 1$. Then

$$\operatorname{var}(W(n) - W'(n)) = \sum_{\mathcal{A}^c} \sigma_I^2 \leq C\, D^{-1/2},$$

by Chebyshev's inequality. However, since condition b) is clearly the condition which is needed in the proof of Th. 2.1.1 (and Th. 2.2.3), we shall use the more restrictive condition. The reader is free to adept the theorems and their proofs to this refinement.

Condition c) may be difficult to check; in Ch. 3 several conditions are given to replace condition c). These conditions are usually more restrictive than condition c), which is sharp in some sense: see Th. 2.2.4 and Th. 3.2.5.

Condition c) may be replaced by the weaker condition

c') $\limsup_{n \to \infty} E\, W(n)^4 \leq 3$.

In Sect. 2.3 it will be shown that under the conditions a) and b) of Th. 2.1.1 condition c') implies condition c).

The example below shows that the assumption that the random variables W(n) are homogeneous sums cannot be dropped. In Sect. 3.3 we shall give a central limit theorem for inhomogeneous sums.

Example 2. We start with the construction of a family of random variables with fourth moment equal to 3. Then we give a sequence of inhomogeneous sums in the Hoeffding decomposition with a non-normal limit distribution which satisfies the conditions a) and b) of Th. 2.1.1 and converges to a member in this family.

Let Y,Z be random variables with $E\,Y = E\,Z = 0$, $E\,Y^2 = 1$, $E\,Z^2 = \sigma^2$, $E\,Y\,Z = 0$, $E\,Y^3 Z \neq 0$ and $E\,Y^4 = 3$. Set $V = Z + \alpha\, Y$, then

$$E\,V^2 = \alpha^2 + \sigma^2,$$

$$E V^4 - 3 E^2 V^2 = 4 \alpha^3 E Y^3 Z + 6 \alpha^2 E Y^2 Z^2 + 4 \alpha E Y Z^3 + E Z^4 - 6 \alpha^2 \sigma^2 - 3 \sigma^4.$$

This is a polynomial in α of degree 3 (since $E\, Y^3 Z \neq 0$) which has at least one real zero, say α_0. For this zero the normed fourth moment of $V = Z + \alpha_0 Y$ equals 3.

Now choose

$$Y_n = n^{-1/2} \sum_{1 \le i \le n} X_i \text{ and } Z_n = n^{-3/2} \sum_{1 \le i < j < k \le n} X_i X_j X_k,$$

with X_i iid N(0,1) random variables. Then Y_n and Z_n are homogeneous sums in the Hoeffding decomposition (of degree 1 and 3 respectively) and satisfy the conditions a) and b) of Th. 2.1.1. Straightforward calculation yields

$$Y_n^3 = 6 Z_n + 3 Y_n \, (n^{-1} \sum_{1 \le i \le n} X_i^2) - 2 n^{-3/2} \sum_{1 \le i \le n} X_i^3,$$

with

$$n^{-1} \sum_{1 \le i \le n} X_i^2 \to 1, \qquad n \to \infty,$$

$$n^{-3/2} \sum_{1 \le i \le n} X_i^3 \to 0, \qquad n \to \infty,$$

$$E\, Y_n^3 Z_n = 6 E Z_n^2 \neq 0.$$

Since Y_n is N(0,1) distributed, we have

$$Z_n + \alpha\, Y_n \xrightarrow{d} Y^3/6 + Y(\alpha - 1/2), \quad n \to \infty,$$

with Y an N(0,1) distributed random variable. Thus we have constructed (with $V_n = (Z_n + \alpha_0 Y_n) / \mathrm{var}^{1/2} (Z_n + \alpha_0 Y_n)$) a sequence of random variables which satisfies the conditions a), b) and c) of Th. 2.1.1, but which has a non-normal limit distribution. (The tail of the limit distribution is determined by $Y^3/6$, which is not normal.)

2.2. Formulation of the main result

In this section the assumption is dropped that there is an underlying sequence of independent random variables. Then there is no Hoeffding decomposition, and for similar results conditions have to be imposed that are satisfied automatically in the Hoeffding decomposition. Two conditions are important here.

In the Hoeffding decomposition two random variables are independent if $I \cap J = \emptyset$. Indeed, components in the Hoeffding decomposition are dissociated (see McGinley and Sibson (1975)). That is, random variables, indexed by subsets of the integers $W_{I_1}, \ldots, W_{I_q}$ are independent if the sets $I_1, \ldots, I_q$ are mutually disjoint. Condition d) of Th. 2.2.3 is weaker than the assumption that the random variables are dissociated.

More important, however, is the definition of the analogue of (1.1.1). Consider the family of random variables indexed by finite subsets of the integers $\mathcal{W} = \{ W_I : I \subset \{1,2,...\} \}$ and the σ-algebras $\mathcal{F}^{(i)} = \sigma \{ W_I \in \mathcal{W} : i \notin I \}$.

Definition 2.2.1. The random variable $W_I \in \mathcal{W}$ is *clean* if $E(W_I \mid \mathcal{F}^{(i)}) = 0$ a.s. for all $i \in I$.

Notice that this definition is not the same as the definition of 'clean' in De Jong (1987), where components in the Hoeffding decomposition are considered. The family $\mathcal{W}$ is clean if all its members are clean; by abuse of language we say '$\sum_{\mathcal{W}} W_I$ is clean' to indicate that the family $\mathcal{W}$ is clean.

Any subset of a clean family is clean. If $\mathcal{W}$ is clean and $\mathcal{W}_0 \subset \mathcal{W}$, then $\mathcal{W}_0$ is clean, since with $\mathcal{F}_0^{(i)} = \sigma \{ W_I \in \mathcal{W}_0 : i \notin I\}$ we have for $W_I \in \mathcal{W}_0$

$$E(W_I \mid \mathcal{F}_0^{(i)}) = E(E(W_I \mid \mathcal{F}^{(i)}) \mid \mathcal{F}_0^{(i)}) = 0 \text{ a.s.}$$

Homogeneous sums in the Hoeffding decomposition (Sect. 2.1) are clean; since $\mathcal{F}^{(i)} \subset \sigma \{X_j : j \neq i \}$, we have for $i \in I$

$$E(W_I \mid \mathcal{F}^{(i)}) = E(E(W_I \mid X_j , j \neq i) \mid \mathcal{F}^{(i)}) = 0 \text{ a.s.}$$

Therefore, we have the following examples (cf. Ex. 2 of Sect. 2.1).

Example 1. The degenerate homogeneous sum in the Hoeffding decomposition $W(n) = W_{12} +...+ W_{1n}$ is clean. We can leave out the index 1. The set of random variables $\{W_j : W_j = W_{1j}, j = 2,...,n\}$ is clean, since

$$E(W_i \mid W_j , j \neq i) = E(E(W_i \mid X_j , j \neq i) \mid W_j , j \neq i) = 0 \text{ a.s.}$$

This can be extended easily.

Example 2. Consider the components in the Hoeffding decomposition $(W_{ij})_{1 \leq i \leq k,\, k<j \leq n}$. This is a rectangular part of the upper triangle of the matrix (W_{ij}). The set of random variables

$$\{ W_j := \sum_{1 \leq i \leq k} W_{ij} : j = k+1,...,n\}$$

is clean:

$$E(W_j \mid W_g , k < g \neq j) = E(E(W_j \mid X_g , k < g \neq j) \mid W_g , k < g \neq j)$$
$$= E (\sum_{1 \leq i \leq k} E(W_{ij} \mid X_g , k < g \neq j) \mid W_g , k < g \neq j) = 0 \text{ a.s.}$$

Both examples above can be generalized for components in the Hoeffding decomposition with d indices.

Notice that a random variable is clean with respect to a given set $\mathcal{W}$. A change in one variable may effect many other variables.

Example 3. Let $U_1,U_2,...$ be iid uniform (0,1) random variables and let $f_1,f_2,...$ be the Rademacher functions ($f_k(x) = 1 - 2\, e_k(x)$, with $e_k(x)$ right continuous "zero-one" and $\sum_k 2^{-k} e_k(x) = x$, $x \in [0,1]$, where $(e_k(x))_k$ are the coefficients in the binary expansion of x). Put $g_k = f_k + 1$; for $i < j$ we define $W_{ij} = f_j(U_i)g_i(U_j)$. Since $\sigma\{ W_{gh} : \{g,h\} \neq \{i,j\} \} \subset \sigma\{ U_g, f_h(U_i), f_k(U_j) : g,h,k \notin \{i,j\} \}$ and since for fixed i the random variables $f_k(U_i)$ are independent, the random variable W_{ij} is independent of the random variables W_{gh}, $\{g,h\} \neq \{i,j\}$. Thus $E(W_{ij} \mid W_{ik}, k \neq j) = E\, W_{ij} = 0$ a.s., whereas $E(W_{ij} \mid U_i) = f_j(U_i)Eg_i(U_j) \neq 0$ a.s. Hence the set $\{ W_{ij} : 1 \leq i < j \leq n \}$ is clean, and the set $\{ W_{ij} : 1 \leq i < j \leq n \} \cup \{ U_1 - 1/2 \}$ is not clean.

The above example shows that the union of two clean families (or the sum of two clean sums) is not necessarily clean.

Example 4. If $\{W_i : i = 1,2,...\}$ is clean, then the multilinear form

$$\sum_{|I| = d} W_I \text{, with } W_I = a_I \prod_{i \in I} W_i \quad ((a_I)_{|I| = d} \text{ real constants })$$

is clean, since with $i \in I$ we have

$$E(W_I \mid \mathcal{F}^{(i)}) = E(E(W_I \mid W_j, j \neq i) \mid \mathcal{F}^{(i)})$$

$$= a_I\, E ((\prod_{i \in I,\, k\neq i} W_k) E(W_i \mid W_j, i\neq j) \mid \mathcal{F}^{(i)})$$

$$= 0 \text{ a.s.}$$

Clean random variables are uncorrelated: With $i \in J \setminus I$ we have $E\, W_I W_J = E\, W_I\, E\, (W_J \mid \mathcal{F}^{(i)}) = 0$. This idea can be extended; we shall use it often in the following form.

Lemma 2.2.2. Let $\{ W_{I_1}, ..., W_{I_q} \}$ be clean and suppose $I_1 \cap (I_2 \cup ... \cup I_q) \neq I_1$ (I_1 is called a *free index* of the q- tuple $(I_1,...,I_q)$). Then $E\, W_{I_1} ... W_{I_q} = 0$ (provided the expectation exists).

Proof. Let $i \in I_1 / (I_2 \cup ... \cup I_q)$, then

$$E\, W_{I_1} ... W_{I_q} = E\, W_{I_2} ... W_{I_q}\, E\, (W_{I_1} \mid \mathcal{F}^{(i)}) = 0.$$

In fact we have shown more:

$$E\, (W_{I_1} ... W_{I_q} \mid W_J \in \mathcal{W}, J \in \mathcal{A}) = 0 \text{ a.s., if } I_1 \cap (I_2 \cup ... \cup I_q \cup (\cup_{\mathcal{A}} J)) \neq I_1.$$

For the sequence of clean finite sums W(n), homogeneous in $|I|$, with var W(n) = 1 - recall that we reserve the symbol W(n) for clean sums that are homogeneous in $|I|$ -

we have the following central limit theorem. Notice that each sum below is clean; nothing is assumed about W(k) + W(m), with $k \neq m$.

Theorem 2.2.3. Let W(1), W(2),... be a sequence of clean homogeneous sums of degree d, $W(n) = \sum_{|I| = d} W_I$ with var W(n) = 1, for n =1,2,... (d fixed). Suppose

a) $\max_i \sum_{I \ni i} \sigma_I^2 \to 0$ for $n \to \infty$,

b) $\max_I EW_I^4 / \sigma_I^4 \leq D$, D not depending on n,

c) $EW(n)^4 \to 3$ for $n \to \infty$,

d) $\sum_{I \cap J = \emptyset} (EW_I^2 W_J^2 - \sigma_I^2 \sigma_J^2) \to 0$ for $n \to \infty$.

Then

$$W(n) \xrightarrow{d} N(0,1) \text{ for } n \to \infty.$$

Remark 1. If we compare Th. 2.1.1 with the above theorem we can see that homogeneous sums in the Hoeffding decomposition are replaced by homogeneous sums of clean random variables satisfying assumption d). If the random variables W_I are dissociated, then we have $E\, W_I^2 W_J^2 = \sigma_I^2 \sigma_J^2$, if $I \cap J = \emptyset$. Thus assumption d) is implied by dissociated.

Remark 2. We shall see below (Ch. 3) that, under the assumptions a) and b), assumption d) is equivalent to

d') $\sum_{|I| = d} W_I^2 \xrightarrow{L^2} 1$ for $n \to \infty$.

Remark 3. If d=1, the clean sum has the martingale property and the reversed martingale property simultaneously:

$$E\,(W_i \mid W_j\,, j < i\,, j > i\,) = 0 \text{ a.s.}$$

In Sect. 2.4 it will be shown (Prop. 2.4.3) that if var W(n) = 1 and $E\, W_I^6 / \sigma_I^6 \leq D$ for all terms in W(n), we have

$$E\, W(n)^6 \leq D\, C_d,$$

with C_d a constant only depending on d (not on n). This implies the following converse result (using Feller II (1971: 251, part e)):

Theorem 2.2.4. Let W(1), W(2),... be a sequence of clean homogeneous sums of degree d, $W(n) = \sum_{|I| = d} W_I$ with var W(n) = 1, for n =1,2,... (d fixed). Suppose

$$\max_I E W_I^6 / \sigma_I^6 \leq D, \text{ D not depending on n.}$$

Then

$$W(n) \xrightarrow{d} N(0,1) \text{ for } n \to \infty,$$

implies

$$EW(n)^4 \to 3 \text{ for } n \to \infty.$$

If d = 1 then the conditions of Th 2.2.3 can be relaxed somewhat.

Corollary 2.2.5. (Th. 2.2.3 in case d = 1.) Let W(1), W(2),... be a sequence of clean homogeneous sums $W(n) = \sum_{1 \leq i \leq n} W_i$ with var W(n) = 1, for n =1,2,... .
Suppose

a) $\max_i \sigma_i^2 \to 0$ for $n \to \infty$.

Then

$$W(n) \xrightarrow{d} N(0,1) \text{ for } n \to \infty,$$

if two of the following three conditions hold:

b) $\max_i EW_i^4 / \sigma_i^4 \leq D$, D not depending on n,

c) $E\,W(n)^4 \to 3$ for $n \to \infty$,

d) $\sum_{i \neq j} (E\,W_i^2 W_j^2 - \sigma_i^2 \sigma_j^2) \to 0$ for $n \to \infty$.

(For the proof see Sect. 2.3.)

As a consequence of corollary 2.2.5 we obtain the statement of Example 1 in Sect.2.1. The sum $(n-1)^{-1/2}(X_1X_2 + \ldots + X_1X_n)$, ($X_i$ iid, $E\,X_1 = 0$, $E\,X_1^2 = 1$) is homogeneous (d = 1) and clean and condition a) is satisfied. If the distribution of X_i $= \pm 1$ with probability equaling 1/2, then the conditions b) and d) are satisfied. If W(n) is N(0,1) distributed then the sixth moment $E\,X_1^6$ is bounded, thus condition b) is satisfied and, by Th. 2.2.4, we have condition c). The conditions b) and c) imply (see proof of corollary 2.2.5) condition d), which implies that the distribution of X_1 is as required.

Thus Th. 2.2.3 (combined with the trick of 'lowering the dimension d' as shown in the examples 1 and 2 above) can be applied in cases which are excluded by condition a) in Th. 2.1.1.

2.3. Proof of Theorem 2.2.3

In this section we shall consider a fixed finite clean sum W(n) . We can write W(n) as a sum of martingale differences

$$W(n) = \sum_{1\le k\le n} U_k \text{, with } U_k = \sum_{I,\ \max I = k} W_I ,$$

with respect to the σ- algebras $\mathcal{F}_k = \sigma\{W_I : \max I \le k\}$, since

$$E(U_k \mid \mathcal{F}_{k-1}) = \sum_{I,\ \max I = k} E(W_I \mid \mathcal{F}_{k-1})$$

$$= \sum_{I,\ \max I = k} E(E(W_I \mid \mathcal{F}^{(k)}) \mid \mathcal{F}_{k-1})$$

$$= 0 \quad \text{a.s. (as } \mathcal{F}_{k-1} \subset \mathcal{F}^{(k)} \text{).}$$

Remark 1. There is some arbitrariness in the definition of the martingale differences: Another ordering of the integers i,...,n, generally gives another set of martingale differences. In the definition of clean no ordering is assumed. There is even more structure. A martingale difference is a sum of martingale differences:

$$U_k = \sum_{1\le j\le k-1} V_{kj} \text{ ,with } V_{kj} = \sum_{I,\ \max I\setminus\{k\} = j} W_I ,$$

with respect to the σ-algebras $\mathcal{F}_{kj} = \sigma\{W_I : \max I = k,\ \max I\setminus\{k\} \le j\}$. Notice that $\mathcal{F}_{kj-1} \subset \mathcal{F}^{(j)}$. This can be repeated d times: W(n) is a sum of sums ... of sums (d times) of martingale differences. This extra structure is not needed in what follows.

Notice that for the sequence W(n) with var W(n) = 1, for n = 1,2,... , lemma 2.2.2 yields

$$\sum_k E\, U_k^2 = \sum_k \sum_{\max I = k} \sigma_I^2 = \sum_{|I| = d} \sigma_I^2 = 1 \text{ and } \max_k E\, U_k^2 \le \max_k \sum_{I \ni k} \sigma_I^2,$$

which can be chosen arbitrarily small by assumption a). By Th. 1 in Heyde and Brown (1971) we have for the sum of martingale differences $\sum_k U_k$ (with $\sum_k E\, U_k^2 = 1$)

$$\sup_x | P\{ \sum_k U_k \le x \} - \Phi(x) | \le C_1 (\sum_k E\, U_k^4 + E(\sum_k U_k^2 - 1)^2) ,$$

with

$$\Phi(x) = (1/2\pi)^{-1/2} \int_{-\infty}^{x} e^{-t^2/2} dt,$$

and C_1 a constant not depending on n. We shall give estimates for $\sum E\, U_k^4$ and var $(\sum U_k^2)$, which vanish under the assumptions of Th. 2.2.3, thus proving the theorem. We start with proving Corollary 2.2.5.

Proof of Corollary 2.2.5. For d = 1 with $W(n) = \sum_k W_k$ and $\operatorname{var} W(n) = 1 = \sum_k \sigma_k^2$ we have the following two equalities

$$E\,W(n)^4 - 3 \overset{(1)}{=} \sum_k (E\,W_k^4 - 3\,\sigma_k^4) + 6 \sum_{k<l} (E\,W_k^2 W_l^2 - \sigma_k^2 \sigma_l^2)$$

$$\overset{(2)}{=} 3 \operatorname{var}(\sum_k W_k^2) - 2 \sum_k E\,W_k^4 .$$

Condition b) implies

$$\sum_k E\,W_k^4 \le D \sum_k \sigma_k^4 \le D \max_k \sigma_k^2 ,$$

which vanishes by assumption a). Thus equality (1) shows the equivalence of the conditions c) and d). And Corollary 2.2.5 follows by equality (2) (since $U_k = W_k$) and the martingale central limit theorem.

The conditions a), c) and d) together imply, by equality (1), $\sum E\,W_k^4 \to 0$, and thus, by equality (2), $\operatorname{var}(\sum W_k^2) \to 0$. This proves again Corollary 2.2.5 (without use of assumption b)). However, the conditions a), c) and d) together do not imply condition b). This ends the proof of Corollary 2.2.5.

The proof of Th. 2.2.3 for $d \ge 2$ is more involved; one reason being that the different partial sums that make up the fourth moment $E\,W(n)^4$ cannot be described explicitly (as in equality (1) above).The fourth moment

$$E\,W(n)^4 = \sum_{(I,J,K,L)} E\,W_I W_J W_K W_L$$

is split into three partial sums according to whether a quadruple (I,J,K,L) is in one of the three (disjoint) collections below:

$\mathcal{F}$ the collection of quadruples with a free index (see Lemma 2.2.2),

$\mathcal{B}$ the collection of quadruples (I,J,K,L) with each element in the union $I \cup J \cup K \cup L$ in exactly two of the sets I,J,K,L. This is the collection of *bifold* quadruples:

$$1_I + 1_J + 1_K + 1_L = 2\,1_{I \cup J \cup K \cup L} ,$$

$\mathcal{T}$ the rest $\mathcal{F}^c \setminus B$; a quadruple in T has no free index and at least one element in the union $I \cup J \cup K \cup L$ is in three or more sets:

$$1_I + 1_J + 1_K + 1_L \gneqq 2\,1_{I \cup J \cup K \cup L} .$$

In Lemma 2.2.2 it is shown that the set $\mathcal{F}$ nor its subsets contribute to the fourth moment E $W(n)^4$. For any subset $\mathcal{F}^* \subset \mathcal{F}$ we have

$$(2.3.1) \qquad \sum_{(I,J,K,L) \in \mathcal{F}^*} E\, W_I W_J W_K W_L = 0\,.$$

The quantities τ and τ^*, defined below, will play an important role in the next chapters:

$$\tau^* = \sum_{(I,J,K,L) \in \mathcal{T}} |E\, W_I W_J W_K W_L|,$$

$$\tau = \sum_{(I,J,K,L) \in \mathcal{T}} \sigma_I \sigma_J \sigma_K \sigma_L$$

The proof of Th. 2.2.3 is split into several propositions, some of which are quite easy to prove. In fact there are only three major problems to be overcome: in the first place the proof that τ and τ^* vanish under the conditions a) and b) of Th. 2.2.3; this is postponed until Sect. 2.4. Further it has to be shown that the bifold quadruples vanish except those quadruples that consist of two pairs of identical indices ($\sum_{I \cap J = \varnothing} W_I^2 W_J^2$). This is settled in Prop. 2.3.6 and Prop. 2.3.5. (If d = 1 this is evident; for $d \geq 2$ much attention has to be paid to these 'extra' bifold quadruples.) Finally it has to be shown that the conditions of Th. 1 in Heyde and Brown (1971) are satisfied. This is formulated in Prop. 2.3.4. The hardest part of the proof of this proposition rests on a symmetry argument (Prop. 2.3.3).

Proposition 2.3.1. For U_k, defined above, we have (with W_I clean)

$$\sum_k E\, U_k^4 \leq \tau^*\,.$$

Proof.

$$\sum_k E\, U_k^4$$

$$= \sum_k \sum_{(I,J,K,L),\ \max I = \max J = \max K = \max L = k} E\, W_I W_J W_K W_L$$

$$= \sum_{(I,J,K,L),\ \max I = \max J = \max K = \max L} E\, W_I W_J W_K W_L.$$

On the right-hand side no bifold quadruples occur, since $I \cap J \cap K \cap L \neq \varnothing$. The conclusion follows by (2.3.1) and the definition of τ^*.

In order to estimate var ($\sum U_k^2$), the collection of bifold quadruples is split again:

$\mathcal{B}(e,f) = \{ (I,J,K,L) \in B : | I \cap J | = e, | I \cap K | = f \}$.

Given the numbers $e = | I \cap J |$ and $f = | I \cap K |$, the number of elements in each intersection of two indices (other intersections are empty) is known. If (I,J,K,L) is bifold, then $| I \cap L | = d-e-f$, since I is the disjoint union $(I \cap J) \cup (I \cap K) \cup (I \cap L)$; and $| I \cap J | = | K \cap L |$, since $I \Delta J = K \Delta L$ and $|I| = |J| = |K| = |L| = d$, etc. Put

$$S(e,f) = \sum_{(I,J,K,L) \in \mathcal{B}(e,f)} E\, W_I W_J W_K W_L .$$

Since the value of $E\, W_I W_J W_K W_L$ is not changed by a permutation of (W_I, W_J, W_K, W_L), we have

$$S(e,f) = S(f,e) = S(e,d-e-f).$$

Put

$$S = \sum_{1 \le e \le d-2} \; \sum_{1 \le f \le d-e-1} S(e,f),$$

$$S_0 = \sum_{1 \le e \le d-1} S(e,0),$$

then we have

$$\sum_{(I,J,K,L) \in \mathcal{B}} E\, W_I W_J W_K W_L = S + 3\, S_0 + 3\, S(0,0).$$

The following quantity will be used frequently in the sequel:

$$\gamma = \Big| \sum_{I \cap J = \varnothing} (E\, W_I^2 W_J^2 - \sigma_I^2 \sigma_J^2) \Big|.$$

We have

(2.3.2) $$| S(0,0) - 1 |$$

$$= \Big| \sum_{I \cap J = \varnothing} E\, W_I^2 W_J^2 - \Big(\sum_{|I| = d} \sigma_I^2 \Big)^2 \Big|$$

$$\le \gamma + \sum_{I \cap J \ne \varnothing} \sigma_I^2 \sigma_J^2$$

$$\le \gamma + \tau .$$

We shall give an estimate for $\operatorname{var}(\sum U_k^2)$ in terms of τ^*, τ, S, S_0 and γ; see (2.3.3). We start with an auxiliary random variable.

Proposition 2.3.2. For clean random variables W_I we have

$$\operatorname{var}\Big(\sum_{I \cap J \ne \varnothing} W_I\, W_J \Big) \le \gamma + \tau + | S + 2\, S_0 | + \tau^*.$$

Proof. Since the random variables are clean, they are orthogonal

$$E\Big(\sum_{I \cap J \ne \varnothing} W_I\, W_J \Big) = \sum_{|I| = d} \sigma_I^2 = 1.$$

Since quadruples with a free index do not contribute by (2.3.1),

$$E\Big(\sum_{I\cap J\neq\emptyset} W_I\, W_J\Big)^2 = \sum_{I\cap J\neq\emptyset\neq K\cap L} E\, W_I W_J W_K W_L$$

$$= \sum_{\mathcal{B}, I\cap J\neq\emptyset} E\, W_I W_J W_K W_L + R$$

$$= \sum_{1\le e\le d}\ \sum_{1\le f\le d-e} S(e,f) + R$$

$$= S(d,0) + S + 2\, S_0 + R,$$

with

$$R = \sum_{\mathcal{T},\, I\cap J\neq\emptyset\neq K\cap L} E\, W_I W_J W_K W_L \text{ and thus } |R| \le \tau^*,$$

by using

$$\sum_{\mathcal{T}^*} |E\, W_I W_J W_K W_L| \le \sum_{\mathcal{T}} |E\, W_I W_J W_K W_L| = \tau^* \text{ if } \mathcal{T}^* \subset \mathcal{T}.$$

Now the proposition follows from (2.3.2) and the symmetry relation s(d,0) = S(0,0).

By the symmetry of the bifold quadruples we obtain:

Proposition 2.3.3. For U_k, defined above, we have (with W_I clean)

$$\mathrm{var}\Big(\sum_{I\cap J\neq\emptyset} W_I W_J - \sum_k U_k^2\Big) \le \tau^* + |\,2/3\, S + S_0\,|.$$

Proof. With R a partial sum over quadruples in $\mathcal{T}$ (and thus $|R| \le \tau^*$) we obtain

$$\mathrm{var}\Big(\sum_{I\cap J\neq\emptyset} W_I W_J - \sum_k U_k^2\Big)$$

$$= \mathrm{var}\Big(\sum_{I\cap J\neq\emptyset} W_I W_J - \sum_{\max I\cup J\,\in\, I\cap J} W_I W_J\Big)$$

$$= E\Big(\sum_{I\cap J\neq\emptyset,\ \max I\cup J\,\notin\, I\cap J} W_I W_J\Big)^2$$

$$= 4 \sum_{\mathcal{B}, I\cap J\neq\emptyset,\ \max I\cup J\,\in\, I\setminus J,\, \max K\cup L\,\in\, K\setminus L} E\, W_I W_J W_K W_L + R$$

$$= 4 \sum_{\mathcal{B}, I\cap J\neq\emptyset,\ \max I\cup J = \max K\cup L\,\in\, K\cap I} E\, W_I W_J W_K W_L + R$$

$$= 4 \sum_{\mathcal{B}, I\cap J\neq\emptyset,\ \max I\cup J\cup K\cup L\,\in\, K\cap I} E\, W_I W_J W_K W_L + R$$

$$= 4 \sum_{\mathcal{B}, I\cap J\neq\emptyset,\ \max I\cup J\cup K\cup L\,\in\, K\cap I,\ I\cap L=\emptyset} E\, W_I W_J W_K W_L$$

$$+4 \sum_{\mathcal{B}, I\cap J \neq \emptyset,\, \max I\cup J\cup K\cup L \in K\cap I,\, I\cap L \neq \emptyset} E\, W_I W_J W_K W_L + R$$

$$= S_0 + 2/3\, S + R,$$

where the last equality sign is explained below. The set

$$\bigcup_{1\le e\le d-1} \mathcal{B}(e,d-e) = \{(I,J,K,L) \in \mathcal{B} : I \cap J \neq \emptyset \neq I \cap K,\, I \cap L = \emptyset\}$$

can be partitioned into 4 subsets according to whether max $I\cup J\cup K\cup L$ is in one of the following intersections $I \cap J$, $I \cap K$, $K \cap L$, $J \cap L$. The sums over these subsets have equal contributions as can be seen from the equality

$$\sum_{I \cap J \neq \emptyset \neq I \cap K,\, I \cap L = \emptyset,\, \max I\cup J\cup K\cup L \in I \cap J} f(I,J,K,L)$$

$$= \sum_{I \cap J \neq \emptyset \neq I \cap K,\, I \cap L = \emptyset,\, \max I\cup J\cup K\cup L \in I \cap K} f(I,K,J,L),$$

with $f(I,J,K,L) = E\, W_I W_J W_K W_L$ and the commutativity of multiplication, $f(I,J,K,L) = f(I,K,J,L)$.This shows that

$$S_0 = 4 \sum_{I \cap J \neq \emptyset \neq I \cap K,\, I \cap L = \emptyset, \max I\cup J\cup K\cup L \in I \cap K} E\, W_I W_J W_K W_L.$$

By the same argument it is shown that

$$S = 6 \sum_{I \cap J \neq \emptyset,\, I \cap K \neq \emptyset,\, I \cap L \neq \emptyset, \max I\cup J\cup K\cup L \in I \cap K} E\, W_I W_J W_K W_L\,.$$

This proves the proposition.

The three propositions above together imply the following technical central limit theorem.

Proposition 2.3.4. Let W(1), W(2),... be a sequence of clean homogeneous sums of degree d, $W(n) = \sum_{|I| = d} W_I$ with var W(n) = 1. If all the following conditions hold

- I a) $\tau \to 0$ for $n \to \infty$,
- b) $\tau^* \to 0$ for $n \to \infty$,
- II a) $S_0 \to 0$ for $n \to \infty$,
- b) $S \to 0$ for $n \to \infty$,
- III $\gamma \to 0$ for $n \to \infty$,

then

$$W(n) \xrightarrow{d} N(0,1) \text{ for } n \to \infty.$$

Proof. Combining Prop. 2.3.2 and 2.3.3 we obtain, by $\mathrm{var}(A+B) \le 2\,\mathrm{var}\, A + 2\,\mathrm{var}\, B$, (2.3.3) $\quad \mathrm{var}(\sum_k U_k^2) \le 2\,\mathrm{var}(\sum_{I \cap J \neq \emptyset} W_I W_J) + 2\,\mathrm{var}(\sum_{I \cap J \neq \emptyset} W_I W_J - \sum_k U_k^2)$

$$\le 4\,\tau^* + 2\,\tau + 2\,\gamma + 10/3\,|S| + 6\,|S_0|.$$

The proposition follows by Th. 1 in Brown and Heyde (1971) and Prop. 2.3.1.

To prove Th. 2.2.3 we shall check the conditions I, II and III. Assumption d) of Th. 2.2.3 is equivalent to condition III. Under the assumption b) of Th. 2.2.3 we can reduce condition Ib) to condition Ia). By the Hölder inequality we have

(2.3.4) $\quad E\,|W_I W_J W_K W_L| \le E^{1/4} W_I^4\, E^{1/4} W_J^4\, E^{1/4} W_K^4\, E^{1/4} W_L^4$

$$\le D\, \sigma_I\, \sigma_J\, \sigma_K\, \sigma_L.$$

This shows $\tau^* \le D\,\tau$. The proof of condition Ia) is postponed to Sect. 2.4.

Condition II will follow from the two propositions below.

Proposition 2.3.5. For clean homogeneous sums W(n) we have

$$|S + 3\,S_0| \le \tau^* + 3\,\tau + 3\,\gamma + |E\,W(n)^4 - 3|.$$

Proof. The fourth moment $E\,W(n)^4$ can be written as:

$$E\,W(n)^4 = \sum_{\mathcal{T}} E\; W_I W_J W_K W_L + \sum_{\mathcal{B}} E\; W_I W_J W_K W_L.$$

Thus

$$E\,W(n)^4 - 3 = \sum_{\mathcal{T}} E\; W_I W_J W_K W_L + S + 3\,S_0 + 3\,S(0,0) - 3.$$

The proposition follows by the triangle inequality and (2.3.2).

The right-hand side in Prop. 2.3.5 vanishes under the conditions of Th. 2.2.3. However, we have to show that S and S_0 vanish separately. Here is a lower bound for $S_0 = S(1,0) + \ldots + S(d-1,0)$ and for $S + 2\,S_0$.

Proposition 2.3.6. For clean homogeneous sums W(n) we have

a) $S(e,o) \ge -\tau^*$ for $1 \le e \le d-1$,

b) $S + 2\,S_0 \ge -\tau^*$.

Proof. We shall show that both left-hand sides are a sum of squares up to a remainder term, which is a sum over a subset of $\mathcal{T}$. Consider two disjoint sets of the integers both of size d - e: $A, A' \subset \{1,2,\ldots\}$ with $A \cap A' = \emptyset$, $|A| = |A'| = d - e$. Then

$$E\Big(\sum_{|I\cap J| = e,\ I\setminus J = A,\ J\setminus I = A'} W_I W_J\Big)^2$$

$$= \sum_{|I\cap J| = e = |K\cap L|,\ I\setminus J = A = L\setminus K,\ J\setminus I = A' = K\setminus L} E\, W_I W_J W_K W_L$$

$$= \sum_{\mathcal{T},\ |I\cap J| = e = |K\cap L|,\ I\setminus J = A = L\setminus K,\ J\setminus I = A' = K\setminus L} E\, W_I W_J W_K W_L$$

$$+ \sum_{\mathcal{B}(e,o),\ I\cap L = A,\ J\cap K = A'} E\, W_I W_J W_K W_L.$$

Summation over the subsets A, A' yields

$$0 \le \sum_{A,\ A' \subset \{1,\dots,n\},\ A\cap A' = \emptyset,\ |A| = |A'| = d-e} E\Big(\sum_{|I\cap J| = e,\ I\setminus J = A,\ J\setminus I = A'} W_I W_J\Big)^2$$

$$(2.3.5)\quad = \sum_{\mathcal{T},\ |I\cap J| = e = |K\cap L|,\ I\setminus J = L\setminus K,\ J\setminus I = K\setminus L} E\, W_I W_J W_K W_L + S(e,0)$$

$$= R_1 + S(e,0),$$

with $|R_1| \le \tau^*$, since R_1 is a sum over a subset of $\mathcal{T}$. This proves a).
The second inequality follows from

$$E\Big(\sum_{I\cap J \ne \emptyset,\ I \ne J} W_I W_J\Big)^2$$

$$= \sum_{\mathcal{B},\ I\cap J \ne \emptyset,\ I \ne J} E\, W_I W_J W_K W_L$$

$$+ \sum_{\mathcal{T},\ I\cap J \ne \emptyset \ne K\cap L,\ I \ne J,\ K \ne L} E\, W_I W_J W_K W_L$$

$$= \sum_{1\le e\le d-1}\ \sum_{0\le f\le d-e} S(e,f) + R_2$$

$$= S + 2\, S_0 + R_2,$$

with $|R_2| \le \tau^*$. This proves the proposition.

By the above proposition we have

$$(2.3.6)\quad -(d-1)\,\tau^* \le S_0 = S + 3\, S_0 - (S + 2\, S_0) \le |S + 3\, S_0| + \tau^*.$$

Hence S_0 and S vanish if $S + 3\, S_0$ and τ^* vanish. This completes the proof of Th. 2.2.3, except for estimates of the quantity τ.

Remark. In the next section it will be shown that τ and τ^* vanish under the conditions a) and b). Thus γ, τ, τ^*, and (S(0,0) - 1) all vanish under the conditions a), b) and d) of Th. 2.2.3. Hence Prop. 2.3.6 implies that, under these conditions,

$$\liminf_{n \to \infty} E\, W(n)^4 \geq 3.$$

2.4. Estimating the quantities τ and τ^*

The main result of this section consists of estimates for sums over quadruples (I,J,K,L), containing no free indices. The conditions we impose here on the random variables are minimal: $E\, W_I = 0$, $E\, W_I^2 = \sigma_I^2$ and $E\, W_I^4 / \sigma_I^4 \leq D$ for all I. Indeed we shall consider the general case of q-tuples $E\, W_{I_1} \ldots W_{I_q}$ rather than quadruples, where the indices are finite subsets of the integers. Consider the family of q-tuples $Q = \{ (I_1,\ldots,I_q) : I_g \subset \{1,\ldots,n\},\ |I_g| \leq d,\ g = 1,\ldots,q\}$ for fixed d (and n large). Notice that the assumption of homogeneity in $|I_g|$ is dropped. Here also q-tuples $(I_1,\ldots,I_q)$ with indices I_g of different cardinality are taken into account.

In Sect. 2.3 the set of homogeneous quadruples was split into the subsets $\mathcal{B}(e,f)$. We shall now split the set Q in a different way , in which the ordering of the underlying set $\{1,\ldots,n\}$ plays an important role.

Definition 2.4.1. The *shadow* of a q-tuple $(I_1,\ldots,I_q) \in Q$ with $I_1 \cup \ldots \cup I_q = \{i_1,\ldots,i_f : i_1 < \ldots < i_f\}$ is the q-tuple $(I'_1,\ldots,I'_q) \in Q$ defined by $I'_g = \{j : i_j \in I_g\}$, $g = 1, \ldots, q$.

Since the shadow of a q-tuple $(I_1,\ldots,I_q)$ is determined by q subsets (of at most d elements) of a set containing at most dq elements the number of distinct shadows with q elements with cardinality $\leq d$ is bounded by $(\binom{qd}{1} + \ldots + \binom{qd}{d})^q \leq (2qd)^{qd}$. Although this bound may not be sharp, it does not depend on n. Since the number of distinct shadows (q,d fixed) does not depend on n it is for purpose of estimation, sufficient to evaluate the sum over all q- tuples with the same shadow $(I'_1,\ldots,I'_q)$. This amounts to sums over all ordered f-tuples $n_1 < \ldots < n_f$ $(f = |I_1 \cup \ldots \cup I_q|)$. It is easier to work in a product space. Therefore we shall sum over all f-tuples $(n_1,\ldots,n_f)$, i.e. integrate over $\mathbb{N}^f$ with respect to the counting measure. The basics from integration theory are sufficient for our goal.

Let F be a finite set (e.g. a subset of the integers) and φ a non-negative measurable function on $\mathbb{R}^F$, $\varphi : \mathbb{R}^F \to [0,\infty)$. With λ_F the Lebesgue measure on $\mathbb{R}^F$ we have by Fubini's theorem on the rearrangement of the integration order

$$\int_{\mathbb{R}^F} \varphi \, d\lambda_F = \int_{\mathbb{R}^{F_q}} \cdots \int_{\mathbb{R}^{F_1}} \varphi \, d\lambda_{F_q} \cdots d\lambda_{F_1},$$

for a partition $(F_1,\ldots,F_q)$ of F into non-empty subsets and λ_{F_g} the Lebesgue measure on $\mathbb{R}^{F_g}$. The theorem also holds trivially for the counting measure μ_F on $\mathbb{N}^F$. We shall use the shorthand notation:

$$\int^A \xi \, d\mu_A := \int_{\mathbb{N}^A} \xi \, d\mu_A$$

for $\xi : \mathbb{N}^F \to [0,\infty)$, $A \subset F$. If $A \neq F$ then $\int^A \xi \, d\mu_A$ is a non-negative function on $\mathbb{N}^{F\setminus A}$ which may be infinite at certain points. By definition we extend this notation:

$$\int^{\varnothing} \xi \, d\mu_{\varnothing} := \xi.$$

Put

$$\| \xi \|_2^2 = \int^F \xi^2 \, d\mu_F,$$

$$\rho(\xi) = \max_{j \in F} \sup_{\mathbb{N}} \int^{F\setminus\{j\}} \xi^2 \, d\mu_{F\setminus\{j\}},$$

$\rho(\xi)$ is the supremum of the integral of ξ^2 over any hyperplane parallel to some coordinate. Notice that $\rho(\xi) \leq \| \xi \|_2^2$, since a sum of non-negative terms dominates all partial sums.

The next lemma (essentially an application of the Cauchy-Schwarz inequality) contains the basic results of this section.

Lemma 2.4.2. Let F_g be finite non-empty sets of the integers and $\xi_g : \mathbb{N}^{F_g} \to [0,\infty)$, $g = 1,\ldots,q$. Suppose $A \subset F = F_1 \cup \ldots \cup F_q$, and

$$1_{F\setminus A} (1_{F_1} + \ldots + 1_{F_q}) \geq 2 \, 1_{F\setminus A}$$

(any element in $F \setminus A$ is contained in at least two subsets F_g), then

a) $\int^A \left(\int^{F\setminus A} \xi_1 \cdots \xi_q \, d\mu_{F\setminus A} \right)^2 d\mu_A \leq \| \xi_1 \|_2^2 \cdots \| \xi_q \|_2^2.$

b) If $F_1 \cap A \cap (F_2 \cup \ldots \cup F_q) \neq \varnothing$, then
$\int^A \left(\int^{F\setminus A} \xi_1 \cdots \xi_q \, d\mu_{F\setminus A} \right)^2 d\mu_A \leq \rho(\xi_1) \| \xi_2 \|_2^2 \cdots \| \xi_q \|_2^2.$

Proof. The proof proceeds by induction on q. For q = 1 both assertions are trivial: since $F_1 = F = A$ we have $\int^F \xi^2 \, d\mu_F = \| \xi_1 \|_2^2$; hence a). If q = 1, then condition b) is empty.

Assume a) holds for q - 1 functions $\xi_2,\ldots,\xi_q$. Put $\xi = \xi_2 \cdots \xi_q$. The set F is divided into 6 disjoint subsets R_i, i = 1,..., 6, according to the scheme below.

$$\begin{array}{ccl} F/A & A & \\ R_1 & R_4 & \left.\right\} F_1 \\ R_2 & R_5 & \\ R_3 & R_6 & \left.\right\} F_2 \cup \dots \cup F_q \end{array}$$

Notice that $R_1 = \emptyset$ and $2\,1_{R_3} \le 1_{R_3}(1_{F_2} + \dots + 1_{F_q})$. By rearrangement of the integration order, application of the Cauchy-Schwarz inequality and again by rearrangement of the integration order, we obtain

$$\int^{A} \Big(\int^{F\setminus A} \xi_1 \cdots \xi_q \, d\mu_{F\setminus A} \Big)^2 d\mu_A$$

$$= \int^{A} \Big(\int^{R_2} (\xi_1)\Big(\int^{R_3} \xi \, d\mu_{R_3} \Big) d\mu_{R_2}\Big)^2 d\mu_A$$

$$\le \int^{A} \Big(\int^{R_2} \xi_1^2 \, d\mu_{R_2}\Big) \Big(\int^{R_2}\Big(\int^{R_3} \xi \, d\mu_{R_3} \Big)^2 d\mu_{R_2}\Big) d\mu_A$$

$$= \int^{R_5} \Big(\int^{R_2\cup R_4} \xi_1^2 \, d\mu_{R_2\cup R_4}\Big) \Big(\int^{R_6\cup R_2} \Big(\int^{R_3} \xi \, d\mu_{R_3}\Big)^2 d\mu_{R_2\cup R_6}\Big)\Big) d\mu_{R_5}$$

$$= \int^{A'} \Big(\int^{F_1\setminus R_5} \xi_1^2 \, d\mu_{F_1\setminus R_5}\Big) \Big(\int^{F'\setminus A'} \xi \, d\mu_{F'\setminus A'}\Big)^2 d\mu_{A'} ,$$

with $F' = F_2\cup\dots\cup F_q$ and $A' = (A\cup F_1) \cap F' = R_2\cup R_5\cup R_6$. Since $F'\setminus A' = R_3$, we have $1_{F'\setminus A'}(1_{F_2} + \dots + 1_{F_q}) \ge 2\,1_{F'\setminus A'}$ and thus by the induction hypothesis

$$\int^{A'} \Big(\int^{F'\setminus A'} \xi \, d\mu_{F'\setminus A'}\Big)^2 d\mu_{A'} \le \|\xi_2\|_2^2 \cdots \|\xi_q\|_2^2.$$

If $F_1 \cap A \cap (F_2\cup\dots\cup F_q) = R_5 \neq \emptyset$, then

$$\int^{F_1\setminus R_5} \xi_1^2 \, d\mu_{F_1\setminus R_5} \le \rho(\xi_1),$$

where the last inequality rests on a simple property of the counting measure: $a_i \ge 0$ implies $\sup_i a_i \le \sum_i a_i$. This, together with the induction hypothesis, proves b).

If $R_5 = \emptyset$, then

$$\int^{F_1 \setminus R_5} \xi_1^2 \, d\mu_{F_1 \setminus R_5} = \| \xi_1 \|_2^2 .$$

This proves the lemma.

Let B_q be the set of bifold q-tuples in Q defined in the same way as for quadruples:

$$\mathcal{B}_q = \{ (I_1,\ldots,I_q) : 1_{I_1}+\ldots+ 1_{I_q} = 2\, 1_{I_1 \cup \ldots \cup I_q} \}$$

and

$$\mathcal{T}_q = \{ (I_1, \ldots, I_q) : 1_{I_1} + \ldots + 1_{I_q} \gneq 2\, 1_{I_1 \cup \ldots \cup I_q} \},$$

the set of q-tuples with each element in the union contained in at least two indices and at least one element in more than two indices. Let $C_B(d,q)$ be the number of different shadows in $\mathcal{B}_q$ and $C_T(d,q)$ be the number of different shadows in $\mathcal{T}_q$.

Proposition 2.4.3. Let the random variables W_I be indexed by subsets of the integers $\{1,\ldots,n\}$ of size $\le d$, $\{W_I : I \subset \{1,\ldots,n\}, |I| \le d\}$, with $E\, W_I = 0$, $EW_I^2 = \sigma_I^2$ and $\sum_{|I| \le d} \sigma_I^2 = 1$. Put

$$D_q = \max_I E\, |W_I|^q / \sigma_I^q ,$$

$$\rho = \max_i \sum_{I \ni i} \sigma_I^2.$$

Then

a) $$\sum_{(I_1,\ldots, I_q) \in \mathcal{B}_q} \sigma_{I_1} \cdots \sigma_{I_q} \le C_B(d,q),$$

b) $$\sum_{(I_1,\ldots, I_q) \in \mathcal{T}_q} \sigma_{I_1} \cdots \sigma_{I_q} \le C_T(d,q)\, \rho^{1/2} .$$

Moreover,

a') $$\sum_{(I_1,\ldots, I_q) \in \mathcal{B}_q} E\, |W_{I_1} \cdots W_{I_q}| \le D_q\, C_B(d,q),$$

b') $$\sum_{(I_1,\ldots, I_q) \in \mathcal{T}_q} E\, |W_{I_1} \cdots W_{I_q}| \le D_q\, C_T(d,q)\, \rho^{1/2}.$$

Remark. The inequalities a') and b') follow from a) and b) respectively, by Hölder's inequality and the definition of D_q. The inequalities a) and b) can be easily deduced from a') and b') respectively by defining $W_I' = \pm\, \sigma_I$ with equal probability; they are included for later reference;

Proof. Consider the fixed shadow $(F_1,\dots,F_q) \in Q$ with $|F_g| = e_g$, $g = 1,\dots,q$ and $F = F_1 \cup \dots \cup F_q = \{1,\dots,f\}$. Define $\varphi_g : \mathbb{N}^{e_g} \to [0,\infty)$ by

$$\varphi_g(n_1,\dots,n_{e_g}) = \begin{cases} \sigma_I & \text{if } I = \{n_1,\dots,n_{e_g}\} \text{ and } n_1 < \dots < n_{e_g} \\ 0 & \text{else} \end{cases}$$

and

$\pi_g : \mathbb{N}^F \to \mathbb{N}^{F_g}$ the natural projection,

$\xi_g = \varphi_g \circ \pi_g$, $g = 1,\dots,q$

(ξ_g is the function φ_g on $\mathbb{N}^{F_g}$ considered as function on $\mathbb{N}^F$). Then $\| \varphi_g \|_2^2 = \sum_{|I| = e_g} \sigma_I^2 \le 1$ and we have

$$\sum_{(I_1,\dots,I_q) \text{ with shadow } (F_1,\dots,F_q)} \sigma_{I_1} \dots \sigma_{I_q}$$

$$= \sum_{1 \le i_1 < \dots < i_f \le n} \xi_1(i_1,\dots,i_f) \dots \xi_q(i_1,\dots,i_f)$$

$$\le \sum_{1 \le i_1,\dots,i_f \le n} \xi_1(i_1,\dots,i_f) \dots \xi_q(i_1,\dots,i_f)$$

$$= \int^F \xi_1 \dots \xi_q \, d\mu_F$$

$$(2.4.1) \quad \le \begin{cases} 1 & \text{if } (F_1,\dots,F_q) \in \mathcal{B}_q, \\ \rho^{1/2} & \text{if } (F_1,\dots,F_q) \in \mathcal{T}_q. \end{cases}$$

The last inequality follows from lemma 2.4.2. If $(I_1,\dots,I_q) \in \mathcal{B}_q$ the conclusion follows from part a) of the lemma with $A = \emptyset$. If the shadow is in $\mathcal{T}_q$ we have $F_1 \cap F_2 \cap F_3 \ne \emptyset$ (or some other triple). By the Cauchy-Schwarz inequality we have

$$\int^F \varphi_1 \dots \varphi_q \, d\mu_F \le \| \varphi_1 \|_2 \left(\int^{F_1} \left(\int^{F \setminus F_1} \varphi_2 \dots \varphi_q \, d\mu_{F \setminus F_1} \right)^2 d\mu_{F_1} \right)^{1/2}$$

and we can apply part b) of the lemma to the functions $\varphi_2,\dots,\varphi_q$ taking into account that $1_{F \setminus F_1} (1_{F_2} + \dots + 1_{F_q}) \ge 2 \, 1_{F \setminus F_1}$ and $F_1 \cap F_2 \cap (F_3 \cup \dots \cup F_q) \supset F_1 \cap F_2 \cap F_3 \ne \emptyset$ and $\rho(\xi_1) \le \rho$. This proves (2.4.1) and thus the a) and b) part of the proposition by the trivial inequality: $a_1 + \dots + a_k \le k \max a_i$, if $a_i \ge 0$. By the Hölder inequality and the assumptions on the qth moments of W_I we have

$$E\, |W_{I_1} \dots W_{I_q}| \le E^{1/q} |W_{I_1}|^q \dots E^{1/q} |W_{I_q}|^q \le D_q \sigma_{I_1} \dots \sigma_{I_q}.$$

This proves the parts a') and b') of the proposition.

We conclude this section with some remarks.

Remark 1. Prop. 2.4.3 implies almost immediately Th. 2.2.4. We have only to show that the sixth moment remains bounded under the assumptions of the theorem. Since W(n) is clean and $E\, W_I^6 / \sigma_I^6 \le D$, we have by Prop. 2.4.3

$$E\, W(n)^6 \le D\, (C_T(d,6) + C_B(d,6)), \text{ since } \rho \le \sum_I \sigma_I^2 = 1.$$

Remark 2. In many cases the fourth moment condition $E\, W_I^4 / \sigma_I^4 \le D$, D not depending on n is not satisfied. If D diverges to infinity for $n \to \infty$, then it can be seen from Prop. 2.4.3 b) that it is sufficient for $\tau^* \to 0$ to impose the combination of the conditions a) and b) of Th. 2.2.3:

$$D_4\, \rho^{1/2} \to 0 \text{ for } n \to \infty.$$

We shall return briefly to this issue in Ch. 3 (Th. 3.1.2).

Remark 3. In Ch. 3 we shall show that for homogeneous random variables in the Hoeffding decomposition W_I we do not need Hölder's inequality to estimate $E\, |W_{I_1} \ldots W_{I_q}|$ for $(I_1, \ldots, I_q) \in \mathcal{B}_q$. In this case we have the sharper inequality

$$E\, |W_{I_1} \ldots W_{I_q}| \le \sigma_{I_1} \ldots \sigma_{I_q} \quad \text{if } (I_1, \ldots, I_q) \in \mathcal{B}_q.$$

3. Extensions and variations

3.0. Introduction

In this chapter two main items are considered. The first two sections are concerned with variations and extensions of the results of the previous chapter; in Sect. 3.3 we give some results on inhomogeneous sums.

One important aspect of the first two sections concerns the condition $E\,W(n)^4 \to 3$. It will be shown that, in some respects, this condition is also a negligibility condition. In Th. 3.1.5 it is shown that the usual negligibility condition and the fourth moment condition in Th. 2.2.3 can be replaced by a stronger negligibility condition

$$\max_i \sum_{I \ni i} \sigma_I \to 0, \quad n \to \infty.$$

From Sect. 3.2 on, we restrict ourselves to the Hoeffding decomposition. Firstly, it is shown how with the family coefficients $(a_I)_{|I| = d}$ several ([d/2]) rectangular matrices can be associated. Then it is shown that the maximal singular values of the matrices all vanish iff $E\,W(n)^4 \to 3$ for the multilinear form $W(n)$ with coefficients (a_I) in independent centered random variables X_i with $E\,X_i^2 = 1$, $E\,X_i^4 \leq D$. If the maximal singular value of a rectangular matrix $\mathbb{R}^m \to \mathbb{R}^n$is small, then the image of every point in the unit ball of $\mathbb{R}^m$ is small. In this respect the fourth moment condition is a strong negligibility condition. Section 3.2 contains some related results on singular values.

Another aim of this section is to formulate central limit theorems without any reference to fourth moments. Above, it is indicated how the fourth moment condition can be repaced. By means of truncation the uniform bound on the fourth moments of $W_I \,/\, \sigma_I$ can be replaced by a uniform bound on the tails of the distribution of these random variables. These results together lead to Th. 3.2.7 and Th. 3.2.8, which generalize results (for d = 2) in De Jong (1987).

In Sect. 3.3 we consider a finite sum of homogeneous sums

$$V(n) = \sum_{1 \leq e \leq d} W^{(e)}(n), \text{ with } W^{(e)}(n) = \sum_{|I| = e} W_I .$$

Suppose that the variance of each e-homogeneous sum $W^{(e)}(n)$ has a finite non-zero limit

$$\operatorname{var} W^{(e)}(n) \to \sigma^2(e) > 0.$$

If $W^{(e)}(n) \,/\, \operatorname{var}^{1/2} W^{(e)}(n)$ satisfies the conditions of Th. 2.1.1, for $1 \leq e \leq d$, then $V(n)$ has a normal limit distribution. However, these conditions may be difficult to check, since it may be hard to obtain the desired information on the separate homogeneous sums in situations where information on the total sum only is available . The chapter is concluded with an example of an inhomogeneous sum: a multilinear form in iid random variables. These random variables have a particular simple Hoeffding decomposition, as is shown in Prop. 3.3.5.

3.1. Miscellaneous results

In this section we shall freely use the quantities τ, τ^*, S_0, etc., as defined in Ch. 2. We start with the proof of Remark 2 in Sect. 2.2 concerning γ.

Proposition 3.1.1. For (not necessarily clean) random variables W_I indexed by subsets of the integers of size d with $\sum_{|I| = d} \sigma_I^2 = 1$ and τ and τ^* both vanishing, we have

$$\gamma \to 0 \text{ for } n \to \infty \text{ iff } \sum_{|I| = d} W_I^2 \xrightarrow{L^2} 1 \text{ for } n \to \infty.$$

Proof. Recall that $\gamma = |\sum_{I \cap J = \emptyset} (E\, W_I^2 W_J^2 - \sigma_I^2 \sigma_J^2)|$. The proposition follows, since

$$\operatorname{var}\Big(\sum_{|I| = d} W_I^2\Big) = \sum_{I \cap J \neq \emptyset} (E\, W_I^2 W_J^2 - \sigma_I^2 \sigma_J^2) + \sum_{I \cap J = \emptyset} (E\, W_I^2 W_J^2 - \sigma_I^2 \sigma_J^2),$$

since $\sum_{I \cap J \neq \emptyset} E\, W_I^2 W_J^2 \leq \tau^*$ and $\sum_{I \cap J \neq \emptyset} \sigma_I^2 \sigma_J^2 \leq \tau$.

The next theorem is a simple extension of Th. 2.2.3 with the help of Prop. 2.4.3.

Theorem 3.1.2. (Corollary to Th. 2.2.3.) Let W(1), W(2),... be a sequence of clean homogeneous sums $W(n) = \sum_{|I| = d} W_I$ with var W(n) = 1, for n = 1,2,... . Define

$$D(n) = \max_I E\, W_I^4 / \sigma_I^4,$$

$$\rho(n) = \max_i \sum_{I \ni i} \sigma_I^2.$$

Suppose

a) $\rho(n)^{1/2} D(n) \to 0$ for $n \to \infty$,

b) $E\, W(n)^4 \to 3$ for $n \to \infty$,

c) $\gamma \to 0$ for $n \to \infty$.

Then

$$W(n) \xrightarrow{d} N(0,1) \text{ for } n \to \infty.$$

Proof. We shall check the conditions of the technical central limit theorem Prop. 2.3.4. Since $D(n) \geq 1$, assumption a) ensures that $\rho(n)$ vanishes. Thus, by Prop. 2.4.3 b) and b') we have that τ and τ^* both vanish. If τ, τ^* and γ all vanish, then assumption b) implies $S + 3S_0 \to 0$ (Prop. 2.3.5). By the propositions 2.3.5 and 2.3.6 and (2.3.6) both S and S_0 vanish. This proves the theorem.

The next proposition will be used often in the sequel. Recall that the sum over all bifold quadruples was split into partial sums:

$$\sum_{\mathcal{B}} E\, W_I W_J W_K W_L = \sum_{0 \leq e \leq d} \; \sum_{0 \leq f \leq d-e} S(e,f).$$

In the proof of Prop. 2.3.6.a) it was shown that each of the quantities $S(e,0)$ ($= S(0,e) = S(e,d-e)$), $e = 1,\ldots,d-1$ can be expressed as a sum of squares up to a remainder term which can be estimated by τ^*. For fixed e we shall write $S(e,0)$ symbolically (with α, β and γ subsets of the pairs (I,J)):

$$S(e,0) = \sum_{\alpha} \sum_{\beta} E\, (\sum_{\gamma} W_I W_J)^2 \; + R_1,$$

where R_1 is a sum over a subset of $\mathcal{T}$, hence $|R_1| \leq \tau^*$. We shall show that, with the same subsets α,β and γ as in the expression for $S(e,0)$ above, we also have

$$\sum_{0 \leq f \leq d-e} S(e,f) \; = \sum_{\alpha} E\, (\sum_{\beta} \; \sum_{\gamma} W_I W_J)^2 \; + R_2.$$

Again $|R_2| \leq \tau^*$. Then we have by the Cauchy-Schwarz inequality

$$(\sum_{\beta} \; \sum_{\gamma} W_I W_J)^2 \leq (\sum_{\beta} 1^2)(\sum_{\beta} \; (\sum_{\gamma} W_I W_J)^2).$$

This leads to the final inequality in the proof of Prop. 3.1.3. Schematically

$$\sum_{0 \leq f \leq d-e} S(e,f) \leq |\beta| \, (S(e,0) + |R_1|) + |R_2|.$$

Proposition 3.1.3. For $e = 1,\ldots,d-1$

$$\sum_{0 \leq f \leq d-e} S(e,f) \leq \binom{2d-2e}{d-e} S(e,0) + \tau^* \left(\binom{2d-2e}{d-e} + 1\right).$$

Proof. For a fixed set $C \subset \{1,\ldots,n\}$, $|C| = 2d - 2e$ we have

$$E\,(\sum_{I \,\Delta\, J = C} W_I \, W_J)^2$$

$$= E\Big(\sum_{A,\, A' \subset C,\, A \cap A' = \emptyset,\, |A| = |A'| = d-e} \; \sum_{I \setminus J = A,\, J \setminus I = A'} W_I W_J \Big)^2$$

$$\le \binom{2d-2e}{d-e} \sum_{A,\, A' \subset C,\, A \cap A' = \emptyset,\, |A| = |A'| = d-e} E\Big(\sum_{I \setminus J = A,\, J \setminus I = A'} W_I W_J \Big)^2,$$

by the Cauchy-Schwarz inequality: $(\sum_i b_i)^2 \le (\sum_i 1^2)(\sum_i b_i^2)$. Working out the left-hand side and summing over all subsets C we obtain

$$\sum_{C \subset \{1,\dots,n\},\, |C| = 2d - 2e} E\Big(\sum_{I \Delta J = C} W_I W_J \Big)^2$$

$$= \sum_{C \subset \{1,\dots,n\},\, |C| = 2d - 2e} \Big(\sum_{\mathcal{B},\, I \Delta J = C} E\, W_I W_J W_K W_L$$

$$+ \sum_{\mathcal{T},\, I \Delta J = C = K \Delta L} E\, W_I W_J W_K W_L \Big)$$

$$= \sum_{\mathcal{B},\, |I \cap J| = e} E\, W_I W_J W_K W_L + R$$

$$= \sum_{0 \le f \le d-e} S(e,f) + R,$$

with $|R| \le \tau^*$. The proposition follows by summation over all subsets C at the right-hand side in the inequality above and (2.3.5).

The above inequality is useful, since it reduces the amount of work if we check condition II of Prop. 2.3.4. This is made explicit in the corollary below.

Corollary 3.1.4. Suppose that $\tau^* \to 0$ for $n \to \infty$. Then

$$S \to 0 \text{ for } n \to \infty \quad \text{if} \quad S_0 \to 0 \text{ for } n \to \infty.$$

Proof. By Prop. 2.3.6 we have (since $\tau^* \to 0$) that S_0 vanishes iff S(e,0) vanishes for $1 \le e < d$. Thus $S_0 \to 0$ implies that $\sum_{1 \le e \le d-1} \sum_{0 \le f \le d-e} S(e,f) = S + 2S_0$ vanishes (by Prop. 3.1.3); hence $S \to 0$ for $n \to \infty$.

Corollary 3.1.4 is applied in the following theorem. It is shown that the fourth moment condition, $E\, W(n)^4 \to 3$, in Th. 2.2.3 can be dropped if the condition on the negligibility of the hyperplanes (condition a)) is strengthened considerably. An example of this strong negligibility is

(3.1.1) $\max_i \sum_{I \ni i} \sigma_I \to 0$ for $n \to \infty$.

Condition a) in Th. 3.1.5 below is a weaker version of (3.1.1).

Theorem 3.1.5. Let W(1), W(2),... be a sequence of clean d-homogeneous sums $W(n) = \sum_{|I| = d} W_I$ with var W(n) = 1, for n = 1,2,... . Suppose

a) $\max_{A \subset \{1,...,n\},\ 1 \le |A| \le d-1} \sum_{I \supset A} (\sigma_I \sum_{J \supset I \setminus A} \sigma_J) \to 0$ for $n \to \infty$,

b) $\max_I E W_I^4 / \sigma_I^4 \le D$, D not depending on n,

c) $\sum_{I \cap J = \varnothing} (E W_I^2 W_J^2 - \sigma_I^2 \sigma_J^2) \to 0$ for $n \to \infty$.

Then

$$W(n) \xrightarrow{d} N(0,1) \text{ for } n \to \infty.$$

Remark. Notice that $J \supset I \setminus A$ iff $J \cup A \supset I$. Hence condition a) is equivalent to

a') $\max_{A \subset \{1,...,n\},\ 1 \le |A| \le d-1} \sum_{|J| = d} (\sigma_J \sum_{A \subset I \subset J \cup A} \sigma_I) \to 0$ for $n \to \infty$,

where for each subset A the summation of J extends over all indices J.

Notice that

$$\sum_{I \supset A} \sigma_I \le \max_i \sum_{I \ni i} \sigma_I \quad \text{for all subsets } A \neq \varnothing.$$

Thus (3.1.1) implies $\sum_{J \supset I \setminus A} \sigma_J \to 0$ and hence condition a) of Th. 3.1.5.

Proof. We shall check the conditions of Prop. 2.3.4. Assumption c) is identical with condition III. Assumption a) with the maximum restricted to $|A| = 1$ and the summation over J for given I restricted to J = I reads:

$$\max_i \sum_{I \ni i} \sigma_I^2 \to 0, \quad n \to \infty.$$

Combined with assumption b) and Prop. 2.4.3 b) and b') this implies that τ and τ^* vanish respectively (condition I). By Prop. 3.1.4, condition II is satisfied if S_0 vanishes.

Consider a bifold shadow $(I',J',K',L') \in \mathcal{B}(e,0)$ and the triples (I',K',L') and (J',K',L'). Since $I' \cup J' \cup K' \cup L' = I' \cup K' \cup L' = J' \cup K' \cup L' (= \{1,...,2d\})$, both triples are shadows. Thus for each quadruple (I,J,K,L) with shadow (I',J',K',L') there is exactly one triple (I,K,L) with shadow (I',K',L') and $I \cup K \cup L = I \cup J \cup K \cup L$. The same holds for (J,K,L). By Hölder's inequality, assumption b) and the inequality $2\sigma_I \sigma_J \le \sigma_I^2 + \sigma_J^2$ we have

$$2|\,E\,W_I W_J W_K W_L\,| \le D(\sigma_I^2\sigma_K\sigma_L + \sigma_J^2\sigma_K\sigma_L).$$

This gives

$$2 \sum_{(I,J,K,L) \text{ with shadow } (I',J',K',L')} |E\, W_I W_J W_K W_L|$$

$$\le \sum_{(I,J,K,L) \text{ with shadow } (I',J',K',L')} D(\sigma_I^2\sigma_K\sigma_L + \sigma_J^2\sigma_K\sigma_L)$$

$$= D\Big(\sum_{(I,K,L) \text{ with shadow } (I',K',L')} \sigma_I^2\sigma_K\sigma_L + \sum_{(J,K,L) \text{ with shadow } (J',K',L')} \sigma_J^2\sigma_K\sigma_L\Big)$$

$$\le 2D \sum_{|I| = d} \sigma_I^2 \sum_{K\cap I = \emptyset} \sigma_K \max_{|L\cap I| = d-e} \sum_{L \setminus I \subset K} \sigma_L$$

$$\le 2D \sum_{|I| = d} \sigma_I^2 \max_{A \subset I,\ |A| = d-e} \sum_{K\cap A = \emptyset} \sigma_K \sum_{A \subset L \subset K\cup A} \sigma_L$$

$$\le 2D \max_{A \subset \{1,\dots,n\},\ |A| = d-e} \sum_{K\cap A = \emptyset} \sigma_K \sum_{A \subset L \subset K\cup A} \sigma_L .$$

Hence we have, with C_e the number of distinct shadows in $\mathcal{B}(e,0)$:

$$|\,S(e,0)\,| \le C_e\, D \max_{A \subset \{1,\dots,n\},\ 1\le|A|\le d-1} \sum_{|K| = d} \sigma_K \sum_{A \subset L \subset K \cup A} \sigma_L .$$

Since neither D nor C_e depend on n, the theorem follows by assumption a').

3.2. Results involving singular values and truncation

We recall some well-known facts from linear algebra. The matrix $A \in \mathbb{R}^{m\times n}$ can be brought into diagonal form by two orthogonal transforms, $U \in \mathbb{R}^{m\times m}$ and $V \in \mathbb{R}^{n\times n}$: $U^T A V = \text{diag}(\mu_1,\dots,\mu_r)$, $r = \min(m,n)$ with $\mu_1 \ge\dots\ge \mu_r \ge 0$, uniquely. This is the singular value decomposition of A (see Golub and Van Loan (1983: 16 ff.)). If $m = n$ the singular values $\mu_1,\dots,\mu_n$ are the absolute values of the eigenvalues of the matrix. Like the eigenvalues, singular values of a matrix A are related to matrix norms for A:

(3.2.1) $\| A \|_{HS}^2 := \sum_{1\le i\le m} \sum_{1\le j\le n} a_{ij}^2 = \mu_1^2 + \dots + \mu_r^2,$

(3.2.2) $\| A \|_2 := \max_{x \in \mathbb{R}^n,\ \|x\|_2 = 1} \| Ax \|_2 = \mu_1,$

with $\| x \|_2^2 = \sum_{1\le i\le n} x_i^2$ the squared Euclidean vector norm on $\mathbb{R}^n$. Notice that $\| Ax \|_2$ is the Euclidean vector norm on $\mathbb{R}^m$. The matrix norm $\| A \|_{HS}$ is the Hilbert -Schmidt (or

Frobenius) norm and $\| A \|_2$ the operator norm. These facts will be used in the proof of Prop. 3.2.1.

Consider the family of real valued constants $(a_I)_{I \subset \{1,...,n\}, |I| = d}$. First we shall extend the family indexed by subsets $I \subset \{1,...,n\}$ to a symmetric 'd-dimensional matrix' $A = (a_{i_1...i_d})_{1 \le i_1,...,i_d \le n}$, indexed by d-tuples:

$$a_{i_1...i_d} = \begin{cases} a_I & \text{if } \{i_1,...,i_d\} = I \\ 0 & \text{else.} \end{cases}$$

From $A \in \mathbb{R}^{n^d}$ we can form rectangular matrices $A(e) \in \mathbb{R}^{n^e \times n^{d-e}}$, $1 \le e \le d-1$,

$$A(e) = (a(e)_{ij})_{i=1,...,n^e,\, j=1,...,n^{d-e}}.$$

The indices i and j are obtained by splitting a d-tuple is split into two parts: $(i_1,...,i_d) = (i_1,...,i_e, j_1,...,j_{d-e})$. Then we have $i = i_e n^{e-1} +...+ i_1 n^0$, $j = j_{d-e} n^{d-e-1} +...+ j_1 n^0$ and

$$a(e)_{ij} = a_{i_1...i_e j_1...j_{d-e}}.$$

Conversely, any number i can be written uniquely in a n-ary expansion, etc. Any other choice of $(i_1,...,i_e)$ among $(i_1,...,i_d)$ results in the same matrix A(e) by the symmetry in A.

The singular values of A(e) are denoted by $\mu(e)_1 \ge ... \ge \mu(e)_r$ with $r = \min(n^e, n^{d-e})$, and by (3.2.1) we have

$$\mu(e)_1^2 +...+ \mu(e)_r^2 = d! \sum_{|I| = d} a_I^2.$$

Since the matrices A(e) are uniquely determined by the family $(a_I)_{|I| = d}$ we shall say $(a_I)_{|I| = d}$ has singular values $\mu(e)_1 \ge ... \ge \mu(e)_r$, $1 \le e \le d-1$ and maximal singular value

$$\mu^* = \max_{1 \le e < d} \mu(e)_1.$$

The singular value decomposition of the square matrix $A(e)^T A(e) \in \mathbb{R}^{n^{d-e} \times n^{d-e}}$ gives (with the orthogonal transforms U,V defined above):

$$V^T A(e)^T A(e) V = V^T A(e)^T U U^T A(e) V = \mathrm{diag}(\mu(e)_1^2,...,\mu(e)_r^2),$$

with $r = \min(n^e, n^{d-e})$. And by (3.2.1) we have

$$\| A(e)^T A(e) \|_{HS}^2 = \mu(e)_1^4 +...+ \mu(e)_r^4.$$

The following proposition links the facts from linear algebra to the quantities defined in Ch.2. Define for the family $(a_I)_{|I| = d}$

$$S(e,0) = \sum_{\mathcal{B}(e,0)} a_I a_J a_K a_L,$$

$$\tau = \sum_{\mathcal{T}} | a_I a_J a_K a_L |,$$

$$S_0 = \sum_{1 \le e \le d-1} S(e,0).$$

Prop. 3.2.1 below is equivalent with Lemma 5.1 in De Jong (1987) in the case d = 2. (If d = 2 then μ^* is the maximal absolute value of the eigenvalues.)

Proposition 3.2.1. Let S_0, τ, A, $\mu(e)_1,\ldots, \mu(e)_r$, μ^* be defined as above for the families $(a_I)_{I \subset \{1,\ldots,n\}, |I| = d}$ with $\sum_{|I| = d} a_I^2 = 1$ for n = 1, 2,... . Then

$$S_0 \to 0,\ n \to \infty \ \text{ and } \max_i \sum_{I \ni i} a_I^2 \to 0,\ n \to \infty,$$

iff

$$\mu^* \to 0,\ n \to \infty.$$

Proof. Let e_i be the ith unit vector in R^n, then

$$\text{(3.2.3)} \quad \max_i \sum_{I \ni i} a_I^2 \le \max_i \| A(d-1)\, e_i \|_2^2 \le \mu(d-1)_1^2,$$

where the final inequality follows from (3.2.2). We have

$$\mu(e)_1^4 \le \mu(e)_1^4 + \ldots + \mu(e)_r^4 = \| A(e)^T A(e) \|_{HS}^2 \le d!\, \mu(e)_1^2,$$

by $\mu(e)_1^2 + \ldots + \mu(e)_r^2 = d!$. This implies

$$\mu(e)_1 \to 0 \quad \text{iff} \quad \| A(e)^T A(e) \|_{HS} \to 0,\ n \to \infty.$$

We need, under the assumption that $\max_i \sum_{I \ni i} a_I^2$ vanishes,

$$\| A(e)^T A(e) \|_{HS} \to 0 \text{ iff } S(e,0) \to 0,\ 1 \le e \le d-1.$$

Consider one element in the square matrix $A(e)^T A(e) = (b_{ij})_{1 \le i,j \le n^{d-e}}$:

$$b_{ij} = \sum_{1 \le k \le n^e} a(e)_{ik} a(e)_{jk}.$$

Writing i,j and k as n-ary numbers and expressing $a(e)_{ik}$ and $a(e)_{jk}$ in $a_{i_1 \ldots i_d}$ we obtain:

$$b_{ij} = b_{i_1 \ldots i_{d-e} j_1 \ldots j_{d-e}} = \sum_{1 \le k_1,\ldots,k_e \le n} a_{i_1 \ldots i_{d-e} k_1 \ldots k_e}\, a_{j_1 \ldots j_{d-e} k_1 \ldots k_e}.$$

Notice that $\{i_1,\ldots,i_{d-e}\} \cap \{k_1,\ldots,k_e\} \ne \emptyset$ implies $a(e)_{ik} = 0$ and $\{j_1,\ldots,j_{d-e}\} \cap \{k_1,\ldots,k_e\} \ne \emptyset$ implies $a(e)_{jk} = 0$. With the convention $a_I = 0$ if $|I| < d$ and using the abbreviations $B = \{i_1,\ldots,i_{d-e}\}$ and $B' = \{j_1,\ldots,j_{d-e}\}$, we have

$$b_{ij} = e! \sum_{A \subset \{1,...,n\},\ |A| = e,\ I=A\cup B,\ J=A\cup B'} a_I a_J \ .$$

Hence

$$b_{ij}^2 = (e!)^2 \sum_{A,A'\subset\{1,...,n\},\ |A| = |A'| = e,\ I=A\cup B,\ J=A\cup B',\ K=A'\cup B',\ L=A'\cup B} a_I a_J a_K a_L$$

$$= (e!)^2 \sum_{\mathcal{B}(e,0),\ I\cap L = B,\ J\cap K = B'} a_I a_J a_K a_L + R_{ij},$$

with $R_{ij} \leq R^*_{ij}$,

$$R^*_{ij} = (e!)^2 \sum_{\mathcal{T},\ I\cap L = B,\ J\cap K = B',\ |I\cap J| = e,\ |K\cap L| = e} |a_I a_J a_K a_L|.$$

Notice that there are no quadruples with a free index; moreover, if $\{i_1,...,i_{d-e}\} \cap \{j_1,...,j_{d-e}\} \neq \varnothing$, there are no bifold quadruples. Thus we have

(3.2.4) $\quad \left| \ \| A(e)^T A(e) \|^2_{HS} - ((d-e)!\ e!)^2 S(e,0) \right| \leq \sum_{1\leq i,\ j\leq n^{d-e}} R^*_{ij}$

and

(3.2.5) $\quad \sum_{1\leq i,\ j\leq n^{d-e}} R^*_{ij} \leq \tau((d-e)!\ e!)^2.$

By Prop. 2.4.3 b) and (3.2.3) we have

$$\tau \leq C_T(d,4)\ \mu(d-1)_1.$$

The rest of the proof is obvious. If $\mu(e)_1 \to 0$ for $1 \leq e \leq d-1$, which is equivalent to $\mu^* \to 0$, then $\tau \to 0$ and $\| A(e)^T A(e) \|^2_{HS} \to 0$ for $1 \leq e \leq d-1$; thus (3.2.4) and (3.2.5) imply $S(e,0) \to 0$, for $1 \leq e \leq d-1$. This proves, by (3.2.3), the if part of the proposition.

On the other hand, if

$$\max_i \sum_{I \ni i} a_I^2 \to 0,$$

then $\tau \to 0$ by Prop. 2.4.3 b). In order to apply Prop. 2.3.6, which holds for clean random variables W_I, we notice that with

$$W_I = a_I \prod_{i\in I} X_i \ \ (X_i = \pm 1 \text{ with probability equaling } 1/2 \text{ and independent})$$

we have $W_I W_J W_K W_L = a_I a_J a_K a_L$ and since $E\, W_I^4 / \sigma_I^4 \leq 1$ we have $\tau^* \leq \tau$. Thus, by Prop. 2.3.6 we have $S_0 \to 0$ implies $S(e,0) \to 0$, (since $\tau \to 0$) hence $\| A(e)^T A(e) \|_{HS} \to 0$ by (3.2.4) and (3.2.5); and consequently $\mu(e)_1 \to 0$ for $1 \leq e \leq d-1$. This completes the proof of the proposition.

Now we return to the Hoeffding decomposition. Recall that on the probability space $(\Omega,\mathcal{F},P)$ a sequence of independent random variables X_1, X_2,... is given. Define for finite subsets of the integers $I \subset \{1,...,n\}$ the σ-algebras $\mathcal{F}_I = \sigma\{X_i : i \in I\}$, $\mathcal{F}_\varnothing = \{\varnothing,\Omega\}$. Then any square integrable $\mathcal{F}_{\{1,...,n\}}$-measurable random variable Z can be written

$$Z = \sum_{I \subset \{1,..., n\}} W_I,$$

where the components in the Hoeffding decomposition W_I are uniquely determined by

a) W_I is $\mathcal{F}_I$ measurable,

b) $E(W_I \mid \mathcal{F}_J) = 0$ a.s. if $I \setminus J \neq \varnothing$.

The next proposition shows that for V, a component in the Hoeffding decomposition, there is a bounded component V' close (in L^2-sense) to V.

Proposition 3.2.2. Let V be any square integrable random variable, satisfying the conditions a) and b) above. For each $C \geq 0$ there is a random variable V' satisfying a) and b) with

$$E(V - V')^2 \leq E V^2 1_{\{|V| > C\}}$$

and

$$|V'| \leq 2^{|I|} C.$$

Proof. Truncate V at C: $V^* = V1_{\{|V| \leq C\}}$. Then V^* can be written in the Hoeffding decomposition as

$$V^* = \sum_{J \subset I} W_J.$$

Call W_I the *clean version* of V^* and put $V' = W_I$. Then V - V' is the clean version of V - V^*, since we have in the Hoeffding decomposition:

$$V - V^* = \sum_{J \subsetneq I} - W_J + (V - V').$$

By the orthogonality of the Hoeffding decomposition

$$E(V - V')^2 \leq E(V - V^*)^2 = E V^2 1_{\{|V| > C\}}.$$

From Prop. 3.2.3 below we see that V' can be written as a sum of $2^{|I|}$ terms each bounded in absolute value by C: $|V'| \leq 2^{|I|} C$. This proves the proposition.

We give an expression for components in the Hoeffding decomposition in terms of conditional expectations (cf. Prop. 2.2 in Karlin and Rinott (1982)).

Proposition 3.2.3. Let Z be $\mathcal{F}_I$-measurable, and $Z = \sum_{J \subset I} W_J$ (Hoeffding decomposition). Then

$$W_I = \sum_{J \subset I} E(Z \mid \mathcal{F}_J)\,(-1)^{|I| - |J|}.$$

Proof. The proof is by induction. For $I = \varnothing$ the assertion is trivial. For $I \neq \varnothing$ we have

$$\begin{aligned}
W_I &= Z - \sum_{J \subsetneq I} W_J \\
&= Z - \sum_{J \subsetneq I} \sum_{K \subset J} E(Z \mid \mathcal{F}_K)\,(-1)^{|J| - |K|} \\
&= Z - \sum_{K \subsetneq I} E(Z \mid \mathcal{F}_K) \sum_{K \subset J \subsetneq I} (-1)^{|J| - |K|} \\
&\overset{(1)}{=} Z - \sum_{K \subsetneq I} E(Z \mid \mathcal{F}_K)\,\left((1-1)^{|I| - |K|} - (-1)^{|I| - |K|}\right) \\
&= \sum_{J \subset I} E(Z \mid \mathcal{F}_J)\,(-1)^{|I| - |J|}.
\end{aligned}$$

Equality (1) follows by (for $K \subsetneq I$, I fixed)

$$\sum_{K \subset J \subsetneq I} (-1)^{|J| - |K|} = \sum_{J' \subsetneq I \setminus K} (-1)^{|J'|} = \sum_{J' \subset I \setminus K} (-1)^{|J'|} - (-1)^{|I \setminus K|}.$$

This proves the proposition.

In estimating the expectation $E\, W_I W_J W_K W_L$ Hölder's inequality was used involving fourth moments $E\, W_I^4$. If the quadruple (I,J,K,L) is bifold and the random variables W_I are components in the Hoeffding decomposition, a sharper inequality, involving only second moments, is available. Consider the case $(I,J,K,L) \in \mathcal{B}(e,0)$, then by the Cauchy-Schwarz inequality we have

$$E\,|W_I W_J W_K W_L| \le E^{1/2}\,(W_I W_K)^2\; E^{1/2}\,(W_J W_L)^2 = \sigma_I \sigma_J \sigma_K \sigma_L,$$

where the last inequality follows by the independence of W_I and W_K ($I \cap K = \varnothing$), and that of W_J and W_L respectively. In fact, the above inequality holds for dissociated random variables. The obvious idea of invoking Prop. 3.1.3 to obtain an inequality for all bifold quadruples fails, since the remainder term in Prop. 3.1.3 involves τ^*, a sum of higher moments. The idea of the proof of Prop. 3.2.4 below is a refinement of the above inequality. It resembles the proof of Lemma 2.4.2.

Proposition 3.2.4. Let $W_{I_1}, \ldots, W_{I_q}$ be components in the Hoeffding decomposition and let $(I_1, \ldots, I_q)$ be a bifold q-tuple. Then

$$E\,|W_{I_1} \cdots W_{I_q}| \le \sigma_{I_1} \cdots \sigma_{I_q}.$$

Proof. We start with an assertion which resembles that of Lemma 2.4.2. If $F = I_1 \cup \ldots \cup I_q$ and $A \subset F$ is covered only once and $F \setminus A$ exactly twice by the sets $I_1, \ldots, I_q$, i.e.

$$1_{I_1} + \ldots + 1_{I_q} = 1_A + 2\, 1_{F \setminus A},$$

then

$$\text{(3.2.6)} \quad E\, E^2(\, |W_{I_1} \ldots W_{I_q}| \mid \mathcal{F}_A) \le \sigma^2_{I_1} \ldots \sigma^2_{I_q}.$$

The proof is by induction. If $q = 1$ the assertion is trivial: $F = I = A$ and $E\, W_I^2 = \sigma_I^2$. Suppose (3.2.6) holds for (q-1)-tuples ($q \ge 2$). Then, with $G = I_2 \cup \ldots \cup I_q$,

$$E\, E^2(\, |W_{I_1} \ldots W_{I_q}| \mid \mathcal{F}_A)$$

$$= E\, E^2(\, |W_{I_1}|\, E\, (|W_{I_2} \ldots W_{I_q}| \mid \mathcal{F}_{A \cup I_1}) \mid \mathcal{F}_A)$$

$$\overset{(1)}{\le} E\, E(\, W_{I_1}^2 \mid \mathcal{F}_A)\; E(\, E^2(\, |W_{I_2} \ldots W_{I_q}| \mid \mathcal{F}_{A \cup I_1}) \mid \mathcal{F}_A)$$

$$\overset{(2)}{=} E\, E(\, W_{I_1}^2 \mid \mathcal{F}_{A \cap I_1})\; E(\, E^2(\, |W_{I_2} \ldots W_{I_q}| \mid \mathcal{F}_{(A \cup I_1) \cap G}) \mid \mathcal{F}_{A \cap G})$$

$$\overset{(3)}{=} \sigma^2_{I_1}\; E\, E^2(\, |W_{I_2} \ldots W_{I_q}| \mid \mathcal{F}_{(A \cup I_1) \cap G}),$$

where inequality (1) is the conditional version of the Cauchy-Schwarz inequality and equality (2) follows since $E(\, W_I \mid \mathcal{F}_J\,) = E(\, W_I \mid \mathcal{F}_{J \cap I}\,)$ a.s. by the independence of the underlying random variables (see Chung (1972, Th. 9.2.1)). Equality (3) follows by the independence, since $(A \cap I_1)$ and $(A \cap G)$ are disjoint by the definition of A in which every element is contained in only one set. This proves (2.3.6) as

$$1_{I_2} + \ldots + 1_{I_q} = 1_{(A \cup I_1) \cap G} + 2\, 1_{G \setminus (A \cup I_1)}.$$

The proposition follows from (3.2.6) with $A = \emptyset$.

Remark 1. Notice that Prop. 3.2.4 remains valid for $\mathcal{F}_I$-measurable, zero-mean square integrable random variables W_I. However, the independence of the underlying random variables X_i is used in an essential way: The equalities (2) and (3) rest on it.

Remark 2. The inequality of Prop. 3.2.4 is sharp: If $W_I = a_I \prod_{i \in I} X_i$, with X_i independent, $E\, X_i = 0$, $E\, X_i^2 = 1$ $a_I \in \mathbb{R}$ and $(I_1, \ldots, I_q)$ is a bifold q-tuple, then

$$E\, W_{I_1} \ldots W_{I_q} = a_{I_1} \ldots a_{I_q} \prod_{i \in I_1 \cup \ldots \cup I_q} E\, X_i^2 = \pm\, \sigma_{I_1} \ldots \sigma_{I_q},$$

according to the sign of $a_{I_1} \ldots a_{I_q}$.

We shall use Prop. 3.2.2 and Prop. 3.2.4 to sharpen Th. 2.2.4 in case W(n) is a d-homogeneous sum in the Hoeffding decomposition. The sufficiency of the fourth moment condition $E\,W(n)^4 \to 3$ is proved without any assumption on sixth moments (as in Th. 2.2.4).

Theorem 3.2.5. Let W(1), W(2),... be a sequence of d-homogeneous sums in the Hoeffding decomposition, $W(n) = \sum_{|I|=d} W_I$, with var W(n) = 1 for n = 1,2,.... Suppose

a) $\max_i \sum_{I \ni i} \sigma_I^2 \to 0, \; n \to \infty$,

b) $\max_I E\,W_I^4 / \sigma_I^4 \le D$, D not depending on n.

Then the following two statements are equivalent

1) $E\,W(n)^4 \to 3, \quad n \to \infty$,

2) $W(n) \xrightarrow{d} N(0,1), \; n \to \infty$.

Proof. Th. 2.1.1 states that 1) implies 2). Assume 2). Put

$$\alpha = (\max_I \sum_{I \ni i} \sigma_I^2)^{-1/12}.$$

Let W'_I be the clean version (See proof of Prop. 3.2.2) of $W_I \, 1_{\{ |W_I| \le \alpha \sigma_I \}}$. Then, with $W'(n) = \sum_{|I|=d} W'_I$, we have $W'(n) - W(n) \xrightarrow{L^2} 0$, since

$$\text{var}\,(W(n) - W'(n)) \overset{(1)}{=} \sum_{|I|=d} E\,(W_I - W'_I)^2$$

$$\overset{(2)}{\le} \sum_{|I|=d} E\,W_I^2 \, 1_{\{ |W_I| > \alpha \sigma_I \}}$$

$$\le \alpha^{-2} \sum_{|I|=d} \sigma_I^{-2} E\,W_I^4$$

$$\le D / \alpha^2,$$

which vanishes, since $\alpha \to \infty$. Equality (1) follows since $(W_I - W'_I)$ are components in the Hoeffding decomposition of W(n) - W'(n) and inequality (2) by Prop. 3.2.2. Thus 2) implies $W'(n) \xrightarrow{d} N(0,1), \; n \to \infty$. If $E\,W'(n)^6$ remains bounded then 2) implies (by Th. 2.4.4) $E\,W'(n)^4 \to 3$. The former will be shown. Recall that $|W'_I| \le 2^d \alpha \, \sigma_I$ by Prop. 3.2.2. For n sufficiently large we have $\sigma_I / \sigma'_I \le 2$. Then

$$E\,(W'_I / \sigma'_I)^6 \le 2^6 \, E\,(W'_I / \sigma_I)^6 \le 2^{6(d+1)} \alpha^6.$$

By Prop. 2.4.3 b') we have

$$\sum_{(I_1,\dots,I_6)\in \mathcal{T}_6} |E\, W'_{I_1}\cdots W'_{I_6}|$$

$$\le C_T(d,6)\, 2^{6(d+1)}\, \alpha^6 (\max_i \sum_{I \ni i} \sigma_I^2)^{1/2}$$

$$= C_T(d,6)\, 2^{6(d+1)}.$$

By Prop. 3.2.4 and Prop. 2.4.3 a) we have

$$\sum_{(I_1,\dots,I_6)\in \mathcal{B}_6} |E\, W'_{I_1}\cdots W'_{I_6}| \le \sum_{(I_1,\dots,I_6)\in \mathcal{B}_6} E\, \sigma'_{I_1}\cdots\sigma'_{I_6}$$

$$\le \sum_{(I_1,\dots,I_6)\in \mathcal{B}_6} E\, \sigma_{I_1}\cdots\sigma_{I_6}$$

$$\le C_B(d,6),$$

where the second inequality follows from

$$(3.2.6)\quad \sigma'^2_I = E\, W'^2_I \le E\, W_I^2\, 1_{\{|W_I| \le \alpha\, \sigma_I\}} \le \sigma_I^2.$$

Hence $E\, W'(n)^4 \to 3$. Combining the assumptions a) and b) with Prop. 2.4.3 b') we have $\tau^* \to 0$ and also $\tau'^* \to 0$ (that is τ^* for W'(n)). Thus $S' + 3S'_0 \to 0$ by Prop. 2.3.5. Since $E\, W(n)^4 = 3\, S(0,0) + 3S_0 + \sum_{\mathcal{T}} E\, W_I W_J W_K W_L$, with $|S(0,0) - 1| \le \tau$ (cf. proof of Prop. 2.3.5) it remains to show $S + 3\, S_0 \to 0$. This follows by

$$|E\, W_I W_J W_K W_L - E\, W'_I W'_J W'_K W'_L|$$

$$\begin{aligned} = |\, & E\, W_I W_J W_K W_L - E\, W'_I W_J W_K W_L \\ & + E\, W'_I W_J W_K W_L - E\, W'_I W'_J W_K W_L \\ & + E\, W'_I W'_J W_K W_L - E\, W'_I W'_J W'_K W_L \\ & + E\, W'_I W'_J W'_K W_L - E\, W'_I W'_J W'_K W'_L \,| \end{aligned}$$

$$\le \sigma''_I \sigma_J \sigma_K \sigma_L + \sigma'_I \sigma''_J \sigma_K \sigma_L + \sigma'_I \sigma'_J \sigma''_K \sigma_L + \sigma'_I \sigma'_J \sigma'_K \sigma''_L,$$

where the last inequality follows by the triangle inequality and Prop. 3.2.4, with $\sigma''^2_I = E\, (W_I - W'_I)^2$. Hence we have by Prop. 2.4.3 a)

$$\sum_{\mathcal{B}} |E\, W_I W_J W_K W_L - E\, W'_I W'_J W'_K W'_L|$$

$$\le \sum_{\mathcal{B}} (\sigma''_I \sigma_J \sigma_K \sigma_L + \sigma_I \sigma''_J \sigma_K \sigma_L + \sigma_I \sigma_J \sigma''_K \sigma_L + \sigma_I \sigma_J \sigma_K \sigma''_L)$$

$$\le 4\, C_B(d,4) \,(\sum_{|I| = d} \sigma''^2_I)^{1/2},$$

which vanishes as is shown above. This proves the theorem.

This theorem yields in combination with Prop. 3.2.1 the following corollary for multilinear forms.

Corollary 3.2.6. Let W(1), W(2),... be homogeneous d-linear forms

$$W(n) = \sum_{|I| = d} a_I \prod_{i \in I} X_i, \text{ with } X_i \text{ independent, } E\,X_i = 0 \text{ and } E\,X_i^2 = 1,$$

and (a_I) a family real constants with $\sum_{|I| = d} a_I^2 = 1$, for n = 1,2,... and maximal singular value μ^*. Suppose

a) $\max_i \sum_{I \ni i} a_I^2 \to 0, \quad n \to \infty,$

b) $E\,X_i^4 \le D < \infty$, for all i, D not depending on n.

Then the following two statements are equivalent

1) $\mu^* \to 0$ for $n \to \infty$,

2) $W(n) \xrightarrow{d} N(0,1)$ for $n \to \infty$.

Proof. Since $E\,W_I^2 = a_I^2$ we have by a)

$$\max_i \sum_{I \ni i} \sigma_I^2 \to 0, \quad n \to \infty,$$

and by b) we have

$$\max_I E\,W_I^4 / \sigma_I^4 \le D^d.$$

Thus we have by Prop. 2.4.3 that both τ and τ^* vanish. This implies that the following five statements are equivalent.

1) $\mu^* \to 0$,

2) $\sum_{1 \le e \le d-1} \sum_{\mathcal{B}(e,0)} a_I a_J a_K a_L \to 0$,

3) $S_0 \to 0$,

4) $E\,W(n)^4 \to 3$,

5) $W(n) \xrightarrow{d} N(0,1)$.

We have

1) $\Leftrightarrow$ 2) follows by Prop. 3.2.1.

2) $\Leftrightarrow$ 3) follows by $E\,W_I W_J W_K W_L = a_I a_J a_K a_L$ for bifold quadruples since $E\,X_i^2 = 1$.

3) $\Rightarrow$ 4) follows from the expression for the fourth moment $E\,W(n)^4$ and by Corollary 3.1.4, since τ and τ^* vanish.

3) $\Leftarrow$ 4) follows from Sect.2.3, since τ and τ^* vanish.

4) $\Leftrightarrow$ 5) follows by Th. 3.2.5.

This finishes the proof of the corollary.

In Th. 3.2.7 below a uniform bound is imposed on the tails of the distributions of W_I / σ_I:

$P\{ | W_I | > x\,\sigma_I \} \le R(x)$ for all I, with R(x) monotone and not depending on n, such that $\int_{(0,\infty)} x\, R(x)\, dx < \infty$.

By partial integration we obtain for the random variable $U = W_I / \sigma_I$

$$E\, U^2\, 1_{\{ |U| > C \}} = \int_{(C,\infty)} x^2 dP\{ | U | \le x \}$$

$$= - \int_{(C,\infty)} x^2 d(1 - P\{ | U | \le x \})$$

$$\overset{(1)}{=} C^2 P\{ | U | > C \} + 2 \int_{(C,\infty)} x\, P\{ | U | > x \}\, dx$$

$$\le C^2\, R(C) + 2 \int_{(C,\infty)} x\, R(x)\, dx$$

$$\overset{(2)}{\le} 3 \int_{(1/2C,\infty)} x\, R(x)\, dx,$$

where inequality (2) follows by the monotony of R(x) which implies

$$\int_{(1/2C,C)} x\, R(x)\, dx \ge R(C) \int_{(1/2C,C)} x\, dx = 3/8\, C^2\, R(C).$$

Since the integral converges, the left-hand side vanishes if C tends to infinity . This proves equality (1). Now we can prove the following theorem.

Theorem 3.2.7. Let W(1), W(2),... be d-homogeneous sums in the Hoeffding decomposition, $W(n) = \sum_{|I| = d} W_I$, with var W(n) = 1. Let (σ_I) have maximal singular value μ^*. Suppose

a) $P\{ | W_I | > x\,\sigma_I \} \le R(x)$ for all I, with R(x) monotone and not depending on n, such that $\int_{(0,\infty)} x\, R(x)\, dx < \infty$,

b) $\mu^* \to 0, \; n \to \infty$.

Then

$$W(n) \xrightarrow{d} N(0,1), \; n \to \infty.$$

Proof. By Prop. 3.2.1 assumption b) implies

$$\max_i \sum_{I \ni i} \sigma_I^2 \to 0 \text{ for } n \to \infty.$$

Let C_n be a sequence such that

$$C_n \to \infty \quad \text{and} \quad C_n^4 \max_i \sum_{I \ni i} \sigma_I^2 \to 0 \text{ for } n \to \infty.$$

Define $W_I^* = W_I \, 1_{\{ |W_I| \leq C_n \sigma_I \}}$ and let W'_I be the clean version of W_I^*. (See proof of Prop. 3.2.2.) As in Th. 3.2.5 we shall show that

$$W'(n) - W(n) \overset{L^2}{\to} 0, \text{ with } W'(n) = \sum_{|I| = d} W'_I .$$

Then we shall check for W'(n) the conditions of Th. 3.1.2.

$$\operatorname{var}(W(n) - W'(n)) \overset{(1)}{=} \sum_{|I| = d} E\,(W_I - W'_I)^2$$

$$\overset{(2)}{\leq} \sum_{|I| = d} \sigma_I^2 \, E\,(W_I^2 / \sigma_I^2)\, 1_{\{ |W_I| > C_n \sigma_I \}}$$

$$\overset{(3)}{\leq} 3 \int_{(1/2\, C_n, \infty)} x\, R(x)\, dx,$$

where the last term vanishes. Equality (1) follows, since $(W_I - W'_I)$ are components in the Hoeffding decomposition; inequality (2) follows by Prop. 3.2.2 and inequality (3) by assumption a). For W'(n) we shall check the conditions a) and c) of Th. 3.1.2 and condition b) reformulated as $E\, W'(n)^4 - 3 \operatorname{var}^2 W'(n) \to 0$. Then the result follows, since $\operatorname{var} W'(n) \to 1$.

Since W'_I and W'_J are independent if $I \cap J = \emptyset$, we have $\gamma' = 0$. Thus by Prop. 2.3.5

$$| E\, W'(n)^4 - 3 \operatorname{var}^2 W'(n) | \leq 3\, \tau' + \tau'^* + | 3\, S'_0 + S' |.$$

First we check condition a) of Th. 3.1.2. Since $E\,(W_I - W'_I)^2 \leq 3\, \sigma_I^2 \int_{(1/2C_n, \infty)} x\, R(x)\, dx$, we have for n sufficiently large: $2\, \sigma'^2_I \geq \sigma_I^2$ for all I. Then

$$E\,(W'^4_I / \sigma'^4_I) \leq \max W'^2_I / \sigma'^2_I$$

$$\overset{(1)}{\leq} (2^d\, C_n)^2 \sigma_I^2 / \sigma'^2_I$$

$$\leq 2^{2d+1}\, C_n^2,$$

where inequality (1) follows by Prop. 3.2.2. Condition a) follows from

$$C_n^4 \max_i \sum_{I \ni i} \sigma'^2_I \leq C_n^4 \max_i \sum_{I \ni i} \sigma_I^2$$

(by (3.2.6)) and from the choice of C_n^4. Thus τ' and τ'^* vanish for $n \to \infty$ by Prop. 2.4.3 b) and b'). In order to show that $3\, S'_0 + S'$ vanishes it is sufficient (by Corollary 3.1.4) to show that S'_0 vanishes. By Prop. 3.2.1 and assumption b) we have

$$\sum_{\mathcal{B}(0,e)} \sigma_I \sigma_J \sigma_K \sigma_L \to 0 \quad \text{for } e = 1, \ldots, d-1,$$

and for a bifold quadruple (I,J,K,L) we have (Prop. 3.2.4)

$$E\,|W'_I W'_J W'_K W'_L| \le \sigma'_I \sigma'_J \sigma'_K \sigma'_L \le \sigma_I \sigma_J \sigma_K \sigma_L,$$

by (3.2.6). Thus S'_0 vanishes. This proves the theorem.

Th. 3.2.7 above may not be the best possible if there is more information about the form of the components W_I. For example, if we apply Th. 3.2.7 to a multilinear form in independent centered random variables with unit variance and coefficients (a_I) we use the maximal singular value of $(\sigma_I) = (|a|_I)$ and hence neglect the signs of (a_I). The maximal singular value of $(|a|_I)$ may be very different from that of (a_I). Th. 3.2.8 below generalizes Th. 5.3 in De Jong (1987). The proof combines ideas that underlie the theorems 3.2.5 and 3.2.7.

Theorem 3.2.8. Let the family (a_I), with $\sum_{|I|=d} a_I^2 = 1$, have maximal singular value μ^* and

a) $\max_i \sum_{I \ni i} a_I^2 \to 0$ for $n \to \infty$.

Let $X_1, X_2, \ldots$ be iid random variables and $w_n(\ldots)$ a symmetric Borel function $\mathbb{R}^d \to \mathbb{R}$ with $E\, w_n(X_1, x_2, \ldots, x_d) = 0$ and $E\, w_n^2(X_1, \ldots, X_d) = 1$ and $P\{\, |w_n(X_1, \ldots, X_d)| > x\,\} \le R(x)$ with $R(x)$ monotone and not depending on n such that

b) $\int_{(0,\infty)} x\, R(x)\, dx < \infty$.

Put $W_I = w_n(X_{i_1}, \ldots, X_{i_d})$ for $I = \{\, i_1, \ldots, i_d \,\}$ and $W(n) = \sum_{|I|=d} a_I\, W_I$. Then

$$W(n) \xrightarrow{d} N(0,1), \quad n \to \infty$$

if either one of the following conditions is satisfied:

c_1) $\mu^* \to 0$, $n \to \infty$.

c_2) $E\, W_I W_J W_K W_L \to 0$, $n \to \infty$ for all bifold quadruples (I,J,K,L) with $|I \cap J| = e$, $I \cap K = \emptyset$, for $1 \le e \le [d/2]$.

Remark 1. Since the random variables X_i are iid and $w_n(\ldots)$ is symmetric in its arguments, we can restrict condition c_2) to the following [d/2] cases:

$I = \{1, \ldots, d\}$

$K = \{d+1, \ldots, 2d\}$

$L = \{e+1, \ldots, e+d\}$

$J = \{1, \ldots, e, e+d+1, \ldots, 2d\}$, for $1 \le e \le [d/2]$.

Proof. By condition a) the sum W(n) defined above satisfies condition a) of Th. 2.2.3. We proceed as in the proof of Th. 3.2.7, except for some minor changes. Let C_n be a sequence such that

$$C_n \to \infty \text{ and } C_n^4 \max_i \sum_{I \ni i} \sigma_I^2 \to 0 \text{ for } n \to \infty.$$

Then with W'_I the clean version of $W_I\, 1_{\{|W_I| \le C_n\}}$ and $W'(n) = \sum_{|I|=d} a_I\, W'_I$, we have as in the proof of Th. 3.2.7 both

$$W'(n) - W(n) \xrightarrow{L^2} 0, \quad n \to \infty$$

and

$$E\, W'^4_I / \sigma'^4_I \le 2^{2d+1}\, C_n^2, \text{ for n sufficiently large.}$$

We shall show that $W'(n) \xrightarrow{d} N(0,1)$ for $n \to \infty$. Condition a) of Th. 3.1.2 is satisfied by the choice of the constants C_n and the above inequality for the fourth moments. By the independence of the random variables X_i we have $\gamma' = 0$. We check condition b). By the assumptions on W_I we have

$$S(e,0) = \sum_{\mathcal{B}(e,0)} a_I a_J a_K a_L\, E\, W_I W_J W_K W_L$$

$$= E\, W_{I'} W_{J'} W_{K'} W_{L'} \sum_{\mathcal{B}(e,0)} a_I a_J a_K a_L,$$

with (I',J',K',L') as in Remark 1 above. Thus $S_0 \to 0$ under either one of the conditions c_1) and c_2). To show $S'_0 \to 0$ we need $|S_0 - S'_0| \to 0$. With $\sigma''^2_I = E\,(W_I - W'_I)^2$ we have by the triangle inequality and by Prop. 3.2.4 as in the proof of Th. 3.2.5

$$E\, |W_I W_J W_K W_L - W'_I W'_J W'_K W'_L|$$

$$\le \sigma''_I \sigma_J \sigma_K \sigma_L + \sigma'_I \sigma''_J \sigma_K \sigma_L + \sigma'_I \sigma'_J \sigma''_K \sigma_L + \sigma'_I \sigma'_J \sigma'_K \sigma''_L$$

$$\le 4\, \sigma''_I \le 4\, E^{1/2}\, W_I^2\, 1_{\{|W_I| > C_n\}}.$$

By assumption b) the last term vanishes and by Prop. 2.4.3 a) we have

$$\sum_{\mathcal{B}} |a_I a_J a_K a_L| \le C_B(d,4),$$

the number of bifold shadows. Thus $|S_0 - S'_0| \to 0$ for $n \to \infty$. This proves the theorem.

3.3. Inhomogeneous sums

Suppose that V(n) has a finite Hoeffding decomposition, that is, for fixed d we have

$$V(n) - E\,V(n) = \sum_{1\le e\le d} \sum_{|I| = e} W_I .$$

In order to obtain a central limit theorem for V(n) we cannot use Th. 2.1.1. In Example 2 it is shown that the assumption that the sums are homogeneous cannot be removed from Th. 2.1.1. We shall impose stronger conditions on V(n). Write V(n) as a sum of homogeneous sums:

$$V(n) - E\,V(n) = W^{(1)}(n) + \ldots + W^{(d)}(n), \text{ with}$$

$$W^{(e)}(n) = \sum_{|I| = e} W_I .$$

Theorem 3.3.1. Let V(n) and $W^{(e)}(n)$ be as above. Suppose that

$$\lim_{n\to\infty} \operatorname{var} W^{(e)}(n) = \sigma^2(e) \text{ exists and is finite for } 1 \le e \le d.$$

If $W^{(e)}(n) / \operatorname{var}^{1/2} W^{(e)}(n)$ satisfies the conditions of Th. 2.1.1 for each e, $1 \le e \le d$, with $\sigma^2(e) > 0$, then

$$V(n) - E\,V(n) \xrightarrow{d} N(0,\sigma^2(1) + \ldots + \sigma^2(d)), \qquad n \to \infty.$$

Remark 1. If V(n) satisfies the conditions of Th. 3.3.1, then also $V'(n) = \lambda_1 W^{(1)}(n) + \ldots + \lambda_d W^{(d)}(n)$, with λ_e real constants. This shows that the joint distribution of $W^{(1)}(n),\ldots,W^{(d)}(n)$ tends to a multivariate normal distribution with vanishing covariances. Thus the sums $W^{(e)}(n)$ are asymptotically independent (see Billingsley (1968: 49)).

Remark 2. In De Jong (1985) a stronger version of Th. 3.3.1 is given for d = 2: Suppose $V(n) = \sum_{1\le i\le n} W_i + \sum_{|I| = 2} W_I$, with $\sigma^2(1) = \sum_{1\le i\le n} E\, W_i^2$ and $\sigma^2(2) = \sum_{|I| = 2} E\, W_I^2$ satisfies

Ia) $\max_i E\, W_i^2 \to 0, \qquad n \to \infty,$

b) $\sum_{1\le i\le n} W_i \xrightarrow{d} N(0,\sigma^2(1)), \qquad n \to \infty,$

IIa) $\max_i \sum_{I \ni i} \sigma_I^2 \to 0, \qquad n \to \infty,$

b) $E\big(\sum_{|I| = 2} W_I\big)^4 - 3\,\sigma^4(2) \to 0, \quad n \to \infty.$

Then

$$V(n) \xrightarrow{d} N(0, \sigma^2(1) + \sigma^2(2)), \qquad n \to \infty.$$

Th. 3.3.1 follows immediately from Prop. 3.3.2 below. This proposition is the analogue of Prop. 2.3.4 for inhomogeneous sums in the Hoeffding decomposition.

Proposition 3.3.2. Let v(n) and $W^{(e)}(n)$ be as above. Suppose

Ia) $\tau \to 0$, $\qquad n \to \infty$,

b) $\tau^* \to 0$, $\qquad n \to \infty$,

II) $S_0 \to 0$, $\qquad n \to \infty$,

for each $W^{(e)}(n)$ with $\limsup_{n \to \infty} \operatorname{var} W^{(e)}(n) \neq 0$.

Then

$$(V(n) - E\, V(n)) / \operatorname{var}^{1/2} V(n) \xrightarrow{d} N(0,1), \quad n \to \infty.$$

Proof of Th. 3.3.1. Eliminating the vanishing homogeneous sums from V(n) we have, for $V'(n) = \sum_{\sigma^2(e) \neq 0} W^{(e)}(n)$, that $V(n) - V'(n) \xrightarrow{L^2} 0$. For V'(n) we have $E\, W_I^4 \leq D(e)\, \sigma_I^4$ for $|I| = e$ and $\rho(e) = \max_i \sum_{|I| = e, i \in I} \sigma_I^2 \to 0$. By Prop. 2.4.3 we have, with $\rho = \max_e \rho(e)$ and $D = \max_e D(e)$, that τ and τ^* vanish for V'(n). Under the conditions of Th. 2.1.1 S_0 vanishes, as is shown in Sect. 2.3; thus Prop. 3.3.2 implies Th.3.3.1.

Proof of Prop. 3.3.2. Without loss of generality we may assume $\limsup_{n \to \infty} \operatorname{var} W^{(e)}(n) \neq 0$, e = 1,...,d and var V(n) = 1. Define for each e the martingale differences $U_k^{(e)} = \sum_{|I| = e, \max I = k} W_I$ with respect to the increasing σ-algebras $\mathcal{F}_k = \sigma\{X_1, \ldots, X_k\}$. Then $U_k = U_k^{(1)} + \ldots + U_k^{(d)}$ are also martingale differences with respect to $\mathcal{F}_k$. Then by Th.1 in Heyde and Brown (1971) we have the required asymptotic normality for $V(n) = \sum_{1 \leq k \leq n} U_k$ if $\sum_{1 \leq k \leq n} E\, U_k^4 \to 0$ and $\operatorname{var}(\sum_{1 \leq k \leq n} U_k^2) \to 0$ for $n \to \infty$ (cf. Sect. 2.3).

By Hölder's inequality we have

$$\sum_{1 \leq k \leq n} E\, U_k^4 = \sum_{1 \leq k \leq n} E\, (\sum_{1 \leq e \leq d} U_k^{(e)})^4$$

$$\leq \sum_{1\leq k\leq n} E\,((\sum_{1\leq e\leq d} 1^{4/3})^{3/4}(\sum_{1\leq e\leq d}(U_k^{(e)})^4)^{1/4})^4$$

$$= d^3 \sum_{1\leq e\leq d}\sum_{1\leq k\leq n} E\,(U_k^{(e)})^4,$$

where the final term vanishes by the assumptions on $W^{(e)}(n)$ and Prop. 2.3.1, since $\tau^* \to 0$. By the Cauchy-Schwarz inequality we have

$$\mathrm{var}\,(\sum_{1\leq k\leq n} U_k^2) = E\,(\sum_{1\leq e,f\leq d}\sum_{1\leq k\leq n}(U_k^{(e)}U_k^{(f)} - E\,U_k^{(e)}U_k^{(f)}))^2$$

$$\leq E\,(\sum_{1\leq e,f\leq d} 1^2)\,(\sum_{1\leq e,f\leq d}(\sum_{1\leq k\leq n}(U_k^{(e)}U_k^{(f)} - E\,U_k^{(e)}U_k^{(f)}))^2)$$

$$= d^2 \sum_{1\leq e,f\leq d} \mathrm{var}\,(\sum_{1\leq k\leq n}(U_k^{(e)}U_k^{(f)}).$$

For the homogeneous case we have, combining Prop. 2.3.2 and Prop. 2.3.3,

$$(3.3.1)\quad \mathrm{var}\,(\sum_{1\leq k\leq n}(U_k^{(e)})^2) = \mathrm{var}\,(\sum_{|I| = |J| = e, \max I\cup J \in I\cap J} W_I W_J)$$

$$\leq 2\,(\tau + \tau^* + |\,S + 2S_0\,| + |\,2/3\,S + S_0\,|),$$

where the right-hand side vanishes under the conditions of Prop 3.3.2. Notice that, by Corollary 3.1.4, S vanishes if τ^* and S_0 vanish.

For the inhomogeneous case (e < f) we have to show, since $E\,W_I W_J = 0$,

$$E\,(\sum_{|I| = e,\ |J| = f, \max I\cup J \in I\cap J} W_I W_J)^2 \to 0.$$

This will follow from Prop. 3.3.3 and Prop.3.3.4. With

$$A = \sum_{|I| = e,\ |J| = f, \max I\cup J \in I\cap J} W_I W_J$$

and

$$B = \sum_{|I| = e,\ |J| = f, \max I\cup J \notin I\cap J} W_I W_J$$

we have $E\,(A + B)\,A - 1/2\,E\,A^2 \to 0$ (by Prop. 3.3.4), hence $E\,(A + B)^2 - E\,B^2 \to 0$. By Prop. 3.3.3 we have $E\,(A + B)^2 \to 0$ and hence $E\,B^2 \to 0$ and $E\,A^2 \to 0$. This proves the proposition.

Proposition 3.3.3. Under the conditions of Prop.3.3.2, with $|\,I\,| = e < f = |\,J\,|$, we have

$$E\,(\sum_{I\cap J \neq \emptyset} W_I\, W_J)^2 \to 0.$$

Proof. Throughout the proof we shall assume $|I| = |K| = e < f = |J| = |L|$. The remainder terms R_i below are sums over subsets of $\mathcal{T}$ and vanish by $|R_i| \leq \tau^*$.

$$E\Big(\sum_{I\cap J\neq\emptyset} W_I W_J\Big)^2 = \sum_{\mathcal{B},\, I\cap J\neq\emptyset} E\, W_I W_J W_K W_L + R_1$$

$$= \sum_{\mathcal{B},\, J\neq L} E\, W_I W_J W_K W_L - \sum_{\mathcal{B},\, J\neq L,\, I\cap J=\emptyset} E\, W_I W_J W_K W_L + R_1,$$

where the first equality follows since in a bifold quadruple $I \Delta J = K \Delta L$ and hence $|I \cap J| = |K \cap L|$ and the last equality follows, since in a bifold quadruple $I \cap J \neq \emptyset$ implies $J \neq L$. We shall show that the last sum is, up to a vanishing remainder term, a sum of squares and hence (asymptotically) non-negative. Then, since the left-hand side is non-negative, it remains to show that the first sum in the final expression above vanishes.

We start with the last sum.

$$\sum_{A\subset\{1,\ldots,n\},\, |A|\leq f-1} E\Big(\sum_{I,L,\, L\setminus I = A} W_I W_L\Big)^2$$

$$= \sum_{A\subset\{1,\ldots,n\},\, |A|\leq f-1} \;\sum_{\mathcal{B},\, J\cap L = A,\, I\cap J=\emptyset} E\, W_I W_J W_K W_L + R_2$$

$$= \sum_{\mathcal{B},\, J\neq L,\, I\cap J=\emptyset} E\, W_I W_J W_K W_L + R_2,$$

where the last inequality follows from $|A| \leq f-1$.

To show that the first sum vanishes we write

$$E\Big(\sum_{I,K} W_I W_K\Big)\Big(\sum_{J\neq L,\, J\cap L\neq\emptyset} W_J W_L\Big) = \sum_{\mathcal{B},\, J\neq L} E\, W_I W_J W_K W_L + R_3.$$

The first factor on the left-hand side has a bounded second moment

$$E\Big(\sum_{I,K} W_I W_K\Big)^2 = E\big(W^{(e)}(n)\big)^4 \leq \tau^* + 3\tau + |S + S_0| + 3\,\mathrm{var}^2 W^{(e)}(n),$$

by (2.3.2), with the right-hand side bounded (since under the assumptions on $W^{(e)}(n)$, all terms vanish except $3\,\mathrm{var}^2 W^{(e)}(n) \leq 3$). For the second factor we have by the Cauchy-Schwarz inequality

$$E\Big(\sum_{J\neq L,\, J\cap L\neq\emptyset} W_J W_L\Big)^2 \leq 2\Big(\mathrm{var}\Big(\sum_{J\cap L\neq\emptyset} W_J W_L\Big) + \mathrm{var}\Big(\sum_{|J|=f} W_J^2\Big)\Big).$$

By Prop. 2.3.2. and Prop. 3.1.1 the right hand side vanishes. Now the proposition follows by Cauchy-Schwarz.

Proposition 3.3.4. Under the conditions of Prop. 3.3.2 we have, with $|I|=|K|= e < f = |J|=|L|$

$$E\Big(\sum_{I\cap J\neq\emptyset} W_I W_J\Big)\Big(\sum_{\max K\cup L\in K\cap L} W_K W_L\Big) - 1/2\, E\Big(\sum_{\max I\cup J\in I\cap J} W_I W_J\Big)^2 \to 0.$$

Proof. The quantities R_i below are sums over subsets of $\mathcal{T}$ and can be estimated by $|R_i| \le \tau^*$.

$$E\Big(\sum_{I\cap J\neq\emptyset} W_I W_J\Big)\Big(\sum_{\max K\cup L\in K\cap L} W_K W_L\Big)$$

$$= \sum_{\mathcal{B},\, \max K\cup L\in K\cap L} E\, W_I W_J W_K W_L + R_1$$

$$= \sum_{\mathcal{B},\, K\cap L\ni \max K\cup L > \max I\cup J} E\, W_I W_J W_K W_L$$

$$+ \sum_{\mathcal{B},\, K\cap L\ni \max K\cup L < \max I\cup J\in I\cap J} E\, W_I W_J W_K W_L + R_1.$$

Notice that $\max I\cup J > \max K\cup L$ implies $\max I\cup J\in I\cap J$ for bifold quadruples.

Consider the second sum in the assertion of the proposition:

$$E\Big(\sum_{\max I\cup J\in I\cap J} W_I W_J\Big)^2$$

$$= 2 \sum_{\mathcal{B},\, K\cap L\ni \max K\cup L < \max I\cup J\in I\cap J} E\, W_I W_J W_K W_L + R_2.$$

To prove the proposition we have to show that the following sum vanishes.

$$\sum_{\mathcal{B},\, K\cap L\ni \max K\cup L > \max I\cup J} E\, W_I W_J W_K W_L$$

$$= E\Big(\sum_{\max I\cup K\in K\setminus I} W_I W_K\Big)\Big(\sum_{\max J\cup L\in L\setminus J,\, J\cap L\neq\emptyset} W_J W_L\Big) - R_3.$$

The second moment of the first factor in the cross product above

$$E\Big(\sum_{\max I\cup K\in K\setminus I} W_I W_K\Big)^2$$

$$= 1/4\, E\Big(\sum_{\max I\cup K\notin K\cap I} W_I W_K\Big)^2$$

$$= 1/4\, E\Big(\sum_{I,K} W_I W_K - \sum_{\max I\cup K\in K\cap I} W_I W_K\Big)^2$$

$$= 1/4\, E\Big((W^{(e)}(n))^2 - \sum_{1\le k\le n} (U_k^{(e)})^2\Big)^2$$

remains bounded, since $E\,(W^{(e)}(n))^4$ remains bounded and since

$$E \sum_{1\le k\le n} (U_k^{(e)})^2 = \operatorname{var} W^{(e)}(n) \le 1 \text{ and } \operatorname{var}\Big(\sum_{1\le k\le n} (U_k^{(e)})^2 \Big) \to 0,$$

under the assumptions on $W^{(e)}(n)$ by (3.3.1). Further, the second moment of the second factor

$$E\Big(\sum_{\max J\cup L \in L\setminus J,\, J\cap L \neq \emptyset} W_J W_L\Big)^2 = 1/4\, E\Big(\sum_{\max J\cup L \notin J\cap L \neq \emptyset} W_J W_L\Big)^2$$

vanishes by the assumptions on $W^{(f)}(n)$ and Prop. 2.3.3. Now the proposition follows by Cauchy-Schwarz.

The example below concerns d-linear forms in iid random variables

$$Z(n) = \sum_{|I| = d} a_I \prod_{i \in I} X_i, \quad X_i \text{ iid.}$$

These multilinear forms have a simple Hoeffding decomposition as can be seen from the following proposition.

Proposition 3.3.5. Let Z be a d-linear form

$$Z = \sum_{|I| = d} a_I \prod_{i\in I} X_i,$$

with X_i independent random variables $E X_i = E X_1$, $\operatorname{var} X_i = \operatorname{var} X_1$. Then Z has the following Hoeffding decomposition:

$$Z - E Z = \sum_{|J| = 1} W_J + \ldots + \sum_{|J| = d} W_J,$$

with for $|J| = e$

$$W_J = (E X_1)^{d-e} \prod_{j\in J} (X_j - E X_1) \sum_{|I| = d,\, J\subset I} a_I, \text{ and}$$

$$\operatorname{var}\Big(\sum_{|J| = e} W_J \Big) = (E X_1)^{2d-2e} (\operatorname{var} X_1)^e \sum_{|J| = e} \Big(\sum_{|I| = d,\, J\subset I} a_I \Big)^2 .$$

Proof. Apply the Hoeffding decomposition to $a_I \prod_{i\in I} X_i$:

$$a_I \prod_{i\in I} X_i = a_I \sum_{J\subset I} (E X_1)^{d - |J|} \prod_{j\in J} (X_j - E X_1).$$

Example. Consider the d-linear ($d \ge 2$) form $Z(n)$ in iid zero-one random variables $X_i \in \{0,1\}$, $E X_i = p_n \le 1/2$. Assume the family (a_I) to be of a very simple form, with $a_I \in \{0,1\}$ subject to the condition

$$\max_{J \subset \{1,\dots,n\},\, |J| = d-1} \sum_{I \supset J} a_I = \mathcal{O}(n\,\rho_n),$$

with ρ_n the mean number of ones in (a_I):

$$\rho_n = \binom{n}{d}^{-1} \sum_{|I| = d} a_I .$$

We use $f(n) = \mathcal{O}(g(n))$ to denote $| f(n) / g(n) | \le C$ for all n and some constant C. Thus for each one-dimensional row $\{I : |I| = d, I \supset J\}$ for fixed J with $|J| = d-1$ the number of a_I equal one is bounded by $C\, n\rho_n$. Notice that the condition on the family (a_I) implies that $a_I = 0$ for all I if $n\rho_n \to 0$, since $a_I \in \{0,1\}$. In fact we have more; since

$$\max_{J \subset \{1,\dots,n\},\, |J|=d-1} \sum_{I \supset J} a_I \ \ge\ \binom{n}{d-1}^{-1} \sum_{|J| = d-1} \sum_{I \supset J} a_I$$

$$= \binom{n}{d-1}^{-1} d \sum_{|I| = d} a_I$$

$$= (n - d + 1)\, \rho_n,$$

we have

$$\max_{J \subset \{1,\dots,n\},\, |J|=d-1} \sum_{I \supset J} a_I \ \propto\ n\,\rho_n,$$

with $f(n) \propto g(n)$ if $1/C \ \le\ | f(n) / g(n) | \ \le\ C$ for all n and some constant $C \ge 1$.

The behaviour of Z(n) can be described in terms of the two parameters p_n and ρ_n. Notice that $E\, Z(n) = \binom{n}{d} \rho_n (p_n)^d$. Hence $\rho_n (np_n)^d \to 0$ implies $Z(n) \xrightarrow{L^1} 0$. We shall assume $\rho_n (np_n)^d \to \infty$. Indeed, we shall show in Prop. 3.3.6 that $\rho_n (np_n)^d \to \infty$ is sufficient for a normal limit distribution of Z(n). Notice that $\rho_n (np_n)^d \to \infty$ implies $np_n \to \infty$, since $0 \le \rho_n \le 1$.

The simple structure of the family (a_I) has an interesting consequence. For fixed J with $|J| = e$ we have

$$\text{(3.3.2)} \qquad \sum_{I \supset J} a_I = \begin{cases} \mathcal{O}(n^{d-e}\rho_n) & \text{if } 1 \le e < d, \\ a_J & \text{if } e = d. \end{cases}$$

This 'discontinuity' at d = e gives, since $p_n \le 1/2$,

$$\operatorname{var}\Big(\sum_{|J| = e} W_J \Big) \propto \begin{cases} (np_n)^{2d-e}\rho_n^2 & \text{if } 1 \le e < d, \\ (np_n)^d \rho_n & \text{if } e = d, \end{cases}$$

where the first estimate follows from Prop. 3.3.5 and the bound on $\sum_{I \supset J} a_I$ for the upperbound, and from the inequality

$$\sum_{|J|=e} \Big(\sum_{I \supset J} a_I - \tbinom{n-e}{d-e} \rho_n \Big)^2 = \sum_{|J|=e} \Big(\sum_{I \supset J} a_I \Big)^2 - \tbinom{n}{e} \Big(\tbinom{n-e}{d-e} \rho_n \Big)^2 \geq 0$$

for the lowerbound. Hence we have for $1 < e < d$

$$\operatorname{var}\Big(\sum_{|J|=e} W_J \Big) \,/\, \operatorname{var}\Big(\sum_{|J|=1} W_J \Big) \;\propto\; (np_n)^{1-e} \to 0.$$

Thus Z(n) can be written as the orthogonal sum of only two (instead of d) homogeneous sums plus a remainder term:

$$Z(n) - E\,Z(n) = \sum_{|J|=1} W_J + \sum_{|J|=d} W_J + R(n),$$

with $\operatorname{var} R(n) / \operatorname{var} Z(n) \to 0$ for $n \to \infty$.
Since

$$\operatorname{var}\Big(\sum_{|J|=1} W_J \Big) / \operatorname{var}\Big(\sum_{|J|=d} W_J \Big) \;\propto\; (np_n)^{d-1} \rho_n,$$

the random variable $Z(n) - E\,Z(n)$ is approximately: 1) a sum of independent random variables if $(np_n)^{d-1}\rho_n \to \infty$, 2) a d-homogeneous sum if $(np_n)^{d-1}\rho_n \to 0$, and 3) a mixture of the two above cases else.

Notice that $(np_n)^{d-1}\rho_n \leq C$ implies $\rho_n \to 0$, since we assume $np_n \to \infty$ and $d > 1$. The speed at which ρ_n vanishes depends on p_n and the value of d. We have the following bound for the 'number of ones per one-dimensional row': $n\rho_n \leq C\, n^{2-d} p_n^{1-d}$. E.g. for $d = 2$ this yields $n\rho_n \leq C\, p_n^{-1}$.

Proposition 3.3.6. Let Z(n), p_n and ρ_n be as above. If $(np_n)^d \rho_n \to \infty$, then

$$(Z(n) - E\,Z(n)) / \operatorname{var}^{1/2} Z(n) \xrightarrow{d} N(0,1), \quad n \to \infty.$$

Remark. The excluded case $d = 1$ falls under the above proposition, since a sum of independent bounded random variables with diverging total variance has a normal limit distribution.

Proof. As shown above we may assume without loss of generality:

$$Z(n) - E\,Z(n) = \sum_{1 \leq i \leq n} b_i\,(X_i - p_n) + \sum_{|I|=d} W_I,$$

with

$$b_i = (p_n)^{d-1} \sum_{I \ni i} a_I,$$

$$W_I = a_I \prod_{i \in I} (X_i - p_n),$$

and

$$\operatorname{var} Z(n) \propto (np_n)^d \rho_n \; (1 + (np_n)^{d-1} \rho_n).$$

We shall apply Prop. 3.3.2. The main part of the proof consists in showing that τ and τ^* vanish. This will be done by direct computation using the special regularity of the family (a_I) and the fact that the random variables are iid. The theory developed in Sect. 2.4 for estimating the quantities τ and τ^* does not (or at most partially) apply.

We start with some special partial sums of τ and τ^*. First we consider homogeneous quadruples in $\mathcal{T}$. Since

$$\text{(3.3.3)} \quad E\,|X_i - p_n|^\alpha \propto p_n \quad \text{for } \alpha \geq 1,$$

we find by (3.3.2):

$$\sum_{1\leq i\leq n} b_i^4 \, E\,(X_i - p_n)^4 / \operatorname{var}^2 Z(n)$$

$$= O(\; (np_n)^{4d-3} \rho_n^4 \; / \; (np_n)^{2d} \rho_n^2 \; (1 + (np_n)^{d-1} \rho_n)^2 \,)$$

$$= O(\; (np_n)^{-1} \, ((np_n)^{d-1} \rho_n \; / \; (1 + (np_n)^{d-1} \rho_n) \,)^2 \,),$$

which vanishes since $np_n \to \infty$. Thus the partial sum of τ^* of homogeneous quadruples with $|I| = 1$ vanishes, as does the corresponding partial sum of τ, since $E^2\,(X_i - p_n)^2 \leq E\,(X_i - p_n)^4$. Next we shall estimate partial sums of τ and τ^* of homogeneous quadruples with $|I| = d$.

By (3.3.3) we have

$$\sigma_I^2 \propto a_I \; (p_n)^d$$

and for a quadruple with $f = |I \cup J \cup K \cup L|$ we have

$$E\,|W_I W_J W_K W_L| \propto a_I a_J a_K a_L \; (p_n)^f,$$

$$\sigma_I \sigma_J \sigma_K \sigma_L \propto a_I a_J a_K a_L \, (p_n)^{2d}.$$

This yields for a quadruple without a free index (and consequently $f \leq 2d$)

$$\sigma_I \sigma_J \sigma_K \sigma_L = O(\, E\,|W_I W_J W_K W_L| \,).$$

Thus we only need to show that partial sums of τ^* of homogeneous quadruples with $|I| = d$ vanish. We start with a special sum.

$$\sum_{|I| = d} E\, W_I^4 / \operatorname{var}^2 Z(n)$$

$$= O((np_n)^d \rho_n \; / \; (np_n)^{2d} \rho_n^2 \; (1 + (np_n)^{d-1} \rho_n)^2 \,)$$

$$= O(((np_n)^d \rho_n \, (1 + (np_n)^{d-1} \rho_n)^2)^{-1}),$$

which vanishes by the assumption $(np_n)^d \rho_n \to 0$.

Consider a shadow (I',J',K',L') with f = | I'∪J'∪K'∪L'|, d < f < 2d. Without restriction we assume | I'∩J' | = e < d. Thus we have | (K'∪L') \ (I'∪J') | = f - 2d + e. Then

$$\sum_{(I,J,K,L) \text{ with shadow } (I',J',K',L')} a_I a_J a_K a_L$$

$$\leq \sum_{(I,J,K,L),\ |I\cap J| = e,\ |I\cup J\cup K\cup L| = f} a_I a_J a_K a_L$$

$$= O(n^{f-2d+e} \sum_{|I\cap J| = e} a_I a_J).$$

Since by (3.3.2)

$$\sum_{|I\cap J| = e} a_I a_J = O(\rho_n^2 n^{2d-e}),$$

we have by (3.3.3)

$$\sum_{(I,J,K,L) \text{ with shadow } (I',J',K',L')} E\,|\,W_I W_J W_K W_L\,| \,/\, \mathrm{var}^2 Z(n)$$

$$= O(\rho_n^2 (np_n)^f / (np_n)^{2d} \rho_n^2 \ (1 + (np_n)^{d-1} \rho_n)^2)$$

$$= O((np_n)^{f-2d} / (1 + (np_n)^{d-1} \rho_n)^2),$$

which vanishes since f < 2d.

Thus the contribution of all homogeneous quadruples to τ and τ^* vanishes. We shall estimate the contribution of the mixed quadruples with indices not all containing the same number of elements by the contribution of the homogeneous quadruples. Define

$$V_I = \begin{cases} W_I & \text{if } |\,I\,| = d \\ b_i (X_i - p_n) & \text{if } I = \{i\} \end{cases}$$

and

$$c_I = \begin{cases} a_I & \text{if } |\,I\,| = d \\ b_i & \text{if } I = \{i\}. \end{cases}$$

Then, since $E\,V_I^2 \propto c_I^2 (p_n)^{|I|}$, we have

$$\sigma_I \sigma_J \sigma_K \sigma_L \propto c_I c_J c_K c_L \ (p_n)^{1/2\,(|I|+|J|+|K|+|L|)},$$

and by (3.3.3)

$$E\,|\,V_I V_J V_K V_L\,| \propto c_I c_J c_K c_L \ (p_n)^{|I\cup J\cup K\cup L|}.$$

By the definition of $\mathcal{T}$ we have 1/2 (|I| + |J| + |K| + |L|) > | I∪J∪K∪L | and thus

$$\sigma_I \sigma_J \sigma_K \sigma_L = O(E \mid V_I V_J V_K V_L \mid).$$

Consider a mixed quadruple (I_1,I_2,I_3,I_4) in $\mathcal{T}$. Then any index with only one element is contained in some index with d elements. For each such mixed quadruple we construct a homogeneous quadruple (J_1,J_2,J_3,J_4) also in $\mathcal{T}$ with $I_1 \cup I_2 \cup I_3 \cup I_4 = J_1 \cup J_2 \cup J_3 \cup J_4$ by

$$J_g = \begin{cases} I_g & \text{if } |I_g| = d \\ I_h & \text{if } |I_g| = 1, |I_h| = d, I_g \subset I_h, \text{ h minimal.} \end{cases}$$

Then with $b^* = \max_i b_i$ and α the number of indices with one element in (I_1,I_2,I_3,I_4) we have

$$c_{I_1} c_{I_2} c_{I_3} c_{I_4} \le a_{J_1} a_{J_2} a_{J_3} a_{J_4} (b^*)^{\alpha}.$$

For a fixed shadow (I',J',K',L') we have

$$\sum_{(I,J,K,L) \text{ with shadow } (I',J',K',L')} E \mid W_I W_J W_K W_L \mid / \operatorname{var}^2 Z(n)$$

$$= O(\rho_n^2 (np_n)^f (b^*)^{\alpha} \; / \; (np_n)^{2d} \rho_n^2 \; (1 + (np_n)^{d-1} \rho_n)^2)$$

$$= O((np_n)^{f-2d} \; ((np_n)^{d-1} \rho_n)^{\alpha} \; / \; (1 + (np_n)^{d-1} \rho_n)^2).$$

If $\alpha \le 2$ then the above estimate vanishes since f < 2d. If $\alpha = 3$ then f = d and we have the estimate

$$(np_n)^{-1} \rho_n \; ((np_n)^{d-1} \rho_n \; / \; (1 + (np_n)^{d-1} \rho_n))^2,$$

which vanishes since $np_n \to \infty$. This shows that τ and τ^* both vanish.

Condition II of Prop. 3.3.2 remains to be checked. For a sum of independent random variables we have $S_0 = 0$. For a bifold quadruple (I,J,K,L) we have $E\ W_I W_J W_K W_L = a_I a_J a_K a_L\ (E\ (X_i - p_n)^2)^{2d} = \sigma_I \sigma_J \sigma_K \sigma_L$. In the proof of Th. 3.1.5 it is shown that condition a) of Th. 3.1.5 implies for clean sums with var W(n) = 1 that $\sum_{\mathcal{B}(0,e)} \sigma_I \sigma_J \sigma_K \sigma_L$ vanishes. We shall check condition a) for $\sum_{|I| = d} W_I \,/\, \operatorname{var}^{1/2} Z(n)$.

By (3.3.2) we have

$$\max_{A \subset \{1,\dots,n\},\ 1 \le |A| \le d-1} \; \sum_{I \supset A} \sigma_I \sum_{J \supset I \setminus A} \sigma_J \; / \operatorname{var} Z(n)$$

$$= O(\rho_n^2 (np_n)^d / \; (np_n)^d \rho_n \; (1 + (np_n)^{d-1} \rho_n))$$

$$= O((np_n)^{1-d} \; ((np_n)^{d-1} \rho_n \; / \; (1 + (np_n)^{d-1} \rho_n)))$$

$$= O((np_n)^{1-d}),$$

which vanishes for d ≥ 2. This proves the proposition.

We shall use the above example to test some of the conditions of the different central limit theorems given above. We start with condition b) of Th. 2.1.1: $E\,W_I^4 / \sigma_I^4 \le D$. It may seem reasonable to impose this condition. However, for the homogeneous components in the Hoeffding decomposition of the multilinear form above it is a severe restriction:

$$E\,W_I^4 / \sigma_I^4 \propto p_n^d / p_n^{2d} = p_n^{-d} \quad (\text{if } a_I \neq 0),$$

which is not bounded if $p_n \to 0$. We have given two ways to circumvent (partially) condition b).

In the first place we have truncation (cf. Prop. 3.2.2) However, truncation is not very useful in case of zero-one valued random variables: Consider a zero-one valued random variable X, with $E\,X = p_n$ ($\text{var}\,X = p_n(1 - p_n)$) The variance of the truncated version

$$\text{var}\,(X\,1_{\{X \le C\,p_n^{1/2}\}} \propto p_n^2 \text{ for } n \to \infty,$$

which vanishes with respect to var X.

Th. 3.1.2 combines the conditions a) and b) of Th. 2.1.1 allowing D to diverge in a controlled way:

$$(\max_I E\,W_I^4 / \sigma_I^4)\,(\max_i \sum_{i \ni I} \sigma_I^2)^{1/2} \to 0 \text{ for } n \to \infty.$$

In the case of the homogeneous components in the Hoeffding decomposition of the multilinear form above we have with $\sigma_I^2 = E\,W_I^2 / \text{var}\,(\sum_{|I| = d} W_I^2)$

$$\max_i \sum_{i \ni I} \sigma_I^2 \propto p_n^d\, n^{d-1}\, \rho_n / (p_n n)^d \rho_n = 1/n.$$

Thus condition a) of Th. 3.1.2 is satisfied if $np_n^{2d} \to \infty$. The latter condition implies in combination with $(p_n n)^{d-1}\rho_n \le C$ (which ensures that the variance of the d-homogeneous sum does not vanish) that the family (a_I) contains elements $a_I = 1$, only if d = 2. If d = 2 we have

$$n\rho_n \le C\,p_n^{-1};$$

if d ≥ 3 we have

$$n\rho_n \le C\,p_n^{d-1}\,n^{2-d} = C\,(n\,p_n^{(d-1)/(d-2)})^{2-d}.$$

Thus $n\rho_n$ vanishes, since $n\,p_n^{(d-1)/(d-2)} \to \infty$ for $d \ge 3$.

In the proof of Prop. 3.3.6 we did not check condition c) of Th.2.1.1 directly. Instead we checked condition a) of Th. 3.1.4. In Sect. 3.1 we also mentioned a more restrictive form of condition a)

$$(3.1.1) \quad \max_i \sum_{I \ni i} \sigma_I \to 0 \text{ for } n \to \infty.$$

If we apply (3.1.1) in the example above we find with $\sigma_I^2 = E\, W_I^2 / \operatorname{var}(\sum_{|I| = d} W_I^2)$:

$$\max_i \sum_{I \ni i} \sigma_I = O(n^{d-1} \rho_n (p_n)^{d/2} / ((np_n)^d \rho_n) p_n^{1/2})$$

$$= O(n^{d/2 - 1} \rho_n^{1/2}),$$

which vanishes if $n^{d-2} \rho_n \to 0$. Again for $d \geq 3$ this implies $n\rho_n \to 0$.

4. W(n) as a Gaussian process

4.0. Introduction

In this chapter we shall we restrict ourselves to homogeneous sums in the Hoeffding decomposition, W(n), with respect to one given sequence $X_1, X_2, \ldots$ of independent random variables. If we mention several d-homogeneous sums (for the same n), these sums are all defined with respect to this one sequence of independent random variables. The main aim of this chapter is to show how the random variables W_I can be embedded as random point masses at points x_I in a suitable product space E^d. Given this embedding, a fairly broad class of functions f is identified such that the stochastic integral

$$\int f \, dW(n) = \sum_{|I| = d} f(x_I) W_I$$

converges to a stochastic integral with respect to a Gaussian process with independent increments. We shall follow the usual approach for defining integrals; we define at first the integral for stepfunctions. Indeed, once the problems with the stochastic integral for stepfunctions have been solved, the extension is by standard approximation techniques.

A stepfunction $f = a_1 1_{A_1} + \ldots + a_q 1_{A_q}$ (A_g disjoint) partitions the space into finitely many subsets. The distribution of the stochastic integral $\int f \, dW(n)$ is the distribution of the linear combination of the partial sums $W^{(g)}(n) = \sum_{I \in \mathcal{A}_g} W_I$ for subsets of the indices $\mathcal{A}_g = \{ I : x_I \in A_g \}$.

Section 4.1 is concerned with questions regarding the distribution of partial sums of W(n), with W(n) homogeneous and satisfying the conditions of Th.2.1.1. Clearly, any partial sum inherits the conditions a) and b) of Th. 2.1.1, since

$$\text{a)} \max_i \sum_{I \ni i, I \in \mathcal{A}} \sigma_I^2 \le \max_i \sum_{I \ni i} \sigma_I^2 ,$$

$$\text{b)} \max_{I \in \mathcal{A}} E\, W_I^4 / \sigma_I^4 \le \max_I E\, W_I^4 / \sigma_I^4 .$$

Condition c) (in the alternative formulation $E\, W'(n)^4 - 3 \operatorname{var}^2 W'(n) \to 0$, $n \to \infty$) need not to be satisfied, as is shown at the end of this section. Thus some caution is needed in constructing the subsets A in such a way that partial sums over A inherit condition c). This construction is carried out in Sect. 4.1 in a rather unexpected way. Instead of using directly a geometrical approach, e.g. devide the index set $\{I: I \subset \{1,\ldots,n\}\}$ into appropriate 'blocks' and consider the (joint) distribution of these blocks, we define two algebraic operations on W(n). Addition is defined in the usual way: W(n) = W'(n) +

W''(n), with W'(n) and W''(n) homogeneous sums in the Hoeffding decomposition. In Th. 4.1.1 it is shown that W(n) has a normal limit distribution if W'(n) and W''(n) both satisfy the conditions of Th. 2.1.1. Further we define 'scalar' multiplication. We consider products of a homogeneous sum W(n) and a family (a_I) of real constants defined by

$$W'(n) = \sum_{|I| = d} a_I W_I .$$

If the family (a_I) has a special structure, then W'(n) satisfies the conditions of Th. 2.1.1, provided W(n) satisfies these conditions. Using these operations we define a class of subsets of the indices such that the partial sums over these subsets have normal limit distribution.

In Sect. 4.2 these results are extended. With the help of the subsets, constructed in Sect. 4.1, stepfunctions are defined on the product space E^d, such that the stochastic integrals of the stepfunctions converge in distribution. By approximation the class of stepfunctions is extended.

In Sect 4.3 a simple condition on W(n) is given, ensuring that any partial sum W'(n) has a normal limit distribution, provided the variance converges. Homogeneous sums satisfying this condition can also be embedded. However, the embedding has a different character. Whereas in the general case the product structure of E^d is used, this structure is not needed to embed homogeneous sums W(n) of the restricted class considered in Sect. 4.3. By the simple extra condition imposed on W(n) the 'dimension of W(n) is lost'.

Example. We shall construct a sequence of matrices $(a_{ij})_{1\leq i,j\leq n}$ with eigenvalues λ_i such that max $\lambda_i^2 / (\lambda_1^2 + ... + \lambda_n^2) \to 0$, $n \to \infty$. Then, as is shown in the introduction, the fourth moment of the quadratic form in independent normal N(0,1) random variables with respect to (a_{ij}) tends to three for $n \to \infty$. We shall give a subset $\mathcal{A}$ of the indices (i,j) such that the maximal eigenvalue of the matrices $(a''_{ij})_{1\leq i,j\leq n} = (a_{ij})_{1\leq i,j\leq n,\, (i,j) \in \mathcal{A}}$ does not vanish. Then the quadratic form W'(n) with respect to (a''_{ij}) does not satisfy condition c) of Th. 2.1.1:

$$\lim_{n \to \infty} (E\, W'(n)^4 - 3\, \mathrm{var}^2\, W'(n)) \neq 0,$$

and hence W'(n) does not have a normal limit distribution.

We construct a symmetric $2^m \times 2^m$ matrix with all entries ± 1 and all rows mutually orthogonal vectors. We start with the construction of the 2^m orthogonal vectors; then we give an enumeration of these vectors that makes the matrix symmetric. Let $e_i(x)$: $[0,1) \to \{0,1\}$ be the ith digit in the binary expansion of $x \in [0,1)$. To make the representation unique take $e_i(x)$ to be right continuous. The functions $r_i = 2e_i - 1$

($-r_i$ is the ith Rademacher function) are independent with respect to the Lebesgue measure, and $r_i^2 = 1$. Notice that

$$r_i = (-1)^{(1-e_i)}.$$

Define for the set $A \subset \{1,...,m\}$

$$f_A = \prod_{g \in A} r_g = \prod_{1 \le g \le m} (r_g)^{1_A(g)} = \prod_{1 \le g \le m} (-1)^{(1-e_g)\, 1_A(g)} .$$

There are 2^m different subsets $A_k \subset \{1,...,m\}$, $k = 1,...,2^m$. For two subsets A, A' we have by independence

$$\int_{[0,1)} f_A(x)\, f_{A'}(x)\, dx = \prod_{i \in A \,\Delta\, A'} \int_{[0,1)} r_i(x) dx = \begin{cases} 1 & \text{if } A = A', \\ 0 & \text{else.} \end{cases}$$

Define for each set A_i the vector with jth entry

$$a_{ij} = f_{A_i}\left(\frac{j-1}{2^m}\right).$$

These vectors are mutually orthogonal:

$$\sum_{1 \le k \le 2^m} a_{ik} a_{jk} = 2^m \int_{[0,1)} f_{A_i}(x)\, f_{A_j}(x)\, dx = 0 \quad \text{if } i \neq j.$$

With the following definition of the set A_i we obtain a symmetric matrix with ith row the vector defined by A_i:

$$A_i = \{g : r_g\left(\frac{i-1}{2^m}\right) = -1\}.$$

Since

$$1_{A_i}(g) = 1 - e_g\left(\frac{i-1}{2^m}\right),$$

we have

$$a_{ij} = \prod_{1 \le g \le m} (-1)^{\left(1-e_g\left(\frac{j-1}{2^m}\right)\right) 1_{A_i}(g)} = \prod_{1 \le g \le m} (-1)^{\left(1-e_g\left(\frac{j-1}{2^m}\right)\right)\left(1-e_g\left(\frac{i-1}{2^m}\right)\right)} .$$

(Thus $a_{ij} = -1$ if the number of zeros in the same position in the binary expansion of (j-1) and (i-1) is odd; and $a_{ij} = 1$ else.)

Thus the real matrix (a_{ij}) is symmetric and orthogonal and has real eigenvalues λ_i with $|\lambda_i| = 2^{m/2}$, $i = 1,...,2^m$ and hence $\max \lambda_i^2 / (\lambda_1^2 + ... + \lambda_n^2) \to 0$ for $n \to \infty$. We need for our example a matrix with zero diagonal. Define the real symmetric matrix (a'_{ij}), with $a'_{ij} = a_{ij}$ if $i \neq j$, $a'_{ii} = 0$, and with eigenvalues λ'_i. Then we have

$$\sum_{1 \le i \le 2^m} \lambda'^2_i = \sum_{1 \le i,j \le 2^m} a'^2_{ij} = 2^{2m} - 2^m \text{ and } \max_i |\lambda'_i| \le \max_i |\lambda_i| + 1$$

by the triangle inequality for matrix norms, since the matrix $(a_{ij} - a'_{ij})$ is a diagonal matrix with diagonal entries ± 1 and hence with maximal absolute eigenvalue 1.

The quadratic form

$$W(2^m) = \sum_{1\le i,j\le 2^m} a'_{ij} X_i\, X_j,$$

with X_i independent normal N(0,1) distributed, satisfies the conditions of Th. 2.1.1, as is shown in the introduction. (This can also be seen directly from the theorems 3.2.5 and 3.2.6.) However, since the matrix $(a''_{ij}) = \max(0,a'_{ij})$ has a non-vanishing maximal eigenvalue $\ge 2^{m-1} - 1$ (use vector with all entries equal 1), the partial sum

$$W'(2^m) = \sum_{1\le i,j\le 2^m} a''_{ij} X_i X_j$$

does not inherit condition c) of Th. 2.1.1. This ends the example.

4.1. Simple operations on W(n)

In this section we give central limit theorems for transforms obtained from homogeneous sums W(n) (several for the same n) by the two operations addition and multiplication. We start with addition.

Theorem 4.1.1. Let $W(n) = W^{(1)}(n) +...+ W^{(q)}(n)$ be a sum of q (q not depending on n) d-homogeneous sums in the Hoeffding decomposition. Suppose that $W^{(g)}(n)$ satisfies the conditions a), b) and c) of Th. 2.1.1, with c) in the alternative formulation:

$$E\, W^{(g)}(n)^4 - 3\, \mathrm{var}^2 W^{(g)}(n) \to 0, \text{ for } g = 1,..., q, \text{ and}$$

$$\lim_{n\to\infty} \mathrm{var}\, W(n) = \sigma^2, \quad (0 < \sigma < \infty).$$

Then

$$W(n) \overset{d}{\to} N(0,\sigma^2), \quad n \to \infty.$$

Proof. Without loss of generality we assume $\sigma^2 = 1$. We shall check the conditions of Prop. 2.3.4. For $W_I = W_I^{(1)} +...+ W_I^{(q)}$ we have by the Cauchy-Schwarz inequality

$$\sigma_I^2 \le q(\sigma_I^{(1)2} +...+ \sigma_I^{(q)2})$$

and hence

$$\max_i \sum_{I\ni i} \sigma_I^2 \le q\,(\max_i \sum_{I\ni i} \sigma_I^{(1)2} +...+ \max_i \sum_{I\ni i} \sigma_I^{(q)2}),$$

where the right-hand side vanishes. Thus we have by Prop. 2.4.3 b') that τ vanishes. Notice that condition b) of Th. 2.1.1 is not closed under finite addition. However, by

$$|\, E\, W_I W_J W_K W_L \,| \le \sum_{1\le i,j,k,l\le q} |\, E\, W_I^{(i)} W_J^{(j)} W_K^{(k)} W_L^{(l)} \,|$$

we can use condition b) for $W^{(g)}(n)$, $1 \le g \le q$. Thus

$$\tau^* = \sum_{(I,J,K,L) \in \mathcal{T}} | E\, W_I W_J W_K W_L |$$

$$\le D \sum_{1 \le i,j,k,l \le q} \sum_{(I,J,K,L) \in \mathcal{T}} \sigma_I^{(i)} \sigma_J^{(j)} \sigma_K^{(k)} \sigma_L^{(l)}$$

$$= D \sum_{(I,J,K,L) \in \mathcal{T}} \left((\sigma_I^{(1)} + ... + \sigma_I^{(q)}) \; ... \; (\sigma_L^{(1)} + ... + \sigma_L^{(q)}) \right)$$

and since

$$\max_i \sum_{I \ni i} (\sigma_I^{(1)} + ... + \sigma_I^{(q)})^2 \le q \,(\max_i \sum_{I \ni i} (\sigma_I^{(1)})^2 + ... + \max_i \sum_{I \ni i} (\sigma_I^{(q)})^2),$$

which vanishes, we have by Prop. 2.4.3 b') that τ^* vanishes.

The random variables W_I are components in the Hoeffding decomposition, hence $\gamma = 0$. We shall now check condition II of Prop. 2.3.4. By Corollary 3.1.4 it is sufficient to show that $S_0 = S(1,0) + ... + S(d-1,0)$ vanishes. As in the proof of Prop. 2.3.6 we rewrite $S(e,0)$ as a sum of squares plus a remainder term (see (2.3.5)). The remainder terms R_i $(i = 1,...,5)$ below are sums over subsets of $\mathcal{T}$, so we have $| R_i | \le \tau^*$.

(4.1.1)
$$S(e,0) = \sum_{A,A' \subset \{1,...,n\},\, |A| = |A'| = d-e,\, A \cap A' = \emptyset} E\Big(\sum_{|I \cap J| = e,\, I \setminus J = A,\, J \setminus I = A'} W_I W_J \Big)^2 + R_1,$$

with $W_I W_J = \sum_{1 \le g,h \le q} W_I^{(g)} W_J^{(h)}$. By the Cauchy-Schwarz inequality we obtain

$$| S(e,0) | \le q^2 \sum_{1 \le g,h \le q} \sum_{A,A' \subset \{1,...,n\},\, |A| = |A'| = d-e,\, A \cap A' = \emptyset} E\Big(\sum_{|I \cap J| = e,\, I \setminus J = A,\, J \setminus I = A'} W_I^{(g)} W_J^{(h)} \Big)^2 + | R_1 |.$$

If $g = h$, the sum of squares vanishes by the assumptions on $W^{(g)}(n)$. It remains to show, with $W'_I = W_I^{(g)}$, $W''_I = W_I^{(h)}$, $g \ne h$, that

$$\sum_{A,A' \subset \{1,...,n\},\, |A| = |A'| = d-e,\, A \cap A' = \emptyset} E\Big(\sum_{|I \cap J| = e,\, I \setminus J = A,\, J \setminus I = A'} W'_I W''_J \Big)^2$$

$$= \sum_{B(e,0)} E\, W'_I W''_J W''_K W'_L + R_2$$

vanishes. The right-hand side equals up to a vanishing remainder term the following sum of covariances

$$\sum_{B(e,0)} E\, W'_I W''_J W''_K W'_L + R_3$$

$$= \sum_{A,A' \subset \{1,\dots,n\},\, |A| = |A'| = e,\, A \cap A' = \emptyset}$$

$$E\Big(\sum_{|I \cap L| = d-e,\, I \setminus L = A,\, L \setminus I = A'} W'_I W'_L \Big)\Big(\sum_{|J \cap K| = d-e,\, J \setminus K = A,\, K \setminus J = A'} W''_J W''_K \Big)$$

$$\le \Big(\sum_{B(e,0)} E\, W'_I W'_J W'_K W'_L + R_4 \Big)^{1/2} \Big(\sum_{B(e,0)} E\, W''_I W''_J W''_K W''_L + R_5 \Big)^{1/2},$$

where the last inequality follows by Cauchy-Schwarz and (2.3.1). This proves the theorem, since the right-hand side vanishes under the assumptions on $W^{(g)}(n)$ and $W^{(h)}(n)$, as is shown in Sect. 2.3.

Next we shall look at multiplication. Consider the homogeneous sum W(n) and its transform

$$W'(n) = \sum_{|I| = d} a_I W_I, \text{ with } |a_I| \le 1.$$

Put $W'_I = a_I W_I$ and $\sigma'^2_I = E\, W'^2_I$. Suppose that W(n) satisfies the conditions a), b) and c) of Th. 2.1.1. Clearly W'(n) inherits the properties a) and b) since

$$\max_i \sum_{I \ni i} \sigma'^2_I \le \max_i \sum_{I \ni i} \sigma^2_I$$

and

$$EW'^4_I / \sigma'^4_I = EW^4_I / \sigma^4_I \text{ for } a_I \ne 0.$$

In general, condition c) is not inherited: Consider $W(n) = \sum_{1 \le i < j \le n} a_{ij} X_i X_j$, with (a_{ij}) and X_i as in the example of the previous section. Then

$$W'(n) = \sum_{1 \le i < j \le n} a^2_{ij} X_i X_j = 1/2 \Big(\big(\sum_{1 \le i \le n} X_i \big)^2 - \sum_{1 \le i \le n} X^2_i \Big)$$

has a non-normal limit distribution.

We introduce a special family (a_I) with a very simple structure: The family $(a_I)_{|I| = d}$ is *of rank 1* if $a_I = \prod_{i \in I} a_i$ for all $|I| = d$, with a fixed sequence $(a_i)_{i = 1,\dots,n}$.

Proposition 4.1.2. Let the homogeneous sums W(n) satisfy the conditions of Th. 2.1.1 and let $(a_I)_{|I| = d}$ be of rank 1 with $|a_i| \le 1$. Then the transform

$$W'(n) = \sum_{|I| = d} a_I W_I$$

inherits the conditions a), b) and c) of Th. 2.1.1 (with the latter condition reformulated as $E\,W'(n)^4 - 3\,\mathrm{var}^2 W'(n) \to 0$ for $n \to \infty$).

Proof. Only condition c) needs a proof. Since $|\,E\,W'(n)^4 - 3\,\mathrm{var}^2\,W'(n)\,| \le |\,3\,S'_0 + S'\,| + 3\,\tau' + \tau'^*$ (see proof of Prop. 2.3.5), and since by the conditions a) and b) and Prop. 2.4.3 both τ' and τ'^* vanish, it suffices to show that S' and S'_0 vanish. By Corollary 3.1.4 it is sufficient to show $S'_0 = S'(1,0) + ... + S'(d-1,0) \to 0$. As in (4.1.1), S'(e,0) is written as a sum of squares plus a remainder term. These remainder terms R_i are sums over subsets of $\mathcal{T}$; hence $|\,R_i\,| \le \tau^*$. Using the fact that a_I is a product we obtain

$$(4.1.2)\quad S'(e,0) = \sum_{A,A' \subset \{1,...,n\},\ |A| = |A'| = d-e,\ A\cap A' = \emptyset} \Big(\prod_{i \in A \cup A'} a_i^2 \Big)$$

$$E\Big(\sum_{|I\cap J| = e,\ I \setminus J = A,\ J \setminus I = A'} W_I W_J \prod_{i \in I \cap J} a_i^2 \Big)^2 - R_1.$$

Since $\sum_i b_i c_i$, $(0 \le b_i, c_i \le 1)$ vanishes, if $\sum_i b_i \to 0$, it is sufficient to show

$$S''(e,0) = \sum_{\mathcal{B}(e,0)} E\ W_I W_J W_K W_L \prod_{i \in (I\cap J) \cup (k\cap L)} a_i^2 \to 0.$$

S''(e,0) is obtained from (4.1.2) by omitting the coefficients $\Big(\prod_{i \in A \cup A'} a_i^2 \Big)$, calculating the squares and summing over the subsets A, A' and finally neglecting the contribution of non-bifolds. As in (4.1.2) we obtain

$$S''(e,0) = \sum_{A,A' \subset \{1,...,n\},\ |A| = |A'| = e,\ A\cap A' = \emptyset} \Big(\prod_{i \in A \cup A'} a_i^2 \Big)$$

$$E\Big(\sum_{|I\cap J| = d-e,\ I \setminus L = A,\ L \setminus I = A'} W_I W_L \Big)^2 - R_2.$$

The right-hand side with the coefficients $\Big(\prod_{i \in A \cup A'} a_i^2 \Big)$ (which are ≤ 1) replaced by 1 equals, up to a vanishing remainder term, S(e,0) (see (4.1.1)). And S(e,0) vanishes by the assumptions on W(n) as is shown in Sect. 2.3. This completes the proof of Prop. 4.1.2.

With the help of Th. 4.1.1 we can extend the class of coefficients in Prop. 4.1.2. A family $(a_I)_{|I| = d}$ is said to be *of finite rank* if $a_I = a_I^{(1)} + ... + a_I^{(q)}$ for all I, with q fixed and $a_I^{(g)}$ of rank 1 for $1 \le g \le q$.

Theorem 4.1.3. Let the family $(a_I)_{|I|=d}$ be of finite rank. Suppose W(n) is a homogeneous sum in the Hoeffding decomposition satisfying the conditions of Th. 2.1.1. Put $W'(n) = \sum_{|I|=d} a_I W_I$ and suppose var $W'(n) \to \sigma^2$ for $n \to \infty$, with $0 < \sigma < \infty$. Then

$$W'(n) \overset{d}{\to} N(0,\sigma^2) \text{ for } n \to \infty.$$

Proof. Obvious from Th. 4.1.1.

In the remainder of this section we shall prove a result mentioned in the introduction. For each n, let $A_1, \ldots, A_q$ be a partition of the integers 1,...,n (q not depending on n). This partition induces a partition $\mathcal{A}_1, \ldots, \mathcal{A}_l$ of the indices $\{I \subset \{1,\ldots,n\}: |I| = d\}$. The elements $\mathcal{A}_h$ are defined in the following way. Consider the *symmetrized product sets* $\mathcal{A}_h^* = \bigcup_\sigma A_{g_{\sigma(1)}} \times \ldots \times A_{g_{\sigma(d)}}$, where $\sigma = (\sigma(1),\ldots,\sigma(d))$ passes through all permutations of 1,...,d and where $h = 1,\ldots,\binom{q+d-1}{d}$ is some enumeration of the d-tuples $\{(g_1,\ldots,g_d) : 1 \le g_1 \le \ldots \le g_d \le q\}$. Then $I = \{i_1,\ldots,i_d\} \in \mathcal{A}_h$ if $(i_1,\ldots,i_d) \in \mathcal{A}_h^*$.

We shall show that for fixed h the family $(1_{\mathcal{A}_h}(I))_{|I|=d}$ is of finite rank. Hence Th. 4.1.3 can be used to show that the partial sum

$$W^{(h)}(n) = \sum_{|I|=d} 1_{\mathcal{A}_h}(I) W_I = \sum_{I \in \mathcal{A}_h} W_I$$

has a normal limit distribution, provided

$$\lim_{n \to \infty} \sum_{I \in \mathcal{A}_h} \sigma_I^2 \text{ is finite and positive.}$$

We shall use a well-known theorem in algebra that states that a d-linear symmetric function $\phi(x_1,\ldots,x_d)$ can be written as a linear combination of diagonals:

$$\phi(x_1,\ldots,x_d) = \sum_{1\le i\le r} \alpha_i \, \phi(y_i,\ldots,y_i),$$

with $y_i = \beta_{i1}x_1 + \ldots + \beta_{id}x_d$.

Let l^∞ be the Banach space of bounded real sequences

$$l^\infty = \{ (a_i)_{i=1,2,\ldots} : \sup_i |a_i| \le C \text{ for some } C < \infty\},$$

equipped with scalar multiplication $\lambda a = (\lambda a_i)_{i=1,2,\ldots}$, addition $a + b = (a_i + b_i)_{i=1,2,\ldots}$ and norm $\|a\| = \sup_i |a_i|$. Let l_d^∞ be the Banach space of bounded real sequences on $\mathbb{N}^d$

$$l_d^\infty = \{ (a_{i_1\ldots i_d})_{i_1,\ldots,i_d=1,2,\ldots} : \sup_{i_1,\ldots,i_d} |a_{i_1\ldots i_d}| \le C \text{ for some } C < \infty\},$$

with addition, scalar multiplication and norm as above. Define $\phi: (l^\infty)^d \to l_d^\infty$ by

$$\phi(a^{(1)},\ldots, a^{(d)})_{i_1\ldots i_d} = \sum_{\sigma} a^{(\sigma(1))}_{i_1}\ldots a^{(\sigma(d))}_{i_d},$$

where $\sigma = (\sigma(1),\ldots, \sigma(d))$ passes through all permutations of $1,\ldots, d$. Then ϕ is symmetric and d-linear in $a^{(1)},\ldots, a^{(d)}$.

Define $a^{(g)} \in l^{\infty}$ by $a^{(g)}_i = 1_{A_g}(i)$ $i = 1,2,\ldots$ for each element A_g in the partition $A_1,\ldots,A_q$. Then we have

$$\phi(a^{(g_1)},\ldots, a^{(g_d)}) = \gamma\, 1_{\mathcal{A}_h^*},$$

with $\mathcal{A}_h^*$ the symmetrized product set of $A_{g_1}\times\ldots\times A_{g_d}$ and γ the number of permutations that leave $A_{g_1}\times\ldots\times A_{g_d}$ invariant. Thus $1_{\mathcal{A}_h^*}$ can be written as a sum of diagonals

$$\gamma\, 1_{\mathcal{A}_h^*} = \sum_{1\le i\le r} \alpha_i\, \phi(b^{(i)},\ldots,b^{(i)}),$$

with $b^{(i)} = \beta_{i1}a^{(g_1)}+\ldots+ \beta_{id}a^{(g_d)}$. Using the definition of ϕ, we have

$$\gamma\, 1_{\mathcal{A}_h^*}(j_1,\ldots,j_d) = \sum_{1\le i\le r} \alpha_i\ d!\ b^{(i)}_{j_1}\ldots b^{(i)}_{j_d}$$

and, for $I = \{j_1,\ldots,j_d\}$, this gives

$$1_{\mathcal{A}_h}(I) = 1_{\mathcal{A}_h^*}(j_1,\ldots,j_d) = (d!/\gamma) \sum_{1\le i\le r} \alpha_i \prod_{j\in I} b^{(i)}_j.$$

Hence it is shown that $(1_{\mathcal{A}_h}(I))_{|I|=d}$ is of finite rank. Thus we have by Th. 4.1.3

$$W^{(h)}(n) \xrightarrow{d} N(0,\sigma_h^2) \text{ for } n \to \infty, \text{ with } \sigma_h^2 = \lim_{n\to\infty} \sum_{I\in\mathcal{A}_h} \sigma_I^2.$$

In fact, we have shown more: The joint limit distribution of the partial sums $W^{(h)}(n)$ is $\binom{q+d-1}{d}$-variate normal with vanishing covariances, provided the variances converge:

$$\lim_{n\to\infty} \sum_{I\in\mathcal{A}_h} \sigma_I^2 = \sigma_h^2, \quad h = 1,\ldots,\binom{q+d-1}{d}.$$

This follows directly from Th. 4.1.3 which implies that any linear combination of partial sums with coefficients $a_1,\ldots,a_l$ $(l = \binom{q+d-1}{d})$ has a normal $N(0, \sum_{1\le h\le l} a_h^2\sigma_h^2)$ limit distribution since $\sum_{1\le h\le l} a_h 1_{\mathcal{A}_h}(I)$ is of finite rank and $\mathrm{var}(\sum_{1\le h\le l} a_h W^{(h)}(n))$ converges. Summarizing we have

Corollary 4.1.4. Let W(n) be homogeneous sums in the Hoeffding decomposition satisfying the conditions of Th. 2.1.1. Let $A_1,\ldots,A_q$ be partitions of the integers $\{1,\ldots,n\}$ (q fixed) and $\mathcal{A}_1,\ldots,\mathcal{A}_l$ the corresponding partition of the indices $\{I \subset \{1,\ldots,n\}: |I| = d\}$ with $l = \binom{q+d-1}{d}$ elements as described above. Define the partial sums

$$W^{(h)}(n) = \sum_{I \in \mathcal{A}_h} W_I .$$

Suppose that for coefficients $a_1,\dots,a_l$ (not depending on n) the variance of the linear combination of partial sums converges,

$$\lim_{n \to \infty} \operatorname{var} \Big(\sum_{1 \le h \le l} a_h W^{(h)}(n) \Big) = \sigma^2,$$

then

$$\sum_{1 \le h \le l} a_h W^{(h)}(n) \xrightarrow{d} N(0,\sigma^2), \quad n \to \infty.$$

Moreover, if the variance of each partial sum converges,

$$\lim_{n \to \infty} \operatorname{var} (W^{(h)}(n)) = \sigma_h^2, \quad h = 1,\dots,\binom{q+d-1}{d},$$

then any linear combination converges in distribution:

$$\sum_{1 \le h \le l} a_h W^{(h)}(n) \xrightarrow{d} N(0, \sum_{1 \le h \le l} a_h^2 \sigma_h^2) \text{ for } n \to \infty.$$

Thus the simultaneous distribution of the l partial sums is l-variate normal with vanishing covariances.

4.2. Convergence to a Gaussian process

It is well known that on a finite measure space $(S,\mathcal{B},\mu)$ a Gaussian process ξ with independent increments can be defined in the following way. Take for ξ a process on a probability space $(\Omega,\mathcal{F},P)$, indexed by elements $B \in \mathcal{B}$, with finite dimensional distributions $\xi(B_1),\dots,\xi(B_q)$ that are q-variate normal with $\operatorname{cov}(\xi(B_g), \xi(B_h)) = \mu(B_g \cap B_h)$. The existence of such a process follows from the Kolmogorov extension theorem. (In fact, $\{\xi(B) : B \in \mathcal{B}\}$ is a collection random variables with prescribed consistent finite dimensional distributions.)

Define the stochastic integral with respect to ξ for stepfunctions $t = b_1 1_{B_1} +\dots+ b_q 1_{B_q}$, with $B_1,\dots,B_q$ measurable and disjoint, by

$$\int t \, d\xi = b_1 \xi(B_1) +\dots+ b_q \xi(B_q).$$

Then

$$\operatorname{var} \Big(\int t \, d\xi \Big) = b_1^2 \mu(B_1) +\dots+ b_q^2 \mu(B_q) = \int t^2 \, d\mu.$$

Thus ξ maps the linear set of stepfunctions isometrically into $L^2(P)$. Since the stepfunctions are dense in $L^2(\mu)$, this isometry has a unique extension to an isometry from $L^2(\mu)$ into $L^2(P)$.

The random variables W_I are embedded in S as real-valued random point masses at points s_I, $I \subset \{1,...,n\}$ $|I| = d$. Define the stochastic integral with respect to W(n) by

$$\int f\, dW(n) = \sum_{|I| = d} f(s_I) W_I .$$

Suppose W(n) satisfies the conditions of Th. 2.1.1. In order that the stochastic integrals converge in distribution at least the variances should converge: Suppose that the discrete probability measures μ_n defined by $\mu_n(\{s_I\}) = \sigma_I^2$, $I \subset \{1,...,n\}$ $|I| = d$, converge weakly to μ. Then $\int f^2\, d\mu_n \to \int f^2\, d\mu$ if f is μ-a.e. continuous (see lemma 4.2.5). However, these restrictions are not sufficient to ensure $\int f\, dW(n) \xrightarrow{d} \int f\, d\xi$, $n \to \infty$, for f μ-a.e. continuous. This is shown in the following example.

Example. Consider the matrix (a_{ij}), constructed in the example in Sect. 4.0. Notice that $n = 2^m$, $m = 1,2,...$ Put $W_{ij} = \binom{n}{2}^{-1/2} a_{ij} X_i X_j$, with X_i iid, N(0,1) for $1 \le i < j \le n$. Now embed (W_{ij}) in the point set $\{s_k = 2k / n(n-1) : k = 1,..., n(n-1) / 2\} \subset [0,1]$. Any enumeration $k = 1,..., n(n-1) / 2$ of the set of indices $\{(i,j) : 1 \le i < j \le n\}$ results in a discrete measure $\mu_n(\{s_k\}) = E\, W_{ij}^2 = \binom{n}{2}^{-1}$ and $\mu_n \to \lambda$ weakly, with λ Lebesgue measure on [0,1]. Choose for each n an enumeration, such that $\{i,j\}$ with $a_{ij} = 1$ is mapped into (1/2,1], and $\{i,j\}$ with $a_{ij} = -1$ into [0,1/2]. Then the stochastic integral

$$\int (1_{(1/2,1]} - 1_{[0,1/2]})\, dW(n) = \binom{n}{2}^{-1/2} \sum_{1 \le i < j \le n} X_i X_j$$

$$= \left(\frac{n}{2(n-1)}\right)^{1/2} \left(\left(\frac{1}{\sqrt{n}} \sum_{1 \le i \le n} X_i\right)^2 - \frac{1}{n} \sum_{1 \le i \le n} X_i^2 \right)$$

does not have a normal limit.

Recall that in the previous section we defined a transform W'(n) of W(n) of the form $W'(n) = \sum_{|I| = d} a_I W_I$. The integral $\int f dW(n) = \sum_{|I| = d} f(s_I) W_I$ is such a transform.

Given the embedding of the above example not all stepfunctions on [0,1] result in transforms W'(n) with the family (a_I) of finite rank. (The stepfunction $1_{(1/2,1]} - 1_{[0,1/2]}$ corresponds with the family $a_I = a_{ij}$.)

We are looking for embeddings such that integrals of stepfunctions result in transforms of finite rank. To achieve this goal it is sufficient to embed W_I in a d-fold product and to carry out the embedding 'coordinatewise': Let S be a d-fold product $S = M^d$, with M a separable metric space and $m_1,...,m_n$ a sequence in M. Put $s_I = (m_{i_1},...,m_{i_d})$ for $I = \{i_1,...,i_d : i_1 < ... < i_d\}$. Measurability in M or M^d is meant with respect to the Borel σ-algebra.

Remark. There is some arbitrariness in this embedding: A renumbering of the points m_i results in a different embedding, since by another enumeration point masses W_I can be moved to another one of the d! points $(m_{i_1},\dots,m_{i_d})$ with $\{i_1,\dots,i_d\} = I$. Therefore, we shall only consider integrands f on M^d that are symmetric in their arguments: $f(m_{i_1},\dots,m_{i_d}) = f(m_{i_{\sigma(1)}},\dots, m_{i_{\sigma(d)}})$ for any permutation $\sigma = (\sigma(1),\dots,\sigma(d))$ of $(1,\dots,d)$. Then it does not matter in which of the d! points $(m_{i_1},\dots,m_{i_d})$ the random mass W_I, $I = \{i_1,\dots,i_d\}$ is placed.

Proceeding from the given embedding, we identify in the two following theorems a broad class of integrands f, for which the integrals $\int f\, dW(n)$ converge in distribution to the stochastic integral $\int f\, d\xi$, with ξ the Gaussian process given above. Loosely speaking, these functions should obey three requirements: 1) f is symmetric, 2) f can be approximated in $L^2(\mu)$ by stepfunctions, 3) $\int f\, d\mu_n \to \int f\, d\mu$ if $\mu_n \to \mu$ weakly. The first requirement is obvious; the second one allows a very broad class: stepfunctions are dense in $L^2(\mu)$. The third requirement is a real restriction.

Theorem 4.2.1. Let M be a separable metric space, $S = M^d$ and μ a probability measure on the Borel σ-algebra of S. Let ξ be the Gaussian process with covariance measure μ given above. For each n, $(m_i)_{i=1,2,\dots}$ is a sequence in M and $s_I = (m_{i_1},\dots,m_{i_d}) \in S$ for $I = \{i_1,\dots,i_d: i_1 < \dots < i_d\}$. Let W(n) be as in Th. 2.1.1, with d as above. Suppose the probability measures μ_n defined by $\mu_n(s_I) = \sigma_I^2$ converge weakly to μ. Then for h, bounded and symmetric and μ-a.e. continuous, the stochastic integrals $\int h\, dW(n) = \sum_{|I|=d} h(s_I)\, W_I$ converge in distribution:

$$\int h\, dW(n) \xrightarrow{d} \int h\, d\xi \quad \text{for } n \to \infty.$$

Proof. The theorem follows immediately from Th. 4.2.2 with $h_n = h$ for all n.

Theorem 4.2.2. Let S, μ, ξ, μ_n, $\{s_I\}$ and W(n) be as in Th. 4.2.1. Suppose that h_n, h are symmetric, uniformly bounded, measurable functions and suppose that $h_n \to$ h in the following sense:

$$\mu\{s \in S : |h_n(s_n) - h(s)| > \delta \text{ for some } \delta > 0 \text{ and some sequence } s_n \to s\} = 0.$$

Then the stochastic integrals converge in distribution:

$$\int h_n\, dW(n) \xrightarrow{d} \int h\, d\xi \text{ for } n \to \infty.$$

Proof. Let T be the set of symmetric stepfunctions that are μ-a.e. continuous and that are defined on measurable rectangles, i.e. $t \in T$ can be written $t = b_1 1_{B_1} + \dots + b_q 1_{B_q}$ with $B_1,\dots,B_q$ disjoint and $B_g = B_g^{(1)} \times \dots \times B_g^{(d)}$, $B_g^{(e)} \subset M$ measurable and $\mu(\partial(B_g)) =$

0, g = 1,..., q, with $\partial(A)$ the boundary of the set A. The proof rests on three lemmas, which will be proved below.

Lemma 4.2.3. For $t \in T$ we have

$$\int t\, dW(n) \xrightarrow{d} \int t\, d\xi, \qquad n \to \infty.$$

Let $L^2_*(\mu)$ be the set of symmetric functions in $L^2(\mu)$.

Lemma 4.2.4. T is dense in $L^2_*(\mu)$.

Lemma 4.2.5. Let h_n, h be as in Th. 4.2.2 then

$$\int h_n^2\, d\mu_n \to \int h^2\, d\mu \qquad n \to \infty.$$

We apply these in the following way. By Lemma 4.2.4 we can find for given $\varepsilon > 0$ some $t \in T$ such that $\int (h - t)^2 d\mu < \varepsilon$. Since t is μ-a.e. continuous we have by Lemma 4.2.5, $\int (h_n - t)^2\, d\mu_n \to \int (h - t)^2\, d\mu$. Choose n_0 such that for $n > n_0$ we have $\int (h_n - t)^2\, d\mu_n < 2\varepsilon$. Then

$$\int h\, d\xi = \int t\, d\xi + \int (h - t)\, d\xi,$$

$$\int h_n\, dW(n) = \int t\, dW(n) + \int (h_n - t)\, dW(n).$$

Since both var $(\int (h - t)\, d\xi)$ and var $(\int (h_n - t)\, dW(n))$ are small, the theorem follows by Lemma 4.2.3.

Proof of Lemma 4.2.3. Consider $t \in T$,

$$t = b_1 1_{B_1} + ... + b_q 1_{B_q} \text{ with } B_1,...,B_q \text{ disjoint and } B_g = B_g^{(1)} \times ... \times B_g^{(d)},$$

with $B_g^{(e)}$ Borel sets in M. The sets $B_g^{(e)}$, $1 \le g \le q$, $1 \le e \le d$, induce a partition $A_1,...,A_r$ of M. Define the symmetric sets $\mathcal{A}_h = \bigcup_\sigma A_{i_{\sigma(1)}} \times ... \times A_{i_{\sigma(d)}}$ with $\sigma = (\sigma(1),...,\sigma(d))$ running through all permutations of (1,...,d), and $h = 1,...,l = \binom{r+d-1}{d}$ an enumeration of the d-tuples $\{(i_1,...,i_d) : 1 \le i_1 \le ... \le i_d \le r\}$. Since t is symmetric in its arguments, it can be rewritten:

$$t = \sum_{1 \le h \le l} a_h 1_{\mathcal{A}_h}.$$

Clearly, the partition $\mathcal{A}_1,..., \mathcal{A}_l$ induces a partition on the grid $\{m_1,..., m_n\}^d$ and thus on $\{1,..., n\}^d$ and thus on the indices $\{ I \subset \{1,..., n\} : |I| = d\}$: the partition of Corollary 4.1.4. Thus Corollary 4.1.4 (first part) can be applied to

$$\int t\, dW(n) = \sum_{1 \le h \le l} a_h W^{(h)}(n), \text{ with } W^{(h)}(n) = \sum_{I,\ s_I \in \mathcal{A}_h} W_I,$$

provided the variance of $\int t\, dW(n)$ converges. This convergence follows, since t is μ-a.e. continuous by definition. Hence, by Lemma 4.2.5, we have $\int t^2\, d\mu_n \to \int t^2\, d\mu$. This proves the lemma.

Proof of Lemma 4.2.4. Symmetric stepfunctions are dense in $L^2_*(\mu)$, since the set $\{a \le f < b\}$ is symmetric if f is symmetric; hence f can be approximated by linear combination of symmetric indicator functions. Let B be a symmetric Borel set in $S = M^d$. Since S is a metric space, μ is regular; that, is for any $\varepsilon > 0$ there is an open set O and a closed set F such that $F \subset B \subset O$ and $\mu(O \setminus F) < \varepsilon$. The topology on S is generated by open rectangles $O^{(1)} \times ... \times O^{(d)}$, $O^{(e)}$ open in M for e = 1,...,d. Cover B by open ε-rectangles $O_\varepsilon(s) = O_\varepsilon(m^{(1)}) \times .. \times O_\varepsilon(m^{(d)})$ around any $s \in B$, with $s = (m^{(1)},...,m^{(d)})$ and $O_\varepsilon(m^{(e)})$ the ε-ball around $m^{(e)}$. It is possible to choose for any given s an ε such that for any permutation σ of 1,..., d we have

a) $O_\varepsilon(m^{(\sigma(1))}) \times ... \times O_\varepsilon(m^{(\sigma(d))}) \subset O$,

b) $\mu\,(\partial(\,O_\varepsilon(m^{(\sigma(1))}) \times ... \times O_\varepsilon(m^{(\sigma(d))})\,)\,) = 0$,

where a) is obvious. For b) notice that $\partial(O_\varepsilon(s)) \cap \partial(O_{\varepsilon'}(s) = \emptyset$ if $\varepsilon \neq \varepsilon'$ and that only countably many disjoint sets can carry mass > 0. Hence it is possible to choose $\varepsilon > 0$ for given s such that a) and b) are satisfied. The product space S is separable, since M is separable. Thus the open cover $\bigcup_{s \in B} O_\varepsilon(s)$ has a countable subcover and O can be approximated by finitely many symmetrized open rectangles such that

$$\mu\,(\,O \setminus (\bigcup_{k \le k_0} O_{\varepsilon_k}(s_k))\,) = \mu\,(\,O \setminus (\bigcup_{k \le k_0} (O_{\varepsilon_k}(s_k) \cup \partial(O_{\varepsilon_k}(s_k)))\,)) < \varepsilon.$$

Hence B can be approximated by symmetrized open rectangles without mass on the boundary. This proves Lemma 4.2.4.

Lemma 4.2.5 is Th. 5.5 in Billingsley (1968).

4.3. The dimension of a multilinear form

Consider a triangular scheme of rowwise independent random variables $(Y_{in})_{i \in A}$, with A a countable set which may depend on n. For each n the random variables $(Y_{in})_{i \in A}$ are independent. We assume $E\, Y_{in} = 0$ and $\sum_{i \in A} E\, Y^2_{in} = 1$.

If $\sum_{i \in A} Y_{in} \xrightarrow{d} N(0,1)$ for $n \to \infty$, then also $\sum_{i \in A} a_{in} Y_{in} \xrightarrow{d} N(0,1)$, with $(a_{in})_{i \in A}$ a family of real coefficients with $| a_{in} | = 1$. This can be seen from the condition below, which is equivalent to asymptotic normality of the rowsums in the triangular scheme

$$\sum_{i \in A} \int_{|x| > \varepsilon} | x | \, | P\{Y_{in} \leq x \} - P\{Y'_{in} \leq x \} | \, dx \to 0, \; n \to \infty,$$

with Y'_{in} normal $N(0, E\, Y^2_{in})$ distributed (see Shiryayev (1984: 326)). The above criterion is invariant under sign changes. If the assumption of independence is dropped, then, in general, a central limit theorem is not invariant under sign changes. E.g. take as (Y_{in}) the d-homogeneous components in the Hoeffding decomposition (W_I), with their sum W(n) satisfying the conditions of Th. 2.1.1. Then the transform

$$W'(n) = \sum_{|I| = d} a_I W_I$$

in general does not have a normal limit as can be seen from the example in the introduction of this chapter. However, we can take as a starting point the property that asymptotic normality is invariant under sign changes. It will be shown that homogeneous sums in the Hoeffding decomposition satisfying this property behave in some respects as sums of independent random variables. This leads to the following definition.

Definition 4.3.1. Let $(W_I)_{|I| = d}$ be components in the Hoeffding decomposition with sum W(n), var W(n)) = 1 and $W(n) \xrightarrow{d} N(0,1)$, $n \to \infty$. The family $(W_I)_{|I| = d}$ is *pseudo independent* if

$$W'(n) = \sum_{|I| = d} a_I W_I \xrightarrow{d} N(0,1), \; n \to \infty$$

for any family of real coefficients $(a_I)_{|I| = d}$ with $| a_I | = 1$.

For pseudo-independent d-homogeneous components Th. 4.2.1 holds with S a separable metric space irrespective of the value of d. Since all arguments concerning the embedding itself are identical to those in the proof of Th. 4.2.1, we give a schematic proof of this embedding (Th. 4.3.3). We start with the analogue of Th. 4.1.3 for pseudo-independent components.

Proposition 4.3.2. Let the pseudo-independent components $(W_I)_{|I| = d}$ satisfy the conditions a) and b) of Th. 2.1.1. If $W(n) \xrightarrow{d} N(0,1)$, $n \to \infty$, then $W'(n) = \sum_{|I| = d} a_I W_I$ satisfies the conditions of Th. 2.1.1 for any family $(a_I)_{|I| = d}$ with $| a_I | = 1$.

Proof. This follows from the definition of pseudo independent and from Th. 3.2.5.

This leads to the following corollary of Th. 4.2.1.

Theorem 4.3.3. Let the pseudo independent components $(W_I)_{|I|=d}$ with their sum satisfying the conditions af Th. 2.1.1 be embedded in a separable metric space M at points m_I such that the probability measures μ_n defined by $\mu_n(\{m_I\}) = \sigma_I^2$ converge weakly to a probability measure μ on M. Let ξ be a Gaussian process with covariance measure μ. Then, for any bounded and μ-a.e. continuous function h the stochastic integrals $\int h\, dW(n) = \sum_{|I|=d} h(m_I)W_I$ converge in distribution:

$$\int h\, dW(n) \xrightarrow{d} \int h\, d\xi,\ n \to \infty.$$

Proof. The key observation is the following consequence of Prop. 4.3.2: For any subset $\mathcal{A} \subset \{ I \subset \{1,...,n\} : |I| = d\}$ we have $\sum_{I \in \mathcal{A}} W_I \xrightarrow{d} N(0,\sigma^2)$, $n \to \infty$, provided the mass converges:

$$\sum_{I \in \mathcal{A}} \sigma_I^2 \to \sigma^2, n \to \infty, 0 < \sigma^2 < \infty.$$

This follows since

$$\sum_{I \in \mathcal{A}} W_I = 1/2(\sum_{|I|=d} b_I W_I + \sum_{|I|=d} W_I),$$

with $b_I = 2\, 1_{\mathcal{A}} - 1 \in \{0,1\}$; the result follows by Th. 4.1.1.

Notice that T defined in the proof of Th. 4.2.2 (the set of symmetric stepfunctions on measurable rectangles without mass on the boundary) in the present case (without explicit product structure) is just the set of μ-a.e. continuous stepfunctions. And for any $t \in T$ we have $\int t\, dW(n) \xrightarrow{d} \int t\, d\xi$, $n \to \infty$, by the above observation and by Th. 4.1.1. This replaces Lemma 4.2.3. Now the theorem follows in exactly the same way as Th. 4.2.2. This completes the proof.

In Ch. 3 several sufficient conditions are given that ensure asymptotic normality for clean random variables. Most of these conditions are more restrictive than the conditions of Th. 2.1.1. In fact, in Prop. 4.3.4 below it is shown that some of these conditions imply pseudo independence.

Proposition 4.3.4. The components in the Hoeffding decomposition $(W_I)_{|I|=d}$ satisfying the conditions a) and b) of Th. 2.1.1 are pseudo independent if any one of the following conditions holds:

1) $\sum_{\mathcal{B}(e,0)} | E\, W_I W_J W_K W_L | \to 0$ for $n \to \infty$, $1 \le e \le d - 1$,

2) $\sum_{\mathcal{B}(e,0)} \sigma_I \sigma_J \sigma_K \sigma_L \to 0$ for $n \to \infty$, $1 \le e \le d - 1$,

3) $\mu^* \to 0$ for $n \to \infty$, with μ^* the maximal singular value of the family $(\sigma_I)_{|I| = d}$,

4) $\max_i \sum_{I \ni i} \sigma_I \to 0$ for $n \to \infty$,

5) $\max_{A \subset \{1,\dots,n\},\, 1 \le |A| < d} \sum_{I \supset A} \sigma_I \sum_{J \supset I \setminus A} \sigma_J \to 0$ for $n \to \infty$.

Proof. All conditions are invariant under sign changes of (W_I). Condition 1) implies $S_0 \to 0$ and thus $S \to 0$ (Corollary 3.1.4) and hence W(n) has a normal limit distribution by Prop. 2.3.4. Further we have

$4) \overset{(1)}{\Rightarrow} 5) \overset{(2)}{\Rightarrow} 2) \overset{(3)}{\Rightarrow} 1)$ and $3) \overset{(4)}{\Rightarrow} 2)$.

(1) follows since $\sum_{J \supset I \setminus A} \sigma_J \le \max_i \sum_{I \ni i} \sigma_I$.

(2) is shown in the proof of Th. 3.1.5.

(3) follows by Prop. 3.2.4.

(4) follows by Prop 3.2.1.

This finishes the proof.

References

Barbour, A.D. and G.K. Eagleson, 1985. Multiple comparisons and sums of dissociated random variables. Adv. Appl. Prob. 17, 147-162.

Beran, R.J., 1972. Rank spectral processes and tests for serial dependence. Ann. Math. Statist. 43, 1749-1766.

Billingsley, P., 1968. Convergence of probability measures. New York: Wiley.

Bloemena, A.R., 1964. Sampling from a graph. Mathematical Centre Tract 2, Amsterdam.

Brown, B.M. and D.G. Kildea, 1978. Reduced U -statistics and the Hodge-Lehmann estimator. Ann. Statist. 6, 828-835.

Chung, K.L., 1974. A course in probability theory. (2nd. edition.) New York: Academic Press.

De Jong, P., 1982. Zwakke convergentie voor stochastische variabelen gedefinieerd op een netwerk. Unpublished Master's thesis.

De Jong, P., 1985. A central limit theorem for generalized quadratic forms. Report 85-03, University of Amsterdam.

De Jong, P., 1987. A central limit theorem for generalized quadratic forms. Probab. Th. Rel. Fields 75, 261-277.

Feller, W., 1971. An introduction to probability theory and its applications II. New York: Wiley.

Golub, G.H. and C.F. Van Loan, 1983. Matrix computations. Oxford: North Oxford Academic.

Hall, P., 1984. Central limit theorem for integrated square error of multivariate nonparametric density estimators. J. Multivar. Anal. 14, 1-16.

Heyde, C.C. and B.M. Brown, 1970. On the departure from normality of a certain class of martingales. Ann. Math. Statist. 41, 2161-2165.

Hoeffding, W., 1948. A class of statistics with asymptotically normal distribution. Ann. Math. Statist. 19, 293-325.

Jammalamadaka, R.S. and S. Janson, 1986. Limit theorems for a triangular scheme of U-statistics with applications to interpoint distances. Ann. Probab. 14, 1347-1358.

Karlin, S. and Y. Rinott, 1982. Applications of ANOVA type decompositions of conditional variance statistics including Jackknife estimates. Ann. Statist. 110, 485-501.

Kester, A., 1975. Asymptotic normality of the number of small distances between random points in a cube. Stochastic Process. Appl. 3, 45-54.

McGinley, W.G. and R. Sibson, 1975. Dissociated random variables. Math. Proc. Cambridge Philos. Soc. 77, 185-188.

Moran, P.A.P., 1948. The interpretation of statistical maps. J. Roy. Statist. Soc. B 10, 243-251.

Noether, G.E., 1970. A central limit theorem with non-parametric applications. Ann. Math. Statist. 41, 1753-1755.

Rotar', V.I., 1973. Some limit theorems for polynomials of second degree. Theor. Probab. Appl. 18, 499-507.

Rotar', V.I., 1979. Limit theorems for polylinear forms. J. Multivar. Anal. 9, 511-530.

Sevast'yanov, B.A., 1961. A class of limit distributions for quadratic forms of normal stochastic variables. Theor. Probab. Appl. 6, 337-340.

Shapiro, C.P. and L. Hubert, 1979. Asymptotic normality of permutation statistics derived from weighted sums of bivariate functions. Ann. Statist. 1, 788-794.

Shiryayev, A.N., 1984. Probability. New York: Springer.

Stein, C., 1970. A bound for the error in the normal approximation of a sum of dependent random variables. Proc. Sixth Berkeley Symposium Math. Statist. Prob. 2, 583-602.

Van Zwet, W.R., 1984. A Berry-Esseen bound for symmetric statistics. Z. Wahrsch. verw. Gebiete 66, 425-440.

Weber, N.C., 1980. A martingale approach to central limit theorems for exchangeable random variables. J. Appl. Prob. 17, 662-673.

Weber, N.C., 1983. Central limit theorems for a class of symmetric statistics. Math. Proc. Cambridge Philos. Soc. 94, 307-313.

Whittle, P., 1964. On the convergence to normality of quadratic forms in independent variables. Theor. Probab. Appl. 9, 113-118.

List of terms

bifold 19
clean 5
clean version 41
dissociated 1
free index 15
Hoeffding decomposition 5
d-homogeneous 11
norm
  -Hilbert-Schmidt 37
  -operator 38
pseudo independent 78
rank 1 69
rank (finite) 71
shadow 26
singular value decomposition 37
symmetrized product sets 71
trace 2

List of symbols

$\mathcal{B}$ 19

$\mathcal{B}(e,f)$ 21

$\mathcal{B}_q$ 29

$\mathcal{F}$ 19

$\mathcal{F}_I$ (σ-algebra) 1

$\mathcal{F}^{(k)}$ (σ-algebra) 5

$\mathcal{F}_k$ (σ-algebra) 18

$\mathcal{T}$ 19

$\mathcal{T}_q$ 29

γ 21

τ 20

τ^* 20

$C_B(d,q)$ 29

$C_T(d,q)$ 29

S 21

S_0 21

S(e,f) 21

O 57

$\propto$ 57

MATHEMATICAL CENTRE TRACTS

1 T. van der Walt. *Fixed and almost fixed points.* 1963.

2 A.R. Bloemena. *Sampling from a graph.* 1964.

3 G. de Leve. *Generalized Markovian decision processes, part I: model and method.* 1964.

4 G. de Leve. *Generalized Markovian decision processes, part II: probabilistic background.* 1964.

5 G. de Leve, H.C. Tijms, P.J. Weeda. *Generalized Markovian decision processes, applications.* 1970.

6 M.A. Maurice. *Compact ordered spaces.* 1964.

7 W.R. van Zwet. *Convex transformations of random variables.* 1964.

8 J.A. Zonneveld. *Automatic numerical integration.* 1964.

9 P.C. Baayen. *Universal morphisms.* 1964.

10 E.M. de Jager. *Applications of distributions in mathematical physics.* 1964.

11 A.B. Paalman-de Miranda. *Topological semigroups.* 1964.

12 J.A.Th.M. van Berckel, H. Brandt Corstius, R.J. Mokken, A. van Wijngaarden. *Formal properties of newspaper Dutch.* 1965.

13 H.A. Lauwerier. *Asymptotic expansions.* 1966, out of print; replaced by MCT 54.

14 H.A. Lauwerier. *Calculus of variations in mathematical physics.* 1966.

15 R. Doornbos. *Slippage tests.* 1966.

16 J.W. de Bakker. *Formal definition of programming languages with an application to the definition of ALGOL 60.* 1967.

17 R.P. van de Riet. *Formula manipulation in ALGOL 60, part 1.* 1968.

18 R.P. van de Riet. *Formula manipulation in ALGOL 60, part 2.* 1968.

19 J. van der Slot. *Some properties related to compactness.* 1968.

20 P.J. van der Houwen. *Finite difference methods for solving partial differential equations.* 1968.

21 E. Wattel. *The compactness operator in set theory and topology.* 1968.

22 T.J. Dekker. *ALGOL 60 procedures in numerical algebra, part 1.* 1968.

23 T.J. Dekker, W. Hoffmann. *ALGOL 60 procedures in numerical algebra, part 2.* 1968.

24 J.W. de Bakker. *Recursive procedures.* 1971.

25 E.R. Paërl. *Representations of the Lorentz group and projective geometry.* 1969.

26 European Meeting 1968. *Selected statistical papers, part I.* 1968.

27 European Meeting 1968. *Selected statistical papers, part II.* 1968.

28 J. Oosterhoff. *Combination of one-sided statistical tests.* 1969.

29 J. Verhoeff. *Error detecting decimal codes.* 1969.

30 H. Brandt Corstius. *Exercises in computational linguistics.* 1970.

31 W. Molenaar. *Approximations to the Poisson, binomial and hypergeometric distribution functions.* 1970.

32 L. de Haan. *On regular variation and its application to the weak convergence of sample extremes.* 1970.

33 F.W. Steutel. *Preservation of infinite divisibility under mixing and related topics.* 1970.

34 I. Juhász, A. Verbeek, N.S. Kroonenberg. *Cardinal functions in topology.* 1971.

35 M.H. van Emden. *An analysis of complexity.* 1971.

36 J. Grasman. *On the birth of boundary layers.* 1971.

37 J.W. de Bakker, G.A. Blaauw, A.J.W. Duijvestijn, E.W. Dijkstra, P.J. van der Houwen, G.A.M. Kamsteeg-Kemper, F.E.J. Kruseman Aretz, W.L. van der Poel, J.P. Schaap-Kruseman, M.V. Wilkes, G. Zoutendijk. *MC-25 Informatica Symposium.* 1971.

38 W.A. Verloren van Themaat. *Automatic analysis of Dutch compound words.* 1972.

39 H. Bavinck. *Jacobi series and approximation.* 1972.

40 H.C. Tijms. *Analysis of (s,S) inventory models.* 1972.

41 A. Verbeek. *Superextensions of topological spaces.* 1972.

42 W. Vervaat. *Success epochs in Bernoulli trials (with applications in number theory).* 1972.

43 F.H. Ruymgaart. *Asymptotic theory of rank tests for independence.* 1973.

44 H. Bart. *Meromorphic operator valued functions.* 1973.

45 A.A. Balkema. *Monotone transformations and limit laws.* 1973.

46 R.P. van de Riet. *ABC ALGOL, a portable language for formula manipulation systems, part 1: the language.* 1973.

47 R.P. van de Riet. *ABC ALGOL, a portable language for formula manipulation systems, part 2: the compiler.* 1973.

48 F.E.J. Kruseman Aretz, P.J.W. ten Hagen, H.L. Oudshoorn. *An ALGOL 60 compiler in ALGOL 60, text of the MC-compiler for the EL-X8.* 1973.

49 H. Kok. *Connected orderable spaces.* 1974.

50 A. van Wijngaarden, B.J. Mailloux, J.E.L. Peck, C.H.A. Koster, M. Sintzoff, C.H. Lindsey, L.G.L.T. Meertens, R.G. Fisker (eds.). *Revised report on the algorithmic language ALGOL 68.* 1976.

51 A. Hordijk. *Dynamic programming and Markov potential theory.* 1974.

52 P.C. Baayen (ed.). *Topological structures.* 1974.

53 M.J. Faber. *Metrizability in generalized ordered spaces.* 1974.

54 H.A. Lauwerier. *Asymptotic analysis, part 1.* 1974.

55 M. Hall, Jr., J.H. van Lint (eds.). *Combinatorics, part 1: theory of designs, finite geometry and coding theory.* 1974.

56 M. Hall, Jr., J.H. van Lint (eds.). *Combinatorics, part 2: graph theory, foundations, partitions and combinatorial geometry.* 1974.

57 M. Hall, Jr., J.H. van Lint (eds.). *Combinatorics, part 3: combinatorial group theory.* 1974.

58 W. Albers. *Asymptotic expansions and the deficiency concept in statistics.* 1975.

59 J.L. Mijnheer. *Sample path properties of stable processes.* 1975.

60 F. Göbel. *Queueing models involving buffers.* 1975.

63 J.W. de Bakker (ed.). *Foundations of computer science.* 1975.

64 W.J. de Schipper. *Symmetric closed categories.* 1975.

65 J. de Vries. *Topological transformation groups, 1: a categorical approach.* 1975.

66 H.G.J. Pijls. *Logically convex algebras in spectral theory and eigenfunction expansions.* 1976.

68 P.P.N. de Groen. *Singularly perturbed differential operators of second order.* 1976.

69 J.K. Lenstra. *Sequencing by enumerative methods.* 1977.

70 W.P. de Roever, Jr. *Recursive program schemes: semantics and proof theory.* 1976.

71 J.A.E.E. van Nunen. *Contracting Markov decision processes.* 1976.

72 J.K.M. Jansen. *Simple periodic and non-periodic Lamé functions and their applications in the theory of conical waveguides.* 1977.

73 D.M.R. Leivant. *Absoluteness of intuitionistic logic.* 1979.

74 H.J.J. te Riele. *A theoretical and computational study of generalized aliquot sequences.* 1976.

75 A.E. Brouwer. *Treelike spaces and related connected topological spaces.* 1977.

76 M. Rem. *Associons and the closure statement.* 1976.

77 W.C.M. Kallenberg. *Asymptotic optimality of likelihood ratio tests in exponential families.* 1978.

78 E. de Jonge, A.C.M. van Rooij. *Introduction to Riesz spaces.* 1977.

79 M.C.A. van Zuijlen. *Emperical distributions and rank statistics.* 1977.

80 P.W. Hemker. *A numerical study of stiff two-point boundary problems.* 1977.

81 K.R. Apt, J.W. de Bakker (eds.). *Foundations of computer science II, part 1.* 1976.

82 K.R. Apt, J.W. de Bakker (eds.). *Foundations of computer science II, part 2.* 1976.

83 L.S. van Benthem Jutting. *Checking Landau's "Grundlagen" in the AUTOMATH system.* 1979.

84 H.L.L. Busard. *The translation of the elements of Euclid from the Arabic into Latin by Hermann of Carinthia (?), books vii-xii.* 1977.

85 J. van Mill. *Supercompactness and Wallman spaces.* 1977.

86 S.G. van der Meulen, M. Veldhorst. *Torrix I, a programming system for operations on vectors and matrices over arbitrary fields and of variable size.* 1978.

88 A. Schrijver. *Matroids and linking systems.* 1977.

89 J.W. de Roever. *Complex Fourier transformation and analytic functionals with unbounded carriers.* 1978.

90 L.P.J. Groenewegen. *Characterization of optimal strategies in dynamic games.* 1981.

91 J.M. Geysel. *Transcendence in fields of positive characteristic.* 1979.

92 P.J. Weeda. *Finite generalized Markov programming.* 1979.

93 H.C. Tijms, J. Wessels (eds.). *Markov decision theory.* 1977.

94 A. Bijlsma. *Simultaneous approximations in transcendental number theory.* 1978.

95 K.M. van Hee. *Bayesian control of Markov chains.* 1978.

96 P.M.B. Vitányi. *Lindenmayer systems: structure, languages, and growth functions.* 1980.

97 A. Federgruen. *Markovian control problems; functional equations and algorithms.* 1984.

98 R. Geel. *Singular perturbations of hyperbolic type.* 1978.

99 J.K. Lenstra, A.H.G. Rinnooy Kan, P. van Emde Boas (eds.). *Interfaces between computer science and operations research.* 1978.

100 P.C. Baayen, D. van Dulst, J. Oosterhoff (eds.). *Proceedings bicentennial congress of the Wiskundig Genootschap, part 1.* 1979.

101 P.C. Baayen, D. van Dulst, J. Oosterhoff (eds.). *Proceedings bicentennial congress of the Wiskundig Genootschap, part 2.* 1979.

102 D. van Dulst. *Reflexive and superreflexive Banach spaces.* 1978.

103 K. van Harn. *Classifying infinitely divisible distributions by functional equations.* 1978.

104 J.M. van Wouwe. *Go-spaces and generalizations of metrizability.* 1979.

105 R. Helmers. *Edgeworth expansions for linear combinations of order statistics.* 1982.

106 A. Schrijver (ed.). *Packing and covering in combinatorics.* 1979.

107 C. den Heijer. *The numerical solution of nonlinear operator equations by imbedding methods.* 1979.

108 J.W. de Bakker, J. van Leeuwen (eds.). *Foundations of computer science III, part 1.* 1979.

109 J.W. de Bakker, J. van Leeuwen (eds.). *Foundations of computer science III, part 2.* 1979.

110 J.C. van Vliet. *ALGOL 68 transput, part I: historical review and discussion of the implementation model.* 1979.

111 J.C. van Vliet. *ALGOL 68 transput, part II: an implementation model.* 1979.

112 H.C.P. Berbee. *Random walks with stationary increments and renewal theory.* 1979.

113 T.A.B. Snijders. *Asymptotic optimality theory for testing problems with restricted alternatives.* 1979.

114 A.J.E.M. Janssen. *Application of the Wigner distribution to harmonic analysis of generalized stochastic processes.* 1979.

115 P.C. Baayen, J. van Mill (eds.). *Topological structures II, part 1.* 1979.

116 P.C. Baayen, J. van Mill (eds.). *Topological structures II, part 2.* 1979.

117 P.J.M. Kallenberg. *Branching processes with continuous state space.* 1979.

118 P. Groeneboom. *Large deviations and asymptotic efficiencies.* 1980.

119 F.J. Peters. *Sparse matrices and substructures, with a novel implementation of finite element algorithms.* 1980.

120 W.P.M. de Ruyter. *On the asymptotic analysis of large-scale ocean circulation.* 1980.

121 W.H. Haemers. *Eigenvalue techniques in design and graph theory.* 1980.

122 J.C.P. Bus. *Numerical solution of systems of nonlinear equations.* 1980.

123 I. Yuhász. *Cardinal functions in topology - ten years later.* 1980.

124 R.D. Gill. *Censoring and stochastic integrals.* 1980.

125 R. Eising. *2-D systems, an algebraic approach.* 1980.

126 G. van der Hoek. *Reduction methods in nonlinear programming.* 1980.

127 J.W. Klop. *Combinatory reduction systems.* 1980.

128 A.J.J. Talman. *Variable dimension fixed point algorithms and triangulations.* 1980.

129 G. van der Laan. *Simplicial fixed point algorithms.* 1980.

130 P.J.W. ten Hagen, T. Hagen, P. Klint, H. Noot, H.J. Sint, A.H. Veen. *ILP: intermediate language for pictures.* 1980.

131 R.J.R. Back. *Correctness preserving program refinements: proof theory and applications.* 1980.

132 H.M. Mulder. *The interval function of a graph.* 1980.

133 C.A.J. Klaassen. *Statistical performance of location estimators.* 1981.

134 J.C. van Vliet, H. Wupper (eds.). *Proceedings international conference on ALGOL 68.* 1981.

135 J.A.G. Groenendijk, T.M.V. Janssen, M.J.B. Stokhof (eds.). *Formal methods in the study of language, part I.* 1981.

136 J.A.G. Groenendijk, T.M.V. Janssen, M.J.B. Stokhof (eds.). *Formal methods in the study of language, part II.* 1981.

137 J. Telgen. *Redundancy and linear programs.* 1981.

138 H.A. Lauwerier. *Mathematical models of epidemics.* 1981.

139 J. van der Wal. *Stochastic dynamic programming, successive approximations and nearly optimal strategies for Markov decision processes and Markov games.* 1981.

140 J.H. van Geldrop. *A mathematical theory of pure exchange economies without the no-critical-point hypothesis.* 1981.

141 G.E. Welters. *Abel-Jacobi isogenies for certain types of Fano threefolds.* 1981.

142 H.R. Bennett, D.J. Lutzer (eds.). *Topology and order structures, part 1.* 1981.

143 J.M. Schumacher. *Dynamic feedback in finite- and infinite-dimensional linear systems.* 1981.

144 P. Eijgenraam. *The solution of initial value problems using interval arithmetic; formulation and analysis of an algorithm.* 1981.

145 A.J. Brentjes. *Multi-dimensional continued fraction algorithms.* 1981.

146 C.V.M. van der Mee. *Semigroup and factorization methods in transport theory.* 1981.

147 H.H. Tigelaar. *Identification and informative sample size.* 1982.

148 L.C.M. Kallenberg. *Linear programming and finite Markovian control problems.* 1983.

149 C.B. Huijsmans, M.A. Kaashoek, W.A.J. Luxemburg, W.K. Vietsch (eds.). *From A to Z, proceedings of a symposium in honour of A.C. Zaanen.* 1982.

150 M. Veldhorst. *An analysis of sparse matrix storage schemes.* 1982.

151 R.J.M.M. Does. *Higher order asymptotics for simple linear rank statistics.* 1982.

152 G.F. van der Hoeven. *Projections of lawless sequences.* 1982.

153 J.P.C. Blanc. *Application of the theory of boundary value problems in the analysis of a queueing model with paired services.* 1982.

154 H.W. Lenstra, Jr., R. Tijdeman (eds.). *Computational methods in number theory, part I.* 1982.

155 H.W. Lenstra, Jr., R. Tijdeman (eds.). *Computational methods in number theory, part II.* 1982.

156 P.M.G. Apers. *Query processing and data allocation in distributed database systems.* 1983.

157 H.A.W.M. Kneppers. *The covariant classification of two-dimensional smooth commutative formal groups over an algebraically closed field of positive characteristic.* 1983.

158 J.W. de Bakker, J. van Leeuwen (eds.). *Foundations of computer science IV, distributed systems, part 1.* 1983.

159 J.W. de Bakker, J. van Leeuwen (eds.). *Foundations of computer science IV, distributed systems, part 2.* 1983.

160 A. Rezus. *Abstract AUTOMATH.* 1983.

161 G.F. Helminck. *Eisenstein series on the metaplectic group, an algebraic approach.* 1983.

162 J.J. Dik. *Tests for preference.* 1983.

163 H. Schippers. *Multiple grid methods for equations of the second kind with applications in fluid mechanics.* 1983.

164 F.A. van der Duyn Schouten. *Markov decision processes with continuous time parameter.* 1983.

165 P.C.T. van der Hoeven. *On point processes.* 1983.

166 H.B.M. Jonkers. *Abstraction, specification and implementation techniques, with an application to garbage collection.* 1983.

167 W.H.M. Zijm. *Nonnegative matrices in dynamic programming.* 1983.

168 J.H. Evertse. *Upper bounds for the numbers of solutions of diophantine equations.* 1983.

169 H.R. Bennett, D.J. Lutzer (eds.). *Topology and order structures, part 2.* 1983.

CWI TRACTS

1 D.H.J. Epema. *Surfaces with canonical hyperplane sections.* 1984.

2 J.J. Dijkstra. *Fake topological Hilbert spaces and characterizations of dimension in terms of negligibility.* 1984.

3 A.J. van der Schaft. *System theoretic descriptions of physical systems.* 1984.

4 J. Koene. *Minimal cost flow in processing networks, a primal approach.* 1984.

5 B. Hoogenboom. *Intertwining functions on compact Lie groups.* 1984.

6 A.P.W. Böhm. *Dataflow computation.* 1984.

7 A. Blokhuis. *Few-distance sets.* 1984.

8 M.H. van Hoorn. *Algorithms and approximations for queueing systems.* 1984.

9 C.P.J. Koymans. *Models of the lambda calculus.* 1984.

10 C.G. van der Laan, N.M. Temme. *Calculation of special functions: the gamma function, the exponential integrals and error-like functions.* 1984.

11 N.M. van Dijk. *Controlled Markov processes; time-discretization.* 1984.

12 W.H. Hundsdorfer. *The numerical solution of nonlinear stiff initial value problems: an analysis of one step methods.* 1985.

13 D. Grune. *On the design of ALEPH.* 1985.

14 J.G.F. Thiemann. *Analytic spaces and dynamic programming: a measure theoretic approach.* 1985.

15 F.J. van der Linden. *Euclidean rings with two infinite primes.* 1985.

16 R.J.P. Groothuizen. *Mixed elliptic-hyperbolic partial differential operators: a case-study in Fourier integral operators.* 1985.

17 H.M.M. ten Eikelder. *Symmetries for dynamical and Hamiltonian systems.* 1985.

18 A.D.M. Kester. *Some large deviation results in statistics.* 1985.

19 T.M.V. Janssen. *Foundations and applications of Montague grammar, part 1: Philosophy, framework, computer science.* 1986.

20 B.F. Schriever. *Order dependence.* 1986.

21 D.P. van der Vecht. *Inequalities for stopped Brownian motion.* 1986.

22 J.C.S.P. van der Woude. *Topological dynamix.* 1986.

23 A.F. Monna. *Methods, concepts and ideas in mathematics: aspects of an evolution.* 1986.

24 J.C.M. Baeten. *Filters and ultrafilters over definable subsets of admissible ordinals.* 1986.

25 A.W.J. Kolen. *Tree network and planar rectilinear location theory.* 1986.

26 A.H. Veen. *The misconstrued semicolon: Reconciling imperative languages and dataflow machines.* 1986.

27 A.J.M. van Engelen. *Homogeneous zero-dimensional absolute Borel sets.* 1986.

28 T.M.V. Janssen. *Foundations and applications of Montague grammar, part 2: Applications to natural language.* 1986.

29 H.L. Trentelman. *Almost invariant subspaces and high gain feedback.* 1986.

30 A.G. de Kok. *Production-inventory control models: approximations and algorithms.* 1987.

31 E.E.M. van Berkum. *Optimal paired comparison designs for factorial experiments.* 1987.

32 J.H.J. Einmahl. *Multivariate empirical processes.* 1987.

33 O.J. Vrieze. *Stochastic games with finite state and action spaces.* 1987.

34 P.H.M. Kersten. *Infinitesimal symmetries: a computational approach.* 1987.

35 M.L. Eaton. *Lectures on topics in probability inequalities.* 1987.

36 A.H.P. van der Burgh, R.M.M. Mattheij (eds.). *Proceedings of the first international conference on industrial and applied mathematics (ICIAM 87).* 1987.

37 L. Stougie. *Design and analysis of algorithms for stochastic integer programming.* 1987.

38 J.B.G. Frenk. *On Banach algebras, renewal measures and regenerative processes.* 1987.

39 H.J.M. Peters, O.J. Vrieze (eds.). *Surveys in game theory and related topics.* 1987.

40 J.L. Geluk, L. de Haan. *Regular variation, extensions and Tauberian theorems.* 1987.

41 Sape J. Mullender (ed.). *The Amoeba distributed operating system: Selected papers 1984-1987.* 1987.

42 P.R.J. Asveld, A. Nijholt (eds.). *Essays on concepts, formalisms, and tools.* 1987.

43 H.L. Bodlaender. *Distributed computing: structure and complexity.* 1987.

44 A.W. van der Vaart. *Statistical estimation in large parameter spaces.* 1988.

45 S.A. van de Geer. *Regression analysis and empirical processes.* 1988.

46 S.P. Spekreijse. *Multigrid solution of the steady Euler equations.* 1988.

47 J.B. Dijkstra. *Analysis of means in some non-standard situations.* 1988.

48 F.C. Drost. *Asymptotics for generalized chi-square goodness-of-fit tests.* 1988.

49 F.W. Wubs. *Numerical solution of the shallow-water equations.* 1988.

50 F. de Kerf. *Asymptotic analysis of a class of perturbed Korteweg-de Vries initial value problems.* 1988.

51 P.J.M. van Laarhoven. *Theoretical and computational aspects of simulated annealing.* 1988.

52 P.M. van Loon. *Continuous decoupling transformations for linear boundary value problems.* 1988.

53 K.C.P. Machielsen. *Numerical solution of optimal control problems with state constraints by sequential quadratic programming in function space.* 1988.

54 L.C.R.J. Willenborg. *Computational aspects of survey data processing.* 1988.

55 G.J. van der Steen. *A program generator for recognition, parsing and transduction with syntactic patterns.* 1988.

56 J.C. Ebergen. *Translating programs into delay-insensitive circuits.* 1989.

57 S.M. Verduyn Lunel. *Exponential type calculus for linear delay equations.* 1989.

58 M.C.M. de Gunst. *A random model for plant cell population growth.* 1989.

59 D. van Dulst. *Characterizations of Banach spaces not containing l^1.* 1989.

60 H.E. de Swart. *Vacillation and predictability properties of low-order atmospheric spectral models.* 1989.

61 P. de Jong. *Central limit theorems for generalized multilinear forms.* 1989.

62 V.J. de Jong. *A specification system for statistical software.* 1989.

63 B. Hanzon. *Identifiability, recursive identification and spaces of linear dynamical systems, part I.* 1989.

64 B. Hanzon. *Identifiability, recursive identification and spaces of linear dynamical systems, part II.* 1989.