The Annals of Probability
1994, Vol. 22, No. 2, 1044–1077

ORDER-OF-MAGNITUDE BOUNDS FOR EXPECTATIONS INVOLVING QUADRATIC FORMS

Victor H. de la Peña[1] and Michael J. Klass[2]

Columbia University and University of California, Berkeley

Let $X_1, X_2, \dots, X_n$ be independent mean-zero random variables and let $a_{ij}, 1 \le i, j \le n$, be an array of constants with $a_{ii} \equiv 0$. We present a method of obtaining the order of magnitude of $E\Phi(\Sigma_{1 \le i,j \le n} a_{ij} X_i X_j)$ for any such $\{X_i\}$ and $\{a_{ij}\}$ and any nonnegative symmetric (convex) function Φ with $\Phi(0) = 0$ such that, for some integer $k \ge 0$, $\Phi(x^{2^{-k}})$ is convex and simultaneously $\Phi(x^{2^{-k-1}})$ is concave on $[0, \infty)$. The approximation is based on decoupling inequalities valid for all such mean-zero $\{X_i\}$ and reals $\{a_{ij}\}$ and a certain further "independentization" procedure.

0. Introduction. Let X_i, $i = 1, \dots, n$, be independent mean-zero random variables and let a_{ij}, $1 \le i, j \le n$, be an arbitrary double sequence of constants with $a_{ii} \equiv 0$. Let Φ be any nonnegative symmetric function with $\Phi(0) = 0$ such that, for some integer $k \ge 0, \Phi(x^{2^{-k}})$ is convex and $\Phi(x^{2^{-k-1}})$ is concave on $[0, \infty)$. In this paper we study the quadratic form

$$S_{n,2} = \sum_{1 \le i,\, j \le n} a_{ij} X_i X_j \tag{0.1}$$

and present a method of approximating the exact order of magnitude of both

$$E\Phi\big(S_{n,2}\big) \text{ and } E\Phi\big(S^*_{n,2}\big) \tag{0.2}$$

where $S^*_{n,2} = \max_{2 \le k \le n} |S_{k,2}|$. When Φ is convex and $\Phi(\sqrt{x})$ is concave, our approximation is based on two quantities. One of them can be expressed as $T_{n,1} \equiv E\Phi(\Sigma_{j=1}^n s_j X_j)$, where $s_j = E|\,\Sigma_{k=1}^n a_{kj} X_k|$. According to Klass (1980), $T_{n,1}$ can be approximated in terms of quantities depending only on aspects of the one-dimensional X_i-distributions as they relate to Φ and the a_{ij}'s. The other quantity depends only on some aspects of the one-dimensional distributions of the "interaction terms" such as $a_{ij} X_i X_j I(|a_{ij} X_i| > 10 s_j, |a_{ij} X_j| > 10 s_i)$ and is similarly approximable. When $\Phi(x^{2^{-k}})$ is convex and $\Phi(x^{2^{-k-1}})$ is concave (for some $k \ge 1$), there are $2k + 2$ approximating quantities which must be constructed, only one of which involves interaction terms. The random variables from which these quantities arise depend on certain polynomials of the X_i's.

Received May 1990; revised December 1992.
[1]Supported by NSF Grant DMS-91-08006.
[2]Supported by NSF Grants DMS-89-06522 and DMS-90-07469.
AMS 1991 *subject classifications.* Primary 60E15, 60F25, 60G50.
Key words and phrases. Quadratic forms of random variables, decoupling, decoupling inequalities, expectations of functions, Khintchine's inequality, maximal inequalities.

The quadratic form $S_{n,2}$ is important in various areas of mathematics. The study of its properties has led to the construction of certain stochastic integrals as seen in Krakowiak and Szulga (1988) and Cambanis, Rosinski and Woyczinski (1985). It was also used by Bourgain and Tzafriri (1987) in dealing with the invertibility of large matrices.

One of the main tools used in the proofs is the so-called decoupling principle of quadratic forms as introduced by McConnell and Taqqu (1984, 1986) and later extended by Zinn (1985), de Acosta (1987), Kwapien (1987), Hitczenko (1988), Zinn (1989) and the just-mentioned works of Bourgain and Tzafriri and of Krakowiak and Szulga. The purpose of decoupling quadratic forms is to facilitate the approximation of (0.2). To decouple a quadratic form with respect to a function Φ means that, for all quadratic forms of independent mean-zero X_j's, (0.2) has the same order of magnitude as

$$E\Phi\left(\sum_{1\le i,j\le n} a_{ij}X_i\tilde{X}_j\right), \tag{0.3}$$

where $\{\tilde{X}_i\}$ is an independent copy of $\{X_i\}$. Heretofore, completely general decoupling results were known in the mean-zero case for $\Phi(x) = |x|^P$ with $P \ge 1$, and for general convex Φ in the symmetric and strictly stable cases. Developing the ideas of Kwapien, Krakowiak and Szulga, and Zinn, we obtain decoupling in the mean-zero case for all symmetric convex functions.

We are able to obtain a full-fledged approximation of $E\Phi(S_{n,2})$ via independentization. The approach used begins by converting the problem to an analogous one involving a decoupled sum. This already introduces more independence. Then we restrict to convex Φ with $\Phi(\sqrt{x})$ concave. Certain truncations are applied so as to separate out terms that can be handled by Jensen's inequality. What remains is a sum having certain enhanced independence-type properties. It is thereby amenable to techniques used to approximate expectations involving sums of independent variates. The restriction on $\Phi(\sqrt{x})$ is then relaxed. An extension to the nonzero-mean case will be presented elsewhere. This endeavor was prefaced by the work of de la Peña (1988), where an accurate approximation of $E|\sum_{1\le i\ne j\le n}X_iX_j|^P$ was obtained in the i.i.d. mean-zero case for $P > 1$.

We also specialize to the quadratic form $\sum_{1\le i\ne j\le n}X_iX_j$, where we obtain more explicit results: Let $X_1, X_2, \ldots, X_n$ be a sequence of independent mean-zero random variables and let $S_n = X_1 + X_2 + \cdots + X_n$. Fix $P \ge 1$. Let $1 \le J_n \le n$ be any index for which $E|X_{J_n}|^P = \max_{1\le i\le n} E|X_i|^P$. Then

$$(96)^{-P}E|S_n|^P E|S_n - X_{J_n}|^P \le E\left|\sum_{1\le i\ne j\le n} X_iX_j\right|^P \tag{0.4}$$

and

$$E\max_{2\le m\le n}\left|\sum_{1\le i\ne j\le m} X_iX_j\right|^P \le 25(16)^P E|S_n|^P E|S_n - X_{J_n}|^P. \tag{0.5}$$

1. Summary of results. For easy reference we now present the main results of this paper. The first result is a so-called decoupling inequality.

THEOREM 2.1. *Let $\{X_i\}$ be a sequence of independent mean-zero random variables with $\{\widetilde{X}_i\}$ an independent copy of $\{X_i\}$. Let $\{a_{ij}\}$, $1 \le i, j \le n$, be a double sequence of constants with $a_{ij} = a_{ji}$, $a_{ii} = 0$ for all i and j. Let $\Phi(\cdot)$ be any symmetric convex function. Then*

$$E\Phi\left(6^{-1}\sum_{1\le i,j\le n} a_{ij}X_i\widetilde{X}_j\right) \le E\Phi\left(\sum_{1\le i,j\le n} a_{ij}X_iX_j\right) \le E\Phi\left(4\sum_{1\le i,j\le n} a_{ij}X_i\widetilde{X}_j\right) \tag{2.2}$$

and

$$E\max_{2\le m\le n}\Phi\left(6^{-1}\sum_{1\le i,j\le m} a_{ij}X_i\widetilde{X}_j\right) \le E\max_{2\le m\le n}\Phi\left(\sum_{1\le i,j\le m} a_{ij}X_iX_j\right) \le E\max_{2\le m\le n}\Phi\left(4\sum_{1\le i,j\le m} a_{ij}X_i\widetilde{X}_j\right). \tag{2.2$'$}$$

Moreover, the upper bounds hold for all convex Φ.

As Corollary 2.1 confirms, the quantities in (2.2) and (2.2′) all have the same order of magnitude.

The next theorem represents a generalization of Khintchine's inequality combined with a decoupled version of this inequality.

Let F_β be the set of nonnegative symmetric convex functions Φ such that $\Phi(0) = 0$ and $\Phi(2x) \le 2^\beta\Phi(x)$.

THEOREM 2.3. *Fix any $1 \le \beta < \infty$. With the same notation as before but with $\Phi \in F_\beta$, we have that*

$$E\Phi\left(\sum_{1\le i,j\le n} a_{ij}X_iX_j\right) \approx_\beta E\Phi\left(\left(\sum_{1\le i,j\le n} a_{ij}^2X_i^2X_j^2\right)^{1/2}\right) \approx_\beta E\Phi\left(\left(\sum_{1\le i,j\le n} a_{ij}^2X_i^2\widetilde{X}_j^2\right)^{1/2}\right). \tag{2.9}$$

The symbol $\approx_\beta$ denotes that the ratio of the adjacent quantities is bounded away from 0 and ∞ by positive constants depending only on β.

The main result of Section 3 (Theorem 3.2) obtains the order of magnitude of $E\Phi(S_{n,2})$ for functions Φ in F_β with the added restriction that $\Phi(\sqrt{x})$ be a concave function for $x \ge 0$. We restate it here.

THEOREM 3.2. *Let $X_1, X_2, \dots, X_n$ be independent mean-zero random variables. Let $\Phi \in F_\beta$ be such that $\Phi(\sqrt{x})$ is concave on $[0, \infty)$. Then, for any reals a_{ij} such that $a_{ij} = a_{ji}$ and $a_{ii} = 0$ and for any fixed $0 < \gamma_0 < \gamma_1 < \infty$,*

$$E\Phi\left(\sum_{i \le i, j \le n} a_{ij} X_i X_j\right) \tag{3.11}$$

$$\approx_\beta \max\left\{E\Phi\left(\sum_{j=1}^{n} s_j' X_j\right), E\Phi\left(\sum_{1 \le i, j \le n} a_{ij} Z_{ij} \varepsilon_{ij}\right)\right\},$$

where

$$s_j = E\left|\sum_{i=1}^{n} a_{ij} X_i\right|, \tag{3.12}$$

$$L(Z_{ij}) = L\Big(X_i X_j I\big(|a_{ij} X_i| > \gamma' s_j', |a_{ij} X_j| > \gamma' s_i'\big)\Big), \tag{3.13}$$

$\gamma_0 s_j \le s_j' \le \gamma_1 s_j$, $\gamma' = 10\gamma_0^{-1}$, *$L(\varepsilon_{ij})$ is symmetric Bernoulli and $\{Z_{ij}, \varepsilon_{ij}: 1 \le i, j \le n\}$ is a set of n^2 mutually independent random variables.*

Since

$$\sum_{1 \le i, j \le n} a_{ij} X_i X_j = \sum_{1 \le i, j \le n} \frac{(a_{ij} + a_{ji})}{2} X_i X_j,$$

Theorem 3.2 can be adapted to arbitrary a_{ij} satisfying $a_{ii} = 0$. In Section 4 we treat the case of Φ as before when $\Phi(x^{2^{-k}})$ is convex and $\Phi(x^{2^{-k-1}})$ is concave for $x \ge 0$, obtaining the following result:

THEOREM 4.1. *Let Φ be a nonnegative, symmetric function with $\Phi(0) = 0$ such that, for some $k \ge 0$, $\Phi(x^{2^{-k}})$ is convex on $[0, \infty)$ and $\Phi(x^{2^{-k-1}})$ is concave on $[0, \infty)$. Take any $n \ge 2$ and any reals a_{ij} such that $a_{ij} = a_{ji}$ and $a_{ii} = 0$. Define $Y_{i,0} = X_i$ and, by induction,*

$$Y_{i,j} = Y_{i,j-1}^2 - EY_{i,j-1}^2 \quad \text{for } j \ge 1$$

and similarly for

$$\widetilde{Y}_{i,0}, \widetilde{Y}_{i,1}, \dots, \widetilde{Y}_{i,k} \quad \text{as derived from } \widetilde{X}_1, \dots, \widetilde{X}_n.$$

We then have $\Phi \in F_\beta$ *for* $\beta = 2^{k+1}$ *and*

$$
\begin{aligned}
&E\Phi\left(\left|\sum_{1\le i,j\le n} a_{ij}X_iX_j\right|\right)\\
&\quad\approx_\beta \max\Bigg\{\max_{1\le m\le k}\Phi\left(\left(\sum_{1\le i,j\le n}(a_{ij})^{2^m} EY_{i,m-1}^2 E\tilde{Y}_{j,m-1}^2\right)^{2^{-m}}\right),\\
&\qquad\qquad \max_{1\le m\le k} E\Phi\left(\left|\sum_{1\le i,j\le n}(a_{ij})^{2^m}(EY_{i,m-1}^2)\tilde{Y}_{j,m}^2\right|^{2^{-m}}\right),\\
&\qquad\qquad\qquad E\Phi\left(\left|\sum_{1\le i,j\le n}(a_{ij})^{2^k} Y_{i,k}\tilde{Y}_{j,k}\right|^{2^{-k}}\right)\Bigg\}.
\end{aligned}
\tag{4.4}
$$

The quantities on the right-hand side of (4.4) can be approximated by means of Klass (1980) and Theorem 3.2. In Section 5 we present a special case where a direct approximation takes a more explicit form. For a statement of the result see the last paragraph of the Introduction.

2. Decoupling inequalities. In this section we extend the usual results on decoupling of quadratic forms to include nonsymmetric random variables. We also present a variant of Khintchine's inequality in Theorem 2.3. Both results will be used repeatedly in Sections 3 and 5. In special cases extensions of decoupling have been obtained by different authors. When the a_{ij} are assumed to be in a Banach space (as a general version of our result proves), Bourgain and Tzafriri (1987) obtained an upper bound for the norm of the quadratic form for random variables assumed to have mean $\delta, 0 \le \delta < 1$. When the variables are assumed to be strictly p-stable, the result appears in Krakowiak and Szulga (1988). Kwapien (1987) dealt with the case of symmetric convex functions and symmetric random variables. Zinn (1989) obtained the result for L^p norms ($p \ge 1$) and all mean-zero variables. However, we believe that the general result for mean-zero variables was not known prior to this paper. The proof of Theorem 2.1 was obtained by combining three lemmas available in the current literature. We also greatly simplified the proof of Lemma 2.1. The first lemma is due to Kwapien (1987) and was communicated to us by Zinn and Hitczenko. The third lemma follows from a "polarization formula" of Mazur and Orlicz (1935). We are indebted to A. de Acosta who pointed out to us the paper of Krakowiak and Szulga (1988) which contains a variant of the second lemma.

We have also augmented these lemmas (and Theorem 2.1 as well) by noticing that they continue to apply to certain maxima [see (2.2′), (2.4′), (2.5′) and Corollary 2.1].

THEOREM 2.1. *Let $\{X_i\}$ be a sequence of independent mean-zero random variables with $\{\tilde{X}_i\}$ an independent copy of $\{X_i\}$. Let*

$$S_{n,2} = \sum_{1 \le i, j \le n} a_{ij} X_i X_j, \tag{2.1}$$

where $\{a_{ij}\}_{1 \le j \le n}$ is a sequence of constants with $a_{ii} = 0$, $a_{ij} = a_{ji}$. Let $\Phi(\cdot)$ be any symmetric convex function. Then

$$\begin{aligned} E\Phi\left(6^{-1} \sum_{1 \le i, j \le n} a_{ij} X_i \tilde{X}_j\right) &\le E\Phi\left(\sum_{1 \le i, j \le n} a_{ij} X_i X_j\right) \\ &\le E\Phi\left(4 \sum_{1 \le i, j \le n} a_{ij} X_i \tilde{X}_j\right) \end{aligned} \tag{2.2}$$

and

$$\begin{aligned} E \max_{2 \le m \le n} \Phi\left(6^{-1} \sum_{1 \le i, j \le m} a_{ij} X_i \tilde{X}_j\right) &\le E \max_{2 \le m \le n} \Phi\left(\sum_{1 \le i, j \le m} a_{ij} X_i X_j\right) \\ &\le E \max_{2 \le m \le n} \Phi\left(4 \sum_{1 \le i, j \le m} a_{ij} X_i \tilde{X}_j\right). \end{aligned} \tag{2.2'}$$

The upper bounds hold for all convex Φ.

PROOF. We require a succession of lemmas.

LEMMA 2.1. *Let $\{X_i\}$ and $\{\tilde{X}_i\}$ be independent copies of independent random variables. Then, for $i \neq j$,*

$$E\left(X_i X_j \mid X_k + \tilde{X}_k,\ k \ge 1\right) = \frac{(X_i + \tilde{X}_i)}{2} \frac{(X_j + \tilde{X}_j)}{2}. \tag{2.3}$$

PROOF. By symmetry and independence,

$$\begin{aligned} E\left(X_i X_j \mid X_k + \tilde{X}_k,\ k \ge 1\right) &= E\left(X_i \tilde{X}_j \mid X_k + \tilde{X}_k,\ k \ge 1\right) \\ &= E\left(\tilde{X}_i X_j \mid X_k + \tilde{X}_k,\ k \ge 1\right) \\ &= E\left(\tilde{X}_i \tilde{X}_j \mid X_k + \tilde{X}_k,\ k \ge 1\right). \end{aligned}$$

Hence, by linearity of conditional expectations,

$$
\begin{aligned}
E\big(X_iX_j \,|\, X_k + \widetilde{X}_k, k \ge 1\big) &= \frac{1}{4}\Big\{E\big(X_iX_j \mid X_k + \widetilde{X}_k,\ k \ge 1\big) \\
&\quad + E\big(X_i\widetilde{X}_j \,|\, X_k + \widetilde{X}_k, k \ge 1\big) + E\big(\widetilde{X}_iX_j \,|\, X_k + \widetilde{X}_k,\ k \ge 1\big) \\
&\quad + E\big(\widetilde{X}_i\widetilde{X}_j \mid X_k + \widetilde{X}_k,\ k \ge 1\big)\Big\} \\
&= E\left(\frac{(X_i + \widetilde{X}_j)}{2}\,\frac{(X_j + \widetilde{X}_j)}{2}\,\middle|\, X_k + \widetilde{X}_k, k \ge 1\right) \\
&\qquad\qquad\qquad\qquad\qquad\qquad \text{(by linearity)} \\
&= \frac{(X_i + \widetilde{X}_i)(X_j + \widetilde{X}_j)}{4} \quad \text{a.s.} \qquad \square
\end{aligned}
$$

LEMMA 2.2. *Let $\{X_i\}$ and $\{\widetilde{X}_i\}$ be independent copies of independent random variables and let $\{a_{ij}\}$ be constants such that $a_{ii} \equiv 0$. Then, for any convex function Φ,*

$$
E\Phi\left(\sum_{1 \le i, j \le n} a_{ij}X_iX_j\right) \ge E\Phi\left(4^{-1}\sum_{1 \le i, j \le n} a_{ij}(X_i + \widetilde{X}_i)(X_j + \widetilde{X}_j)\right) \tag{2.4}
$$

and

$$
\begin{aligned}
&E\max_{2 \le m \le n}\Phi\left(\sum_{1 \le i, j \le m} a_{ij}X_iX_j\right) \\
&\quad \ge E\max_{2 \le m \le n}\Phi\left(4^{-1}\sum_{1 \le i, j \le m} a_{ij}(X_i + \widetilde{X}_j)(X_j + \widetilde{X}_j)\right).
\end{aligned} \tag{2.4$'$}
$$

PROOF.

$$
\begin{aligned}
E\Phi\left(\sum_{1 \le i, j \le n} a_{ij}X_iX_j\right) &= E\left(E\Phi\left(\sum_{1 \le i, j \le n} a_{ij}X_iX_j\right)\,\middle|\, X_k + \widetilde{X}_k,\ k \ge 1\right) \\
&\ge E\Phi\left(\sum_{1 \le i, j \le n} a_{ij}E\left(X_iX_j \,\middle|\, X_k + \widetilde{X}_k,\ k \ge 1\right)\right) \\
&\qquad \text{by the conditional Jensen inequality since } \Phi \text{ is convex} \\
&= E\Phi\left(\sum_{1 \le i, j \le n} a_{ij}\frac{(X_i + \widetilde{X}_i)(X_j + \widetilde{X}_j)}{4}\right) \\
&\qquad \text{by Lemma 2.1.}
\end{aligned}
$$

Similarly,

$$
\begin{aligned}
E \max_{2\le m\le m} \Phi\Biggl(\sum_{1\le i,j\le n} a_{ij}X_iX_j\Biggr)
&= E\Biggl(E \max_{2\le m\le n} \Phi\Biggl(\sum_{1\le i,j\le m} a_{ij}X_iX_j\Biggr)\Biggm| X_k+\widetilde{X}_k,\ k\ge 1\Biggr)\\
&\ge E \max_{2\le m\le n}\Biggl(E\Phi\Biggl(\sum_{1\le i,j\le m} a_{ij}X_iX_j\Biggr)\Biggm| X_k+\widetilde{X}_k,\ k\ge 1\Biggr)\\
&= E \max_{2\le m\le n} \Phi\Biggl(\sum_{1\le i,j\le m} a_{ij}\frac{(X_i+\widetilde{X}_i)(X_j+\widetilde{X}_j)}{4}\Biggr). \qquad \square
\end{aligned}
$$

LEMMA 2.3. *Let $\{X_i\}$ and $\{\widetilde{X}_i\}$ be independent copies of independent mean-zero random variables and let $\{a_{ij}\}$ be constants with $a_{ij} = a_{ji}$, $a_{ii} = 0$. Then, for any symmetric convex function Φ,*

$$
(2.5)\quad E\Phi\Biggl(4^{-1}\sum_{1\le i,j\le n} a_{ij}(X_i+\widetilde{X}_i)(X_j+\widetilde{X}_j)\Biggr) \ge E\Phi\Biggl(6^{-1}\sum_{1\le i,j\le n} a_{ij}X_i\widetilde{X}_j\Biggr)
$$

and

$$
(2.5')\quad
\begin{aligned}
&E \max_{2\le m\le n}\Phi\Biggl(4^{-1}\sum_{1\le i,j\le m} a_{ij}(X_i+\widetilde{X}_i)(X_j+\widetilde{X}_j)\Biggr)\\
&\qquad\ge E \max_{2\le m\le n}\Phi\Biggl(6^{-1}\sum_{1\le i,j\le m} a_{ij}X_i\widetilde{X}_j\Biggr).
\end{aligned}
$$

PROOF. Since $\sum_{1\le i,j\le n} a_{ij}X_i\widetilde{X}_j = \sum_{1\le i,j\le n} a_{ij}\widetilde{X}_iX_j$, we have that

$$
\begin{aligned}
E\Phi\Biggl(\sum_{1\le i,j\le n} a_{ij}X_i\widetilde{X}_j\Biggr)
&= E\Phi\Biggl(\tfrac12\sum_{1\le i,j\le n} a_{ij}(X_i+\widetilde{X}_i)(X_j+\widetilde{X}_j)\\
&\qquad -\tfrac12\sum_{1\le i,j\le n} a_{ij}X_iX_j - \tfrac12\sum_{1\le i,j\le n} a_{ij}\widetilde{X}_i\widetilde{X}_j\Biggr)\\
&\le \tfrac13 E\Phi\Biggl(\tfrac32\sum_{1\le i,j\le n} a_{ij}(X_i+\widetilde{X}_i)(X_j+\widetilde{X}_j)\Biggr)\\
&\qquad + \tfrac13 E\Phi\Biggl(\tfrac32\sum_{1\le i,j\le n} a_{ij}X_iX_j\Biggr) + \tfrac13 E\Phi\Biggl(\tfrac32\sum_{1\le i,j\le n} a_{ij}\widetilde{X}_i\widetilde{X}_j\Biggr)
\end{aligned}
$$

by Jensen's inequality and the fact that Φ is convex and symmetric. Notice that

$$E\left(\sum_{1\le i,j\le n} a_{ij}(X_i+\widetilde{X}_i)(X_j+\widetilde{X}_j)\mid X_1,\dots,X_n\right)=\sum_{1\le i,j\le n} a_{ij}X_iX_j.$$

Hence

$$E\Phi\left(\sum_{1\le i,j\le n} a_{ij}(X_i+\widetilde{X}_i)(X_j+\widetilde{X}_j)\right)\ge E\Phi\left(\sum_{1\le i,j\le n} a_{ij}X_iX_j\right)$$

by the conditional Jensen inequality, and so (2.5) holds. Similarly,

$$\begin{aligned}
&E\max_{2\le m\le n}\Phi\left(\sum_{1\le i,j\le m} a_{ij}X_i\widetilde{X}_j\right)\\
&\quad\le \tfrac{1}{3}E\max_{2\le m\le n}\Phi\left(\tfrac{3}{2}\sum_{1\le i,j\le m} a_{ij}(X_i+\widetilde{X}_i)(X_j+\widetilde{X}_j)\right)\\
&\qquad+\tfrac{2}{3}E\max_{2\le m\le n}\Phi\left(\tfrac{3}{2}\sum_{1\le i,j\le m} a_{ij}X_iX_j\right).
\end{aligned}$$

Then, since

$$\begin{aligned}
&E\max_{2\le m\le n}\Phi\left(\sum_{1\le i,j\le m} a_{ij}X_iX_j\right)\\
&\quad=E\max_{2\le m\le n}\Phi\left(E\sum_{1\le i,j\le m} a_{ij}(X_i+\widetilde{X}_i)(X_j+\widetilde{X}_j)\,\Bigg|\,(X_1,\dots,X_n)\right)\\
&\quad\le EE\left(\max_{2\le m\le n}\Phi\left(\sum_{1\le i,j\le m} a_{ij}(X_i+\widetilde{X}_i)(X_j+\widetilde{X}_j)\right)\Bigg|\,(X_1,\dots,X_n)\right)\\
&\quad=E\max_{2\le m\le n}\Phi\left(\sum_{1\le i,j\le m} a_{ij}(X_i+\widetilde{X}_i)(X_j+\widetilde{X}_j)\right),
\end{aligned}$$

the lemma is complete. □

Combining Lemmas 2.2 and 2.3, one gets (for any symmetric convex Φ)

$$\begin{aligned}
E\Phi\left(\sum_{1\le i,j\le n} a_{ij}X_iX_j\right)&\ge E\Phi\left(4^{-1}\sum_{1\le i,j\le n} a_{ij}(X_i+\widetilde{X}_i)(X_j+\widetilde{X}_j)\right)\\
&\ge E\Phi\left(6^{-1}\sum_{1\le i,j\le n} a_{ij}X_i\widetilde{X}_j\right),
\end{aligned}$$

which gives the lower bound in (2.2) and similarly for the lower bound in (2.2′).

PROOF OF THE UPPER BOUND. Lemma 2.1 gives

$$E\big(X_i\tilde{X}_j \mid X_k+\tilde{X}_k,\ k\geq 1\big)=\frac{(X_i+\tilde{X}_i)}{2}\,\frac{(X_j+\tilde{X}_j)}{2}\quad \text{for } i\neq j. \tag{2.6}$$

Therefore, if $a_{ii}\equiv 0$,

$$\begin{aligned}
E\Phi\Bigg(\sum_{1\leq i,j\leq n} a_{ij}X_i\tilde{X}_j\Bigg) &= EE\left(\Phi\Bigg(\sum_{1\leq i,j\leq n} a_{ij}X_i\tilde{X}_j\Bigg)\,\Bigg|\,X_k+\tilde{X}_k,\ k\geq 1\right)\\
&\geq E\Phi\Bigg(\sum_{1\leq i,j\leq n} a_{ij}E\big(X_i\tilde{X}_j \mid X_k+\tilde{X}_k,\ k\geq 1\big)\Bigg)\\
&\qquad\text{(by the conditional Jensen inequality)}\\
&= E\Phi\Bigg(\sum_{1\leq i,j\leq n} a_{ij}\frac{(X_i+\tilde{X}_i)}{2}\,\frac{(X_j+\tilde{X}_j)}{2}\Bigg)\\
&\geq E\Phi\Bigg(4^{-1}\sum_{1\leq i,j\leq n} a_{ij}X_iX_j\Bigg)\\
&\qquad\text{(again by the conditional Jensen inequality).}
\end{aligned}$$

To verify the right-hand side of (2.2′), note that

$$\begin{aligned}
&E\max_{2\leq m\leq n}\Phi\Bigg(\sum_{1\leq i,j\leq m} a_{ij}X_i\tilde{X}_j\Bigg)\\
&\quad = EE\left(\max_{2\leq m\leq n}\Phi\Bigg(\sum_{1\leq i,j\leq m} a_{ij}X_i\tilde{X}_j\Bigg)\,\Bigg|\,X_k+\tilde{X}_k,\ k\geq 1\right)\\
&\quad \geq E\max_{2\leq m\leq n}E\left(\Phi\Bigg(\sum_{1\leq i,j\leq m} a_{ij}X_i\tilde{X}_j\Bigg)\,\Bigg|\,X_k+\tilde{X}_k,\ k\geq 1\right)\\
&\quad \geq E\max_{2\leq m\leq n}\Phi\left(E\Bigg(\sum_{1\leq i,j\leq m} a_{ij}X_i\tilde{X}_j\Bigg)\,\Bigg|\,X_k+\tilde{X}_k,\ k\geq 1\right)\\
&\quad = E\max_{2\leq m\leq n}\Phi\Bigg(\sum_{1\leq i,j\leq m} a_{ij}\frac{(X_i+\tilde{X}_i)(X_j+\tilde{X}_j)}{4}\Bigg)\\
&\quad \geq E\max_{2\leq m\leq m}\Phi\Bigg(4^{-1}\sum_{1\leq i,j\leq m} a_{ij}X_iX_j\Bigg).
\end{aligned}$$

In Klass (1993) it is shown that, for any nonnegative convex symmetric function Φ and independent mean-zero random variables $Y_1, \dots, Y_n$,

$$E \max_{1 \le m \le n} \Phi\left(\sum_{j=1}^{m} Y_j\right) \le 5E\Phi\left(\sum_{j=1}^{n} Y_j\right).$$

By conditioning $\sum_{1 \le i, j \le n} a_{ij} X_i \widetilde{X}_j$ on $\{X_i\}$ and applying the preceding fact to $Y_j = \sum_i a_{ij} X_i \widetilde{X}_j$ and then unconditioning, it follows that

$$E \max_{2 \le m \le n} \Phi\left(\sum_{1 \le i, j \le m} a_{ij} X_i \widetilde{X}_j\right) \le 5E\Phi\left(\sum_{1 \le i, j \le n} a_{ij} X_i \widetilde{X}_j\right).$$

Invoking Theorem 2.1 yields the following corollary.

COROLLARY 2.1. *If Φ is a nonnegative symmetric convex function,*

$$E \max_{2 \le m \le n} \Phi\left(\sum_{1 \le i, j \le m} a_{ij} X_i X_j\right) \le 5E\Phi\left(24 \sum_{1 \le i, j \le n} a_{ij} X_i X_j\right). \tag{2.7}$$

The following extension of Khintchine's inequality can be found in McConnell and Taqqu [(1986), Lemma 2.1].

THEOREM 2.2. *Let $\{\varepsilon_i\}$ be i.i.d. random variables satisfying $P(\varepsilon_i = 1) = P(\varepsilon_i = -1) = \frac{1}{2}$ and let $\{\widetilde{\varepsilon}_i\}$ be an independent copy of $\{\varepsilon_i\}$. Let $\{a_{ij}\}$ be constants such that $a_{ij} = a_{ji}$ and $a_{ii} = 0$. Fix $1 \le \beta < \infty$. Let Φ be in F_β. Then*

$$\Phi\left(\left(\sum_{1 \le i, j \le n} a_{ij}^2\right)^{1/2}\right) \approx_\beta E\Phi\left(\sum_{1 \le i, j \le n} a_{ij} \varepsilon_i \widetilde{\varepsilon}_j\right). \tag{2.8}$$

The next result represents a variant of Khintchine's inequality.

THEOREM 2.3. *Let $\{X_i\}$ be a sequence of independent mean-zero random variables. Let $\{\widetilde{X}_i\}$ be an independent copy of $\{X_i\}$. Fix $1 \le \beta < \infty$ and let Φ be in F_β. Let $\{a_{ij}\}$ be constants such that $a_{ij} = a_{ji}$ for all i, j and $a_{ii} = 0$ for all i. Then*

$$\begin{aligned} E\Phi\left(\sum_{1 \le i, j \le n} a_{ij} X_i X_j\right) &\approx_\beta E\Phi\left(\left(\sum_{1 \le i, j \le n} a_{ij}^2 X_i^2 X_j^2\right)^{1/2}\right) \\ &\approx_\beta E\Phi\left(\left(\sum_{1 \le i, j \le n} a_{ij}^2 X_i^2 \widetilde{X}_j^2\right)^{1/2}\right). \end{aligned} \tag{2.9}$$

PROOF. Let $\{\varepsilon_i\}$ be a sequence of independent random variables with $P(\varepsilon_i = 1) = P(\varepsilon_i = -1) = \frac{1}{2}$ and let $\{\widetilde{\varepsilon}_i\}$ be an independent copy of $\{\varepsilon_i\}$. Construct them so that both $\{\varepsilon_i\}$ and $\{\widetilde{\varepsilon}_i\}$ are independent of $\{X_i\}$ and $\{\widetilde{X}_i\}$. From Theorem 2.1,

$$E\Phi\left(\sum_{1\le i,j\le n} a_{ij}X_iX_j\right) \approx_\beta E\Phi\left(\sum_{1\le i,j\le n} a_{ij}X_i\widetilde{X}_j\right)$$

$$= E\Phi\left(\sum_{j=1}^n\left(\sum_{i=1}^n a_{ij}X_i\right)\widetilde{X}_j\right)$$

$$\approx_\beta E\Phi\left(\sum_{j=1}^n\left(\sum_{i=1}^n a_{ij}X_i\right)\widetilde{X}_j\widetilde{\varepsilon}_j\right)$$

(by conditioning on $X_1,\dots,X_n$ and then using Lemma 6.2)

$$\approx_\beta E\Phi\left(\sum_{i=1}^n\left(\sum_{j=1}^n a_{ij}\widetilde{X}_j\widetilde{\varepsilon}_j\right)X_i\varepsilon_i\right)$$

(as in the preceding line)

$$\approx_\beta E\Phi\left(\sum_{1\le i,j\le n} a_{ij}X_i\varepsilon_iX_j\varepsilon_j\right) \tag{2.10}$$

(by using Theorem 2.1)

$$\approx_\beta E\Phi\left(\sum_{1\le i,j\le n} a_{ij}X_iX_j\varepsilon_i\widetilde{\varepsilon}_j\right)$$

(by conditioning on $X_1,\dots,X_n$ and using using Theorem 2.1)

$$\approx_\beta E\Phi\left(\left(\sum_{1\le i,j\le n} a_{ij}^2X_i^2X_j^2\right)^{1/2}\right)$$

(by conditioning on $X_1,\dots,X_n$ and using Khintchine's inequality in Theorem 2.2)

$$\approx_\beta E\Phi\left(\left(\sum_{j=2}^n\left(\sum_{i=1}^{j-1} a_{ij}^2X_i^2\right)X_j^2\right)^{1/2}\right),$$

by using monotonicity for the one side and the triangle inequality and the fact that $a_{ij} = a_{ji}$, $a_{ii} = 0$ for the other,

$$\approx_\beta E\Phi\left(\left(\sum_{j=2}^n\left(\sum_{i=1}^{j-1} a_{ij}^2X_i^2\right)\widetilde{X}_j^2\right)^{1/2}\right), \tag{2.11}$$

by using Lemma 6.1 since $\{\sum_{j=2}^{n}(\sum_{i=1}^{j-1}a_{ij}^2X_i^2)\tilde{X}_j^2\}$ is tangent to $\{\sum_{j=2}^{n}(\sum_{i=1}^{j-1}a_{ij}^2 X_i^2)X_j^2\}$

$$\approx_\beta E\Phi\left(\left(\sum_{j=1}^{n}\sum_{i=1}^{n}a_{ij}^2X_i^2\tilde{X}_j^2\right)^{1/2}\right) \quad \text{(as in a line above).} \qquad \square$$

3. Uniform bounds (mean-zero random variables). Let $X_1, X_2, \ldots, X_n$ be a sequence of independent mean-zero random variables. Let $\{a_{ij}\}$ be a double sequence of constants such that $a_{ij} = a_{ji}$, $a_{ii} = 0$, for $1 \le i, j \le n$. In the first part of this section we present a method of obtaining the order of magnitude of

$$E\Phi\left(\left|\sum_{1\le i,j\le n} a_{ij}X_iX_j\right|\right)$$

for functions Φ such that $\Phi(0) = 0, \Phi(x)$ is convex and increasing on $[0,\infty)$ and $\Phi(\sqrt{x})$ is concave on $[0,\infty)$. Later these results will be extended to include more general Φ. To improve readability, we have placed three technical lemmas needed in the proofs at the end of the section.

As we have previously indicated, quantities such as

$$E\Phi\left(\sum_{k=1}^{n}a_kY_k\right) \quad \text{and} \quad E\max_{1\le m\le n}\Phi\left(\sum_{k=1}^{m}a_kY_k\right)$$

can be accurately approximated whenever the Y_k, $k = 1,\ldots,n$, are independent (see Lemmas 6.2 and 6.3). But here we no longer have such total independence. However, we can create just enough. The decoupling results of Section 2 are an important step in this direction. They permit us to consider $E\Phi(\sum_{1\le i,j\le n}a_{ij}X_i\tilde{X}_j)$ instead of the original quadratic form. Conditioning on $\{X_j\}$, the K-function (see Lemma 6.2) approximation to

$$E\left(\Phi\left(\sum_{1\le i,j\le n}a_{ij}X_i\tilde{X}_j\right)\middle|\{X_i\}\right)$$

is the same as that of

$$E\left(\Phi\left(\sum_{j=1}^{n}\left|\sum_{i=1}^{n}a_{ij}X_i\right|\tilde{X}_j\right)\middle|\{X_i\}\right).$$

Hence the two conditional expectations agree to within a proportionality factor uniformly bounded away from 0 and ∞ by constants independent of $\{X_i\}$. Hence so do the unconditional expectations. Thus

$$E\Phi\left(\sum_{1\le i,j\le n}a_{ij}X_iX_j\right)\approx_\beta E\Phi\left(\sum_{j=1}^{n}\left|\sum_{i=1}^{n}a_{ij}X_i\right|\tilde{X}_j\right).$$

The expression $T_{n,1}$ contributes a quick lower bound as follows: Let $s_j \equiv E|\sum_{i=1}^n a_{ij}X_i|$. Then, by Jensen's inequality,

$$E\Phi\left(\sum_{j=1}^n\left(\left|\sum_{i=1}^n a_{ij}X_i\right|\right)\tilde{X}_j\right) \ge E\Phi\left(\left|\sum_{j=1}^n s_j\tilde{X}_j\right|\right) \equiv T_{n,1}. \tag{3.0}$$

Moreover, for any $1 \le \gamma < \infty$ and any $s_j' \in [\gamma^{-1}s_j, \gamma s_j]$, there exist $0 < C_{\beta,\gamma,1} < C_{\beta,\gamma,2} < \infty$ (depending only on β and γ) such that

$$C_{\beta,\gamma,1}E\Phi\left(\left|\sum_{j=1}^n s_j'\tilde{X}_j\right|\right) \le T_{n,1} \le C_{\beta,\gamma,2}E\Phi\left(\left|\sum_{j=1}^n s_j'\tilde{X}_j\right|\right).$$

Hence there exists $C_{\beta,\gamma,3} > 0$ such that, for any such s_j',

$$E\Phi\left(\sum_{1\le i,j\le n} a_{ij}\,X_iX_j\right) \ge C_{\beta,\gamma,3}E\Phi\left(\sum_{j=1}^n s_j'X_j\right). \tag{3.1}$$

Notice that both s_j and $T_{n,1}$ can be approximated directly in terms of one-dimensional integrals involving each component of the random vector $(X_1,\dots,X_n)$ separately. In fact, we regard s_j' as resulting from just such an approximation.

The lower bound $T_{n,1}$ can be insufficient by itself. It is incomplete because it suffers from two defects. First of all, regardless of the growth of Φ, the contribution of $|\sum_{i=1}^n a_{ij}X_i|$ is assessed as s_j. Second, it ignores the possible presence of a significant interaction effect from the terms $a_{ij}X_i\tilde{X}_j$.

We will illustrate these two points by means of two examples.

EXAMPLE 1. The growth of Φ must be considered.

Take $P \ge 1$ and let $\{X_j\}$ be i.i.d. mean zero with $E|X|^P < \infty$. Put $S_n = \sum_{j=1}^n X_j$. Let $\Phi(x) = |x|^P$ and $a_{ij} = I(i \ne j)$. Then

$$\begin{aligned}
E\Phi\left(\sum_{1\le i,j\le n} a_{ij}X_i\tilde{X}_j\right) &= EE\left(\left|\sum_{j=1}^n\left(\sum_{\substack{i=1\\ i\ne j}}^n X_i\right)\tilde{X}_j\right|^P \,\middle|\, S_n, \tilde{X}_1,\dots,\tilde{X}_n\right)\\
&\ge E\left|\sum_{j=1}^n\left(\sum_{\substack{i=1\\ i\ne j}}^n E(X_i \mid S_n)\right)\tilde{X}_j\right|^P\\
&= E\left|\frac{n-1}{n}S_n\sum_{j=1}^n\tilde{X}_j\right|^P\\
&= \left(\frac{n-1}{n}\right)^P (E|S_n|^P)^2.
\end{aligned}$$

De la Peña (1988) shows that $(E|S_n|^P)^2$ does, in fact, provide the order of magnitude of $E|\sum_{1\le i,j\le n} X_i X_j|^P$. When $P = 1$, this is indeed also given by $T_{n,1}$. However, in general,

$$T_{n,1} = \left(E|S_{n-1}|\right)^P E|S_n|^P \approx \left(E|S_n|\right)^P E|S_n|^P,$$

since here the X_j are i.i.d.

Since $(E|S_n|)^P/E|S_n|^P$ can be arbitrarily small, $T_{n,1}$ can greatly underrepresent the actual order of magnitude of $E\Phi(\sum_{1\le i,j\le n} a_{ij} X_i X_j)$.

EXAMPLE 2. Even if $\Phi(x) = |x|$, the interaction effect must be considered.

Again assume $\{X_j\}$ are i.i.d. mean zero and put $S_k = X_1 + \cdots + X_k$. For $1 \le k \le n$ let $a_{2k-1,2k} = a_{2k,2k-1} = 1$. For all other (i,j) let $a_{ij} = 0$. With $\Phi(x) = |x|$,

$$\begin{aligned} E\Phi\left(\sum_{1\le i,j\le 2n} a_{ij} X_i \tilde{X}_j\right) &= E\left|\sum_{k=1}^{n} \left(X_{2k-1}\tilde{X}_{2k} + X_{2k}\tilde{X}_{2k-1}\right)\right| \\ &= E\left|\sum_{j=1}^{2n} X_j \tilde{X}_j\right| \\ &\approx K_{X_1X_2}(2n) \quad \text{[by Klass (1980)]}, \end{aligned}$$

where, for any random variable Z, $K_Z(y) = \sup\{K \ge 0\colon yE(Z^2 \wedge |Z|K) \ge K^2\}$.

On the other hand,

$$T_{2n,1} = E|X|E|S_{2n}| \approx E|X|K_X(2n).$$

We need to construct an X-distribution for which

$$E|X|K_X(2n) \ll K_{X_1X_2}(2n).$$

Let X be symmetric and satisfy

$$P\left(|X| > y\right) = \frac{e^2 \log y}{y^2} \quad \text{for } y \ge e \text{ (logs are taken with base } e).$$

Then $K_X(n) \sim (e\sqrt{n}\log n)/2$, so $T_{2n,1} \approx \sqrt{2n}\log 2n$. A calculation shows that

$$P(|X_1X_2| > y) \sim \frac{e^4 \log^3 y}{3y^2} \quad \text{as } y \to \infty.$$

Hence

$$K_{X_1X_2}(n) \sim \frac{e^2\sqrt{n}}{4\sqrt{6}} \log^2 n$$

so indeed

$$T_{2n,1} \ll E\Phi\left(\sum_{1\le i,j\le 2n} a_{ij}X_i\widetilde{X}_j\right).$$

Thus, even when $\Phi(y) = |y|, T_{2n,1}$ can be too small.

Clearly, a supplementary lower bound is needed. Notice that the (i, j)th term of $\sum_{1\le i,j\le n} a_{ij}X_i\widetilde{X}_j$ is independent of all but at most $2(n-1)$ other terms: those on the same "row" or "column." What is somehow needed is a means of creating more independence. How can this be accomplished? Prior to answering this question, it seems wise to narrow the search via various truncations. Thereby we partition $\sum_{1\le i,j\le n} a_{ij}X_i\widetilde{X}_j$ into those components we can handle and those which may still be elusive. Note that, from Theorem 2.3,

$$E\Phi\left(\left|\sum_{1\le i,j\le n} a_{ij}X_iX_j\right|\right) \approx_\beta E\Phi\left(\left|\sum_{1\le i,j\le n} a_{ij}X_i\widetilde{X}_j\right|\right)$$
$$\approx_\beta E\Phi\left(\sqrt{\sum_{1\le i,j\le n} a_{ij}^2X_i^2\widetilde{X}_j^2}\right)$$

Write $a_{ij}^2X_i^2\widetilde{X}_j^2$ as Z_{ij} and note that

$$\max\{Z_{ij,1}, Z_{ij,2}, Z_{ij,3}\} \le Z_{ij} \le Z_{ij,1} + Z_{ij,2} + Z_{ij,3},$$

where

$$Z_{ij,1} = a_{ij}^2X_i^2\widetilde{X}_j^2 I(|a_{ij}X_i| \le \gamma s_j),$$
$$Z_{ij,2} = a_{ij}^2X_i^2\widetilde{X}_j^2 I(|a_{ij}\widetilde{X}_j| \le \gamma s_i),$$
$$Z_{ij,3} = a_{ij}^2X_i^2\widetilde{X}_j^2 I(|a_{ij}\widetilde{X}_i| > \gamma s_j,\ |a_{ij}\widetilde{X}_j| > \gamma s_i).$$

Let $Z_n = \sum_{1\le i,j\le n} Z_{ij}$ and $Z_{n,k} = \sum_{1\le i,j\le n} Z_{ij,k}$, $k = 1,2,3$. Conditioning on $\{\widetilde{X}_j\}$ and applying Jensen's inequality followed by Lemma 3.3,

$$\begin{aligned}
&E\Phi\left(\sqrt{Z_{n,1}}\right)\\
&\quad\equiv E\Phi\left(\sqrt{\sum_{j=1}^n\left(\sum_{i=1}^n a_{ij}^2X_i^2 I(|a_{ij}X_i| \le \gamma s_j)\right)\widetilde{X}_j^2}\right)\\
&\quad\le E\Phi\left(\sqrt{\sum_{j=1}^n \widetilde{X}_j^2 E\sum_{i=1}^n a_{ij}^2X_i^2 I(|a_{ij}X_i| \le \gamma s_j)}\right) \qquad \text{(by Jensen's inequality)}\\
&\quad\le E\Phi\left(\sqrt{\sum_{j=1}^n \left(q_\gamma s_j\widetilde{X}_j\right)^2}\right) \qquad \text{(by Lemma 3.3)}\\
&\quad\approx_\beta E\Phi\left(\sum_{j=1}^n s_j\widetilde{X}_j\right)\\
&\quad= T_{n,1}.
\end{aligned}$$

Similarly,

$$E\Phi\left(\sqrt{Z_{n,2}}\right) \equiv E\Phi\left(\sqrt{\sum_{i=1}^{n}\left(\sum_{j=1}^{n} a_{ij}^2 \widetilde{X}_j^2 I\left(|a_{ij}\widetilde{X}_j| \le \gamma s_j\right)\right) X_i^2}\right)$$
$$= O(T_{n,1}).$$

Next, note that whenever

$$Z_{n,1} \vee Z_{n,2} \vee Z_{n,3} \le Z_n \le Z_{n,1} + Z_{n,2} + Z_{n,3}$$

(for $Z_{n,j} \ge 0$) we have

$$\max_{1\le j\le 3} E\Phi\left(\sqrt{Z_{n,j}}\right) \le E\Phi\left(\sqrt{Z_n}\right)$$
$$\le E\Phi\left(\sqrt{Z_{n,1}}\right) + E\Phi\left(\sqrt{Z_{n,2}}\right) + E\Phi\left(\sqrt{Z_{n,3}}\right) \quad (\text{since } \Phi(\sqrt{x}) \text{ is concave})$$
$$\le 3 \max_{1\le j\le 3} E\Phi\left(\sqrt{Z_{n,j}}\right).$$

Therefore, since we also know that $T_{n,1} = O(E\Phi(\sum_{1\le i,j\le n} a_{ij} X_i \widetilde{X}_j))$,

$$(3.2)\qquad \begin{aligned} &E\Phi\left(\sqrt{\sum_{1\le i,j\le n} a_{ij}^2 X_i^2 \widetilde{X}_j^2}\right) \\ &\approx_\beta T_{n,1} \vee E\Phi\left(\sqrt{\sum_{1\le i,j\le n} a_{ij}^2 X_i^2 \widetilde{X}_j^2 I\left(|a_{ij}X_i| > \gamma s_j, |a_{ij}\widetilde{X}_j| > \gamma s_i\right)}\right). \end{aligned}$$

Now the crucial random quantity to be analyzed is

$$\sum_{1\le i,j\le n} a_{ij}^2 X_i^2 \widetilde{X}_j^2 I\left(|a_{ij}X_i| > \gamma s_j, |a_{ij}\widetilde{X}_j| > \gamma s_i\right).$$

Its (i, j)th term is still independent of the other $n^2 - 2n + 1$ terms involving (i', j') for $i' \ne i$ and $j' \ne j$. Concerning the terms (i, j') and (i', j'), however, an important development has taken place. The terms along a row (same index i) or column (same index j) enjoy a quasi-independence property. Specifically, if $\gamma > 2$, then (see Lemma 3.2)

$$(3.3)\qquad \sum_{i=1}^{n} P\left(|a_{ij}X_i| > \gamma s_j\right) \le \log\left(\gamma/(\gamma-2)\right)$$

and

$$(3.4)\qquad \sum_{j=1}^{n} P\left(|a_{ij}\widetilde{X}_j| > \gamma s_i\right) \le \log\left(\gamma/(\gamma-2)\right),$$

so that for $\gamma \geq 2e^{1/4}/(e^{1/4}-1)$ there is typically at most one nonzero term on any row or column. This kind of independence by default will permit us to use moment inequalities to obtain a kind of reverse Chebyshev inequality (rather like Kolmogorov's second inequality). This is perhaps the key ingredient to bounding

$$E\Phi\left(\sqrt{\sum_{1\leq i,j\leq n} a_{ij}^2 X_i^2 \widetilde{X}_j^2 I\big(|a_{ij}X_i| > \gamma s_j, |a_{ij}\widetilde{X}_j| > \gamma s_i\big)}\right)$$

(for suitable γ), as we now verify.

THEOREM 3.1. *Let* $\gamma \geq 2e^{1/4}/(e^{1/4}-1)$. *Let* $\{X_i, \widetilde{X}_i, a_{ij}, s_i;\ 1 \leq i, j \leq n\}$ *be as defined previously. Put* $\Psi(x) = \Phi(\sqrt{x})$ *for* $x \geq 0$ *so that* $\Psi(\cdot)$ *is concave and nondecreasing on* $[0,\infty)$ *with* $\Psi(0) = 0$. *Set*

$$v_n \equiv \sup\left\{v \geq 0\colon \sum_{1\leq i,j\leq n} E(Y_{ij} \wedge v_n) \geq \frac{v_n}{2}\right\}, \tag{3.5}$$

where

$$Y_{ij} = a_{ij}^2 X_i^2 \widetilde{X}_j^2 I\big(|a_{ij}X_i| > \gamma s_j, |a_{ij}\widetilde{X}_j| > \gamma s_i\big),$$

and put

$$z_n = \sum_{1\leq i,j\leq n} E\Psi(Y_{ij}) I(Y_{ij} > v_n). \tag{3.6}$$

Then

$$(80)^{-1}(\Psi(v_n) + z_n) \leq E\Psi\left(\sum_{1\leq i,j\leq n} Y_{ij}\right) \leq \Psi(v_n) + z_n. \tag{3.7}$$

Moreover, if $\{Y_{ij}^*\}$ *are independent random variables with* $L(Y_{ij}^*) = L(Y_{ij})$, *then*

$$(50)^{-1}\big(\Psi(v_n) + z_n\big) \leq E\Psi\left(\sum_{1\leq i,j\leq n} Y_{ij}^*\right) \leq \Psi(v_n) + z_n. \tag{3.8}$$

PROOF. The upper bound is straightforward and will be proved first.

$$\begin{aligned} E\Psi\left(\sum_{1\leq i,j\leq n} Y_{ij}\right) &\leq E\Psi\left(\sum (Y_{ij} \wedge v_n)\right) + \sum_{1\leq i,j\leq n} E\Psi(Y_{ij}) I(Y_{ij} > v_n) \\ &\leq \Psi\left(E \sum_{1\leq i,j\leq n} (Y_{ij} \wedge v_n)\right) + z_n \\ &\leq \Psi\left(\frac{v_n}{2}\right) + z_n \\ &\leq \Psi(v_n) + z_n. \end{aligned}$$

We now want to bound $E\Psi(\Sigma_{1\le i,j\le n}Y_{ij})$ below in terms of $\Psi(v_n)$. Notice that

$$E(Y_{i_1j_1}\wedge v_n)(Y_{i_2j_2}\wedge v_n) \le \begin{cases} E(Y_{i_1j_1}\wedge v_n)E(Y_{i_2j_2}\wedge v_n), & \text{if } i_1\ne i_2 \text{ and } j_1\ne j_2,\\ E(Y_{i_1j_1}\wedge v_n)v_nP\big(|a_{i_2j_2}\widetilde{X}_{j_2}|>\gamma s_{i_2}\big), & \text{if } i_1=i_2 \text{ and } j_1\ne j_2,\\ E(Y_{i_1j_1}\wedge v_n)v_nP\big(|a_{i_2j_2}X_{i_2}|>\gamma s_{j_2}\big), & \text{if } i_1\ne i_2 \text{ and } j_1=j_2,\\ E(Y_{i_1j_1}\wedge v_n)v_n, & \text{if } i_1=i_2 \text{ and } j_1=j_2.\end{cases}$$

Hence

$$\begin{aligned}
E&\Bigg(\sum_{1\le i,j\le n}(Y_{ij}\wedge v_n)\Bigg)^2\\
&=\sum_{1\le i_1,i_2,j_1,j_2\le n}E(Y_{i_1j_1}\wedge v_n)(Y_{i_2j_2}\wedge v_n)\\
&\le\Bigg(\sum_{1\le i,j\le n}E(Y_{ij}\wedge v_n)\Bigg)^2+\sum_{i=1}^n\sum_{1\le j_1\ne j_2\le n}E(Y_{ij_1}\wedge v_n)(Y_{ij_2}\wedge v_n)\\
&\quad+\sum_{j=1}^n\sum_{1\le i_1\ne i_2\le n}E(Y_{i_1j}\wedge v_n)(Y_{i_2j}\wedge v_n)+\sum_{1\le i,j\le n}E(Y_{ij}\wedge v_n)^2\\
&\le\Big(\frac{v_n}{2}\Big)^2+\sum_{i=1}^n\sum_{j=1}^nE(Y_{ij}\wedge v_n)v_n\sum_{\substack{j_2=1\\ j_2\ne i\\ j_2\ne j}}^{n}P\big(|a_{ij_2}\widetilde{X}_{j_2}|>\gamma s_i\big)\\
&\quad+\sum_{j=1}^n\sum_{i=1}^nE(Y_{ij}\wedge v_n)v_n\sum_{\substack{i_2=1\\ i_2\ne j\\ i_2\ne i}}^{n}P\big(|a_{i_2j}X_{i_2}|>\gamma s_j\big)+\sum_{1\le i,j\le n}v_nE(Y_{ij}\wedge v_n)\\
&\le\Big(\frac{v_n}{2}\Big)^2+\frac{v_n}{4}\sum_{1\le i,j\le n}E(Y_{ij}\wedge v_n)\\
&\quad+\frac{v_n}{4}\sum_{1\le i,j\le n}E(Y_{ij}\wedge v_n)+v_n\sum_{1\le i,j\le n}E(Y_{ij}\wedge v_n)\quad(\text{by choice of }\gamma)\\
&=v_n^2\Big(\frac14+\frac18+\frac18+\frac12\Big)\\
&=v_n^2.
\end{aligned}$$

These inequalities used the fact that (see Lemma 3.2), for $\gamma\ge 2e^{1/4}/(e^{1/4}-1)$, we have

$$\sum_{j=1}^nP\big(|a_{ij}\widetilde{X}_j|>\gamma s_i\big)\le\tfrac14$$

and

$$\sum_{i=1}^{n} P\big(|a_{ij}X_i| > \gamma s_j\big) \le \tfrac{1}{4}.$$

Putting $W = \sum_{1\le i,j\le n}(Y_{ij}\wedge v_n)$, we have proved that

$$EW^2 \le 4(EW)^2.$$

Suppose $v_n > 0$. Since

$$\begin{aligned}\frac{EW}{2} &\le EWI\left(W \ge \frac{EW}{2}\right)\\ &\le \sqrt{EW^2 P\left(W \ge \frac{EW}{2}\right)} \quad \text{(by Cauchy–Schwarz)}\\ &\le 2EW\sqrt{P\left(W \ge \frac{EW}{2}\right)},\end{aligned}$$

it follows that

$$P\left(W \ge \frac{EW}{2}\right) \ge 4^{-2}.$$

Consequently,

$$\begin{aligned}E\Psi\left(\sum_{1\le i,j\le n} Y_{i,j}\right) &\ge E\Psi(W)\\ &\ge \Psi\left(\frac{EW}{2}\right)P\left(W \ge \frac{EW}{2}\right)\\ &\ge \Psi\left(\frac{v_n}{4}\right)4^{-2}\\ &\ge 4^{-3}\Psi(v_n) \quad \text{(since } \Psi \text{ is concave)},\end{aligned}$$

and this inequality holds even if $v_n = 0$.

It remains to find a lower bound for

$$E\Psi\left(\sum_{1\le i,j\le n} Y_{ij}I\big(Y_{ij} > v_n\big)\right).$$

Toward this end, let

$$\begin{aligned}A_{ij} &= \{Y_{i_1 j_1} \le v_n \text{ for all } (i_1, j_1) \text{ with } i_1 \ne i \text{ and } j_1 \ne j\},\\ B_{ij} &= \{|a_{i_1 j}X_{i_1}| \le \gamma s_j \text{ for all } i_1 \ne i \text{ or } j\},\\ \widetilde{B}_{ij} &= \{|a_{i j_1}\widetilde{X}_{j_1}| \le \gamma s_i \text{ for all } j_1 \ne i \text{ or } j\}.\end{aligned}$$

Notice that

$$P(A_{ij}^c) \le \lim_{\varepsilon \downarrow 0} \sum_{1 \le i_1, j_1 \le n} P(Y_{i_1, j_1} > v_n + \varepsilon)$$
$$\le \lim_{\varepsilon \downarrow 0} \sum_{1 \le i_1, j_1 \le n} E \frac{(Y_{i_1, j_1} \wedge v_n + \varepsilon)}{v_n + \varepsilon} \le \frac{1}{2}.$$

Also,

$$P(B_{ij}) \ge \inf \left\{ \prod_{k=1}^{n-2} (1 - x_k) \colon x_k \ge 0, \sum_{k=1}^{n-2} x_k \le 4^{-1} \right\}$$
$$= 1 - 4^{-1} = \tfrac{3}{4}.$$

Similarly,

$$P(\widetilde{B}_{ij}) \ge \tfrac{3}{4}.$$

Therefore,

$$\begin{aligned} P(A_{ij} \cap B_{ij} \cap \widetilde{B}_{ij}) &\ge P(B_{ij} \cap \widetilde{B}_{ij}) - P(A_{ij}^c) \\ &= P(B_{ij}) P(\widetilde{B}_{ij}) - P(A_{ij}^c) \\ &\ge \tfrac{3}{4}^2 - \tfrac{1}{2} \\ &= 16^{-1} > 0. \end{aligned}$$

By construction, if $\{Y_{ij} > v_n\} \cap A_{ij} \cap B_{ij} \cap \widetilde{B}_{i,j}$ occurs, then

$$\{Y_{i_1 j_1} > v_n\} \cap A_{i_1 j_1} \cap B_{i_1 j_1} \cap \widetilde{B}_{i_1, j_1}$$

cannot occur for any other pair (i_1, j_1). Hence

$$\begin{aligned} E\Psi\left(\sum_{1 \le i, j \le n} Y_{ij} \right) &\ge E\Psi\left(\sum_{1 \le i, j \le n} Y_{ij} I\left(\{Y_{ij} > v_n\}, A_{ij}, B_{ij}, \widetilde{B}_{ij} \right) \right) \\ &= \sum_{1 \le i, j \le n} E\Psi(Y_{ij}) I(Y_{ij} > v_n) I(A_{ij}, B_{ij}, \widetilde{B}_{ij}) \\ &= \sum_{1 \le i, j \le n} E\Psi(Y_{ij}) I(Y_{ij} > v_n) P(A_{ij} \cap B_{ij} \cap \widetilde{B}_{ij}) \\ &\ge (16)^{-1} z_n. \end{aligned}$$

Hence

$$\begin{aligned} E\Psi\left(\sum_{1 \le i, j \le n} Y_{ij} \right) &\ge \left(4^{-3} \Psi(v_n) \right) \vee 4^{-2} z_n \\ &\ge (80)^{-1} \left(\Psi(v_n) + z_n \right) \end{aligned}$$

(by a simple optimization argument).

By the same arguments as for $\sum Y_{ij}, E\Psi(\sum_{1\le i,j\le n} Y^*_{ij}) \le \Psi(v_n) + z_n$. To find a lower bound, consider

$$\begin{aligned} E\Bigg(\sum_{1\le i,j\le n} (Y^*_{ij}\wedge v_n)\Bigg)^2 &= \sum_{1\le i_1,i_2,j_1,j_2\le n} E(Y^*_{i_1 j_1}\wedge v_n)(Y^*_{i_2 j_2}\wedge v_n) \\ &\le \Bigg(\sum_{1\le i,j\le n} E(Y^*_{ij}\wedge v_n)\Bigg)^2 + \sum_{1\le i,j\le n} E(Y^{2}_{ij}\wedge v_n^2) \\ &\le \left(\frac{v_n}{2}\right)^2 + \sum_{1\le i,j\le n} v_n E(Y_{ij}\wedge v_n) \\ &= \frac{3v_n^2}{4}. \end{aligned}$$

Putting $W^* = \sum_{1\le i,j\le n}(Y^*_{ij} \wedge v_n)$, we have proved that $E(W^*)^2 \le 3(EW^*)^2$. Hence

$$P\left(W^* \ge \frac{EW^*}{2}\right) \ge \frac{(EW^*)^2}{4E(W^*)^2} \ge \frac{1}{12}$$

and so

$$\begin{aligned} E\Psi\Bigg(\sum_{1\le i,j\le n} Y^*_{ij}\Bigg) &\ge E\Psi(W^*) \\ &\ge \Psi\left(\frac{EW^*}{2}\right) P\left(W^* \ge \frac{EW^*}{2}\right) \\ &\ge (12)^{-1}\Psi\left(\frac{v_n}{4}\right) \\ &\ge (48)^{-1}\Psi(v_n). \end{aligned}$$

Note that (3.5) entails $\sum_{1\le i,j\le n} P(Y^*_{ij} > v_n) \le \frac{1}{2}$. Therefore,

$$\begin{aligned} &E\Psi\Bigg(\sum_{1\le i,j\le n} Y^*_{ij} I(Y^*_{ij} > v_n)\Bigg) \\ &\ge \sum_{1\le i,j\le n} E\Psi(Y^*_{ij}) I(Y^*_{ij} > v_n) I\Bigg(\bigcap_{\{(i',j')\ne(i,j)\}} \{Y^*_{i'j'} \le v_n\}\Bigg) \\ &\ge \sum_{1\le i,j\le n} E\Psi(Y^*_{ij}) I(Y^*_{ij} > v_n) P\Bigg(\bigcap_{i',j'} \{Y^*_{i',j'} \le v_n\}\Bigg) \\ &\ge (z_n)\Bigg(1 - \sum_{1\le i,j\le n} P(Y^*_{ij} > v_n)\Bigg) \ge \frac{z_n}{2}. \end{aligned}$$

Hence

$$E\Psi\left(\sum_{1\le i,\,j\le n} Y_{ij}^*\right) \ge \left((48)^{-1}\Psi(v_n)\vee 2^{-1}z_n\right)$$
$$\ge (50)^{-1}(\Psi(v_n)+z_n)$$
(by a simple optimization argument). □

REMARK 3.1. As a corollary,

$$E\Psi\left(\sum_{1\le i,\,j\le n} Y_{ij}\right) \approx E\Psi\left(\sum_{1\le i,\,j\le n} Y_{ij}^*\right). \tag{3.9}$$

Moreover, even if each s_j were replaced by s_j' for s_j' satisfying $\gamma_0 s_j \le s_j' \le \gamma_1 s_j$ for some fixed $0<\gamma_0<\gamma_1<\infty$, the same essential result would be obtained.

REMARK 3.2. Theorem 2.1 of de la Peña (1988) shows that for Φ convex and non-decreasing on $[0,\infty)$ with $\Phi(0)=0$ and $\Psi(x)=\Phi(\sqrt{x})$ concave, $E\Phi(|\sum_{j=1}^n d_j|) \le$ const $E\Phi(|\sum_{j=1}^n d_j^*|)$ where $\{d_j\}$ are martingale increments and $\{d_j^*\}$ are independent random variables with $\mathcal{L}(d_j^*)=\mathcal{L}(d_j)$ for $j=1,2,\dots,n$. Motivated by this result, both B. Davis and Johnson and Schechtman (1989) independently showed that for arbitrary non-negative random variables $Y_1,\dots,Y_n$,

$$E\Psi(Y_1+\cdots+Y_n) \le \text{ const } E\Psi(Y_1^*+\cdots+Y_n^*) \tag{3.10}$$

where $\{Y_j^*\}$ are independent and $\mathcal{L}(Y_j^*)=\mathcal{L}(Y_j)$. This inequality is recorded as Proposition A.1 in de la Peña (1988). As indicated by (3.9), Theorem 3.1 represents an instance in which this inequality can be reversed.

Combining Theorem 2.3, (3.2) and Theorem 3.1, we obtain the following fundamental approximation theorem. Whereas Theorems 2.1 and 2.3 provide an initial or one-step decoupling of $\sum a_{ij}X_iX_j$, Theorem 3.2 identifies the full extent to which these variates may be decoupled (or independentized). It thereby constitutes an *ultimate decoupling*. Interestingly, the process of successive independentization has produced magnitude-equivalent quantities which are approximable.

THEOREM 3.2. *Let $X_1,X_2,\dots,X_n$ be independent mean-zero random variables. Let $\Phi\in F_\beta$ be such that $\Phi(\sqrt{x})$ is concave on $[0,\infty)$. Then, for any reals a_{ij} such that $a_{ij}=a_{ji}$ and $a_{ii}=0$ and for any fixed $0<\gamma_0<\gamma_1<\infty$,*

$$\begin{aligned} &E\Phi\left(\sum_{1\le i,\,j\le n} a_{ij}X_iX_j\right)\\ &\quad\approx_\beta \max\left\{E\Phi\left(\sum_{j=1}^n s_j'X_j\right), E\Phi\left(\sum_{i\le i,\,j\le n} a_{ij}Z_{ij}\varepsilon_{ij}\right)\right\}, \end{aligned} \tag{3.11}$$

where

$$s_j = E\left|\sum_{i=1}^{n} a_{ij} X_i\right|, \tag{3.12}$$

$$L(Z_{ij}) = L\Big(X_i X_j I\big(|a_{ij} X_i| > \gamma' s_j', |a_{ij} X_j| > \gamma' s_i'\big)\Big), \tag{3.13}$$

$\gamma_0 s_j \le s_j' \le \gamma_1 s_j, \gamma' = 10\gamma_0^{-1}$, *$L(\varepsilon_{ij})$ is symmetric Bernoulli and* $\{Z_{ij}, \varepsilon_{ij}\colon 1 \le i, j \le n\}$ *is a set of n^2 mutually independent random variables.*

REMARK 3.3. The quantity $E\Phi(\sum_{1 \le i, j \le n} a_{ij} Z_{ij} \varepsilon_{ij})$ in Theorem 3.2 cannot be replaced by $\sum_{1 \le i, j \le n} E\Phi(a_{ij} Z_{ij})$. To see this, let $X, X_1, X_2, \dots$ be i.i.d. mean-zero random variables and let a_{ij} be defined as in Example 2. Then, as shown in Example 2,

$$E\left|\sum_{1 \le i, j \le 2n} a_{ij} X_i X_j\right| = 2E\left|\sum_{k=1}^{n} X_{2k-1} X_{2k}\right| \approx 2K_{X_1 X_2}^{(n)}.$$

Observe that $s_j \equiv E|X|$. Take any such X with

$$P\big(|X| > 10E|X|\big) > 0.$$

Then

$$\begin{aligned}\sum_{1 \le i, j \le 2n} E|a_{ij} Z_{ij}| &= 2nE|X_1 X_2| I\big(|X_1| > 10E|X|, |X_2| > 10E|X|\big) \\ &= 2n\Big(E|X| I\big(|X| > 10E|X|\big)\Big)^2 \gg K_{X_1 X_2}^{(n)} \quad \text{as } n \to \infty.\end{aligned}$$

Hence $\sum E|a_{ij} Z_{ij}|$ cannot be substituted in place of $E|\sum_{1 \le i,j \le n} \varepsilon_{ij} a_{ij} Z_{ij}|$ without sometimes causing Theorem 3.2 to fail.

In what follows we introduce the three missing lemmas.

LEMMA 3.1. *Take any $0 \le \lambda \le 1$ and any integer $n \ge 1$. Let $C = \{\bar{x} = (x_1, \dots, x_n)$ in R^n: $0 \le x_j \le 1$ and $\sum_{j=1}^{n} x_j = \lambda\}$. Then*

$$\sup_{\bar{x} \in C} \prod_{j=1}^{n} (1 - x_j) = \left(1 - \frac{\lambda}{n}\right)^n \le e^{-\lambda}.$$

The proof is left to the reader.

LEMMA 3.2. *If $Y_1, \dots, Y_n$ are independent, then, for all $c > 1$, $P > 0$,*

$$\prod_{j=1}^{n} P\left(|Y_j|^P < 2cE\left|\sum_{i=1}^{n} Y_i\right|^P\right) \ge \prod_{j=1}^{n} P\big(|Y_j|^P < cM^P\big) \ge 1 - c^{-1},$$

where $M^P = E\max_{1\le j\le n}|Y_j|^P$ *and so, by Lemma* 3.1,

$$\sum_{j=1}^{n} P\left(|Y_j|^P \ge 2cE\left|\sum_{i=1}^{n} Y_i\right|^P\right) \le \sum_{j=1}^{n} P(|Y_j|^P \ge cM^P) \le \log\left(\frac{c}{c-1}\right).$$

PROOF. Let

$$\tau_c = \begin{cases} \min\{j : 1 \le j \le n, |Y_j|^P| \ge cM^P\}, \\ \infty, \qquad \text{if no such } j \text{ exists.} \end{cases}$$

Note that

$$M^P \ge E|Y_{\tau_c}|^P I(\tau_c < \infty) \ge E\big(cM^P I(\tau_c < \infty)\big) = cM^P P(\max_{1\le j\le n} |Y_j|^P \ge cM^P),$$

so $P(\max_{1\le j\le n}|Y_j|^P < cM^P) \ge 1 - c^{-1}$. Solving for λ in the right-hand side of Lemma 3.1, we have $\sum_{j=1}^n x_j \le -\log(\prod_{j=1}^n (1-x_j))$ and so it follows that

$$\begin{aligned}\sum_{j=1}^{n} P(|Y_j|^P \ge cM^P) &\le \log\left(\frac{1}{P(\tau_c = \infty)}\right) \\ &\le \log\left(\frac{c}{c-1}\right).\end{aligned}$$

□

LEMMA 3.3. *Let* $Y_1, Y_2, \dots, Y_k$ *be independent mean-zero random variables and put* $s = E|Y_1 + \cdots + Y_k|$. *Then for any* $\gamma \ge 0$ *there exists* $q_\gamma < \infty$ *depending only on* γ *(and not on* k *or* $\{Y_j\}$*) such that*

$$E\sum_{j=1}^{k} Y_j^2 I(|Y_j| \le \gamma s) \le (q_\gamma s)^2.$$

PROOF. Let $Z = \sum_{j=1}^n Y_j^2 I(|Y_j| \le \gamma s)$ and $t^2 = EZ$. By Marshall's inequality, $P(Z \ge EZ - \sqrt{\operatorname{var} Z}) \ge \frac{1}{2}$. Note also that

$$\begin{aligned}&\operatorname{Var}\left(\sum_{j=1}^{k} Y_j^2 I(|Y_j| \le \gamma s)\right) \\ &\le \sum_{j=1}^{k} EY_j^4 I(|Y_j| \le \gamma s) \\ &\le \gamma^2 s^2 \sum_{j=1}^{k} EY_j^2 I(|Y_j| \le \gamma s) \\ &\equiv \gamma^2 s^2 t^2.\end{aligned}$$

Furthermore, note that $E\sqrt{\sum_{j=1}^n Y_j^2} \le C^* s$ for some $C^* < \infty$ independent of k

and $\{Y_j\}$. Hence, if $t \geq (\gamma + 2C^*)s$, then

$$\begin{aligned} C^*s &\geq E\left(\sum_{j=1}^{k} Y_j^2 I(|Y_j| \leq \gamma s)\right)^{1/2} \\ &\geq \left(t^2 - \gamma st\right)^{1/2} P\left(\sum_{j=1}^{k} Y_j^2 I(|Y_j| \leq \gamma s) \geq t^2 - \gamma st\right) \\ &\geq \frac{1}{2}\left(t(t - \gamma s)\right)^{1/2} \quad \text{(by Marshall's inequality)} \\ &\geq \frac{s}{2}\left((\gamma + 2C^*)2C^*\right)^{1/2} \\ &> C^*s, \end{aligned}$$

which gives a contradiction. Hence $t \leq (\gamma + 2C^*)s$. □

4. Higher orders of convexity. We extend the results of Section 3 to convex functions Φ for which $\Phi(x^{2^{-k-1}})$ is concave for some $k \geq 0$. Suppose both $\Phi(x)$ and $\Psi(x) = \Phi(\sqrt{|x|})$ are convex functions on $[0, \infty)$ with $\Phi(0) = 0$ and $\Phi(2x) \leq 2^{\beta}\Phi(x)$ for all $x \geq 0$ and some $2 \leq \beta < \infty$. Then, as shown in Section 2,

$$E\Phi\left(\left|\sum_{1 \leq i,j \leq n} a_{ij} X_i X_j\right|\right) \approx_{\beta} E\Psi\left(\sum_{1 \leq i,j \leq n} a_{ij}^2 X_i^2 \tilde{X}_j^2\right).$$

Let

$$W_n = \sum_{1 \leq i,j \leq n} a_{ij}^2 X_i^2 \tilde{X}_j^2. \tag{4.1}$$

By Jensen's inequality, $E\Psi(W_n) \geq \Psi(EW_n)$. Also,

$$\begin{aligned} E\Psi\left(|W_n - EW_n|\right) &\leq \tfrac{1}{2}\left(E\Psi(2W_n) + \Psi(2EW_n)\right) \quad \text{(by convexity)} \\ &\leq 2^{\beta-1}\left(E\Psi(W_n) + \Psi(EW_n)\right) \\ &\leq 2^{\beta} E\Psi(W_n). \end{aligned}$$

Therefore,

$$\frac{\max\left\{E\Psi\left(|W_n - EW_n|\right), \Psi(EW_n)\right\}}{E\Psi(W_n)} \leq 2^{\beta}.$$

We find the following upper bound for $E\Psi(W_n)$:

$$\begin{aligned} E\Psi(W_n) &\leq E\Psi(W_n - EW_n + EW_n) \\ &\leq \tfrac{1}{2}\left(E\Psi(2|W_n - EW_n|) + \Psi(2EW_n)\right) \quad \text{(by convexity)} \\ &\leq 2^{\beta-1}\left\{E\Psi\left(|W_n - EW_n|\right) + \Psi(EW_n)\right\} \\ &\leq 2^{\beta} \max\left\{E\Psi\left(|W_n - EW_n|\right), \Psi(EW_n)\right\}. \end{aligned}$$

Therefore,

$$E\Psi(W_n) \approx_\beta \max\{E\Psi(|W_n - EW_n|), \Psi(EW_n)\}.$$

The quantity $E\Psi(|W_n - EW_n|)$ can be further simplified (i.e., transformed).
Let

$$(4.2)\qquad \begin{aligned} W_{n,1} &= \sum_{1\le i,j\le n} a_{ij}^2(X_i^2 - EX_i^2)(\tilde{X}_j^2 - E\tilde{X}_j^2),\\ W_{n,2} &= \sum_{1\le i,j\le n} a_{ij}^2(X_i^2 - EX_i^2)E\tilde{X}_j^2,\\ W_{n,3} &= \sum_{1\le i,j\le n} a_{ij}^2(\tilde{X}_j^2 - E\tilde{X}_j^2)EX_i^2. \end{aligned}$$

Notice that

$$\begin{aligned} W_n - EW_n &= \sum_{1\le i,j\le n} a_{ij}^2(X_i^2\tilde{X}_j^2 - EX_i^2E\tilde{X}_j^2)\\ &= W_{n,1} + W_{n,2} + W_{n,3} \end{aligned}$$

and $L(W_{n,2}) = L(W_{n,3})$.
Conditioning on $\{X_i\}$ and using Jensen's inequality,

$$E\Psi(|W_n - EW_n|) \ge E\Psi(|W_{n,2}|).$$

Moreover, by the convexity of Ψ,

$$\begin{aligned} E\Psi(|W_{n,1}|) &= E\Psi(|W_n - EW_n - W_{n,2} - W_{n,3}|)\\ &\le \frac{1}{3}(E\{\Psi(3|W_n - EW_n|) + \Psi(3|W_{n,2}|) + \Psi(3|W_{n,3}|)\})\\ &\le \frac{2^\beta}{3}E\Psi(|W_n - EW_n|) + \frac{2^{\beta+1}}{3}E\Psi(|W_{n,2}|)\\ &\qquad\text{(since } W_{n,2} \text{ and } W_{n,3} \text{ are identically distributed)}\\ &\le 2^\beta E\Psi(|W_n - EW_n|). \end{aligned}$$

Therefore,

$$\max\{E\Psi(|W_{n,1}|), E\Psi(|W_{n,2}|))\} \le 2^\beta E\Psi(|W_n - EW_n|).$$

Upper-bounding $E\Psi(|W_n - EW_n|)$,

$$\begin{aligned} E\Psi(|W_n - EW_n|) &= E\Psi(|W_{n,1} + W_{n,2} + W_{n,3}|)\\ &\le \frac{1}{3}\{E\Psi(3|W_{n,1}|) + E\Psi(3|W_{n,2}|) + E\Psi(3|W_{n,3}|)\}\\ &\le \frac{2^\beta}{3}E\Psi(|W_{n,1}|) + \frac{2^{\beta+1}}{3}E\Psi(|W_{n,2}|)\\ &\le 2^\beta \max\{E\Psi(|W_{n,1}|), E\Psi(|W_{n,2}|). \end{aligned}$$

Combining our results,

$$E\,\Psi(|W_n|) \approx_\beta \max\{\Psi(E\,W_n),\, E\Psi(W_{n,1}),\, E\Psi(W_{n,2})\}. \tag{4.3}$$

The right-hand side of (4.3) represents a simplification (!) in the approximation of $E\Psi(W_n)$. The result is threefold: (1) The term $\Psi(EW_n)$ is (presumably) computable, and (2) so is $E\Psi(|W_{n,2}|)$. The reason $E\Psi(|W_{n,2}|)$ can be regarded as straightforward is that it is but a linear combination of independent mean zero-variables. (3) $E\Psi(|W_{n,1}|)$ involves a decoupled quadratic form of mean-zero variates for which Ψ is more nearly concave than Φ was.

By iterating this process we obtain the following extension of Theorem 3.2.

THEOREM 4.1. *Let Φ be a nonnegative, symmetric function with $\Phi(0) = 0$ such that, for some $k \geq 0$, $\Phi(x^{2^{-k}})$ is convex on $[0,\infty)$ and $\Phi(x^{2^{-k-1}})$ is concave on $[0,\infty)$. Take any $n \geq 2$ and any reals a_{ij} such that $a_{ij} = a_{ji}$ and $a_{ii} = 0$. Define $Y_{i,0} = X_i$, and, by induction,*

$$Y_{i,j} = Y_{i,j-1} - E\,Y^2_{i,j-1} \quad \textit{for } i = 1,\dots n,\ j \geq 1$$

and similarly for

$$\widetilde{Y}_{i,0},\ \widetilde{Y}_{i,1},\dots,\widetilde{Y}_{i,k} \quad \textit{starting from } \overline{X}_i.$$

We then have (with $\beta = 2^{k+1}$)

$$\begin{aligned} E\Phi\left(\left|\sum_{1\leq i,j\leq n} a_{ij}X_iX_j\right|\right) &\approx_\beta \max\Biggl\{\max_{1\leq m\leq k}\Phi\left(\left(\sum_{1\leq i,j\leq n}(a_{ij})^{2^m}\,E\,Y^2_{i,m-1}\,E\,\widetilde{Y}^2_{j,m-1}\right)^{2^{-m}}\right), \\ &\qquad \max_{1\leq m\leq k} E\Phi\left(\left|\sum_{1\leq i,j\leq n}(a_{ij})^{2^m}\left(E\,Y^2_{i,m-1}\right)\widetilde{Y}^2_{j,m}\right|^{2^{-m}}\right), \\ &\qquad E\Phi\left(\left|\sum_{1\leq i,j\leq n}(a_{ij})^{2^k}\,Y_{i,k}\,\widetilde{Y}_{j,k}\right|^{2^{-k}}\right)\Biggr\}. \end{aligned} \tag{4.4}$$

Of the $2k+1$ quantities whose maximum is required, the first k are constants directly computable from the X_i, the second k involve expectations of functions of sums of independent variables, and so they too are computable from the one-dimensional X_i-distributions. The quantity

$$E\Phi\left(\left(\sum_{1\leq i,j\leq n}(a_{ij})^{2^k}\,Y_{i,k}\,\widetilde{Y}_{i,k}\right)^{2^{-k}}\right)$$

remains. It is approximated by the two quantities described in Theorem 3.2.

Finally, to see that $\beta = 2^{k+1}$ will do, note that since $\Phi(x^{2^{-k-1}})$ is concave, $\Phi((2x)^{2^{-k-1}}) \le 2\Phi(x^{2^{-k-1}})$. Equivalently, $\Phi(2^{2^{-k-1}}x) \le 2\Phi(x)$. Hence, by induction, we have

$$\Phi(2x) = \Phi\left(x \prod_{j=1}^{2^{k+1}} 2^{2^{-k-1}}\right) \le \Phi(x) \prod_{j=1}^{2^{k+1}} 2 = 2^{2^{k+1}}\Phi(x).$$

5. Application of decoupling: order of magnitude for a nontrivial martingale. This section presents a special case of the quadratic form for which a more elegant solution is possible.

THEOREM 5.1. *Let $X_1, X_2, \ldots$ be independent mean-zero random variables. Let $\{\tilde{X}_i\}$ be an independent copy of $\{X_i\}$. For $P \ge 1$ assume $E|X_1|^P \le E|X_2|^P \le \cdots \le E|X_n|^P$. Then*

$$(5.1)\quad \begin{aligned}(96)^{-P}E\left|\sum_{j=1}^{n-1}X_j\right|^P E\left|\sum_{j=1}^{n}X_j\right|^P &\le E\left|\sum_{1\le i\ne j\le n} X_iX_j\right|^P \\ &\le E\max_{2\le k\le n}\left|\sum_{1\le i\ne j\le k} X_iX_j\right|^P \\ &\le 25(16)^P E\left|\sum_{i=1}^{n-1}X_j\right|^P E\left|\sum_{j=1}^{n}X_j\right|^P.\end{aligned}$$

The corresponding qualitative result for i.i.d. random variables was obtained by de la Peña (1988). A seemingly more general version of the next theorem can be found in the introduction. However, it is easy to see by reordering the variates that both versions are the same. We state the theorem in this way since the proof becomes more transparent.

PROOF. Let $S_k = X_1 + \cdots + X_k$. We obtain the lower bound first. By using

Theorem 2.1, one gets

$$
\begin{aligned}
E\left|\sum_{1\le i<j\le n} X_iX_j\right|^P &= \frac{1}{2^P}E\left|\sum_{1\le i\ne j\le n} X_iX_j\right|^P \\
&\ge \frac{1}{12^P}E\left|\sum_{1\le i\ne j\le n} X_i\tilde{X}_j\right|^P \\
&= \frac{1}{12^P}E\left|\sum_{j=1}^{n}(S_n - X_j)\tilde{X}_j\right|^P.
\end{aligned}
$$

By conditioning on $(X_1,\dots,X_{k_n})$, $(\tilde{X}_{k_n+1},\dots,\tilde{X}_n)$ and using Jensen's inequality, the preceding quantity is not less than

$$
\begin{aligned}
&\frac{1}{12^P}E\left|\sum_{j=k_n+1}^{n} S_{k_n}\tilde{X}_j\right|^P \\
&= \frac{1}{12^P}E\left|(S_{k_n})(\tilde{S}_n - \tilde{S}_{k_n})\right|^P \qquad (5.2)\\
&= \frac{1}{12^P}E|S_{k_n}|^P E|S_n - S_{k_n}|^P.
\end{aligned}
$$

We can obtain the lower bound by considering two cases.

Case 1. Assume

$$
E|X_n|^P > \frac{E|S_{n-1}|^P}{2^{2P}}. \qquad (5.3)
$$

Letting $k_n = n-1$ above, we get

$$
E\left|\sum_{1\le i<j\le n} X_iX_j\right|^P \ge \frac{1}{12^P}E|S_{n-1}|^P E|X_n|^P.
$$

We need to bound $E|X_n|^P$ below in terms of $E|S_n|^P$:

$$
\begin{aligned}
E|S_n|^P &\le 2^{P-1}\left(E|X_n|^P + E|S_{n-1}|^P\right) \\
&\le 2^{P-1}\left(1+2^{2P}\right)E|X_n|^P \\
&\le 8^P E|X_n|^P
\end{aligned}
$$

and hence

$$E\left|\sum_{1\le i<j\le n} X_i X_j\right|^P \ge (96)^{-P} E|S_{n-1}|^P E|S_n|^P.$$

Case 2. Now assume $E|X_n|^P \le E|S_{n-1}|^P/2^{2P}$ and let

$$k_n = \min\left\{k: E|S_k|^P \ge \frac{3E|S_{n-1}|^P}{2^{2P+1}}\right\}.$$

Then we have

$$\begin{aligned} E|S_{k_n}|^P &\le 2^{P-1}\left(E|S_{k_n-1}|^P + E|X_{k_n}|^P\right) \\ &\le 2^{P-1}\left(\frac{3}{2^{2P+1}} + \frac{1}{2^{2P}}\right)E|S_{n-1}|^P \qquad (5.4) \\ &\le \frac{5}{2^{P+2}} E|S_n|^P. \end{aligned}$$

We also have that

$$\begin{aligned} 2^{1-P}E|S_n|^P &\le E|S_n - S_{k_n}|^P + E|S_{k_n}|^P \\ &\le E|S_n - S_{k_n}|^P + \frac{5}{2^{P+2}} E|S_n|^P. \end{aligned}$$

Therefore,

$$E|S_n - S_{k_n}|^P \ge \frac{3E|S_n|^P}{2^{P+2}}.$$

Then, from (5.2),

$$\begin{aligned} E\left|\sum_{1\le i<j\le n} X_i X_j\right|^P &\ge (12)^{-P} E|S_{k_n}|^P E|S_n - S_{k_n}|^P \\ &\ge (12)^{-P} \frac{3E|S_{n-1}|^P}{2^{2P+1}}\left(\frac{3E|S_n|^P}{2^{P+2}}\right) \\ &\ge (96)^{-P} E|S_{n-1}|^P E|S_n|^P. \end{aligned}$$

PROOF OF THE UPPER BOUND. What follows is an alternative proof of the upper bound from the one we had in mind. It gives better constants as $P \to \infty$.

The approach was suggested by P. Hitczenko:

$$
\begin{aligned}
&E \max_{2\le k\le n}\left|\sum_{1\le i\ne j\le k} X_i X_j\right|^P \\
&\le 2^{P-1}\left\{E \max_{2\le k\le n}\left|\sum_{1\le i<j\le k} X_i X_j\right|^P + E \max_{2\le k\le n}\left|\sum_{1\le j<i\le k} X_i X_j\right|^P\right\} \\
&= 2^P E \max_{2\le k\le n}\left|\sum_{1\le i<j\le k} X_i X_j\right|^P \\
&\le 8^P E \max_{2\le k\le n}\left|\sum_{j=2}^{k} S_{j-1}\tilde{X}_j\right|^P \\
&\qquad\qquad \text{(by the upper-bound in (2.2) of Theorem 2.1)} \\
&\le 5(8)^P E\left|\sum_{j=2}^{n} S_{j-1}\tilde{X}_j\right|^P \quad \text{(by conditioning on } \{S_{j-1}\}_{2\le j\le n}, \\
&\qquad\qquad \text{using Lemma 6.3 and unconditioning)} \\
&\le 5(16)^P E \max_{2\le j\le n}|S_{j-1}|^P E\left|\sum_{j=2}^{n}\tilde{X}_j\right|^P \\
&\qquad\qquad \text{(by Lemma 6.4, the contraction principle)} \\
&\le 25(16)^P E|S_{n-1}|^P E|S_n|^P \quad \text{(in part by Lemma 6.3).}
\end{aligned}
$$

□

6. Supplementary lemmas. Our first lemma is a decoupling result due to Hitczenko (1988).

LEMMA 6.1. *Let $U = \{U_j\}$, $V = \{V_j\}$ be adapted to A_j, where U_j, $V_j \ge 0$. Then U and V are said to be tangent if*

$$
P(U_j > t \mid A_{j-1}) \overset{ac}{=} P(V_j > t \mid A_{j-1}) \quad \forall\, t > 0. \tag{6.1}
$$

If U and V are tangent, then

$$
E\Phi\left(\sum U_j\right) \approx_\alpha E\Phi\left(\sum V_j\right) \tag{6.2}
$$

for all Φ such that $\Phi: R_+ \to R_+$, Φ increasing, continuous and $\Phi(2x) \le \alpha\Phi(x)$, $x \ge 0$, for some $\alpha < \infty$.

In (7.3) of Remark 7.1 of Klass (1980), a method of obtaining the order of magnitude of expectations of functions of sums of independent random variables is presented in the form we need. The result is as follows.

LEMMA 6.2. *Let $\overline{Z}_n = (Z_1, Z_2, \ldots, Z_n)$ be a vector-valued random variable, where Z_i are independent mean-zero random variables. Let Φ be in F_β. Define $K_\Phi(\overline{Z}_n)$ as the unique nonnegative real number K such that*

$$K^2 = \sum_{i=1}^{n} EZ_i^2 I\big(|Z_i| \le K\big) + \frac{K^2}{\Phi(K)} \sum_{i=1}^{n} E\Phi(Z_i) I\big(|Z_i| > K\big). \tag{6.3}$$

Then

$$E\Phi\left(\sum_{i=1}^{n} Z_i\right) \approx_\beta E \max_{1 \le m \le n} \Phi\left(\sum_{i=1}^{m} Z_i\right) \approx_\beta \Phi\big(K_\Phi(\overline{Z}_n)\big). \tag{6.4}$$

From this it is easily seen that $(Z_1, Z_2, \ldots, Z_n)$ has the same K function as $(Z_1\varepsilon_1, Z_2\varepsilon_2, \ldots, Z_n\varepsilon_n)$, where $\{\varepsilon_i\}$ is a sequence of i.i.d. r.v.'s such that $P(\varepsilon_i = 1) = P(\varepsilon_i = -1) = \frac{1}{2}$ and $\{\varepsilon_i\}$ is independent of $\{Z_i\}$. This gives us the result

$$E\Phi\left(\sum_{i=1}^{n} Z_i\right) \approx_\beta E\Phi\left(\sum_{i=1}^{n} Z_i\varepsilon_i\right). \tag{6.5}$$

In the next lemma we restate Theorem 1 of Klass (1993).

LEMMA 6.3. *Let $X_1, X_2, \ldots, X_n$ be a sequence of mean-zero random variables and let $S_n = X_1 + \cdots + X_n$. Also let Φ be any nonnegative convex function on $[0, \infty)$. Then*

$$E\Phi\left(\max_{1 \le j \le n} |S_j|\right) \le 5E\Phi\big(|S_n|\big). \tag{6.6}$$

Moreover, the 5 can be replaced by 3 if the X_j's are i.i.d.

The final lemma, known as the contraction principle, is Lemma 4.1 of Hoffman-Jørgensen (1974).

LEMMA 6.4 (Hoffman-Jørgensen). *Let $X_1, X_2, \ldots, X_n$ be independent mean-zero Banach space valued random variables. Then for all real numbers $a_1, \ldots, a_n$ and $P \ge 1$ such that $\max_{1 \le j \le n} E\|X_j\|^P < \infty$,*

$$E\left\|\sum_{j=1}^{n} a_j X_j\right\|^P \le 2^P \max_{1 \le j \le n} |a_j|^P E\left\|\sum_{j=1}^{n} X_j\right\|^P. \tag{6.7}$$

REFERENCES

BOURGAIN, J. and TZAFRIRI, L. (1987). Invertibility of "large" submatrices with applications to the geometry of Banach spaces and harmonic analysis. *Israel J. Math.* **57** 137–224.

BURKHOLDER, D. L., DAVIS, B. J. and GUNDY, R. F. (1972). Inequalities for convex functions of operators on martingales. *Proc. Sixth Berkeley Symp. Math. Statist. Probab.* **2** 223–240. Univ. California Press, Berkeley.

CAMBANIS, S. ROSINSKI, J. and WOYCZINSKI, W. A. (1985). Convergence of quadratic forms in p-stable random variables and θ_p-radonifying operators. *Ann. Probab.* **13** 885–897.

DE ACOSTA, A. (1987). A decoupling inequality for multilinear functions of stable vectors. *Probab. Math. Statist.* **8** 71–76.

DE LA PEÑA, V. H. (1988). L-Bounds of best possible type for martingales, degenerate U-statistics, and certain linear forms. Ph.D. dissertation, Dept. Statistics, Univ. California, Berkeley.

DE LA PEÑA, V. H. (1990). Bounds on the expectation of functions of martingales and sums of positive R. V.'s in terms of norms of sums of independent random variables. *Proc. Amer. Math. Soc.* **108** 233–239.

HITCZENKO, P. (1988). Comparison of moments of tangent sequences of random variables. *Probab. Theory Related Fields* **78** 223–230.

HOFFMAN-JØRGENSEN, J. (1974). Sums of independent Banach space valued random variables. *Studia Math.* **52** 159–186.

JOHNSON, W. B. and SCHECHTMAN, G. (1989). Sums of independent random variables in rearrangement invariant function spaces. *Ann. Probab.* **17** 789–808.

KLASS, M. J. (1980). Precision bounds for the relative error in the approximation of $E|S_n|$ and extensions. *Ann. Probab.* **8** 350–367.

KLASS, M. J. (1981). A method of approximating expectations of functions of sums of independent random variables. *Ann. Probab.* **9** 413–428.

KLASS, M. J. (1993). Ratio prophet inequalities for convex functions of partial sums. *Statist. Probab. Lett.* **17** 205–209.

KRAKOWIAK, W. and SZULGA, J. (1988). On a p-stable multiple integral. II. *Probab. Theory Related Fields* **78** 449–453.

KWAPIEN, S. (1987). Decoupling inequalities for polynomial chaos. *Ann. Probab.* **15** 1062–1071.

MAZUR, S. and ORLICZ, W. (1935). Grundlegende Eigenschaften der polynomischen Operationen. *Studia Math.* **5** 50–60, 179–189.

MCCONNELL, T. and TAQQU, M. (1984). Double integration with respect to symmetric stable processes. Technical Report 618, Cornell Univ.

MCCONNELL, T. and TAQQU, M. (1986). Decoupling inequalities for multilinear forms in independent symmetric random variables. *Ann. Probab.* **14** 943–954.

ZINN, J. (1985). Comparison of martingale difference sequences. *Probability in Banach Spaces V. Lecture Notes in Math.* **1153** 453–457. Springer, New York.

ZINN, J. (1989). Personal communication.

DEPARTMENTS OF MATHEMATICS
AND STATISTICS
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720