# Symmetrization and Decoupling of Combinatorial Random Elements

Lutz Dümbgen*

published 1998 in *Statistics & Probability Letters* ***39****(4)*, 355-361.

**Abstract.** Let $\Phi = (\phi_{ij})_{1\le i,j\le n}$ be a random matrix whose components $\phi_{ij}$ are independent stochastic processes on some index set $\mathcal{T}$. Let $S = \sum_{i=1}^n \phi_{i\Pi(i)}$, where $\Pi$ is a random permutation of $\{1,2,\ldots,n\}$, independent from $\Phi$. This random element is compared with its symmetrized version $S^o := \sum_{i=1}^n \xi_i\phi_{i\Pi(i)}$ and its decoupled version $\widetilde{S} := \sum_{i=1}^n \phi_{i\widetilde{\Pi}(i)}$. Here $\xi = (\xi_i)_{1\le i\le n}$ is a Rademacher sequence and $\widetilde{\Pi}$ is uniformly distributed on $\{1,2,\ldots,n\}^n$ such that $\Phi$, $\Pi$, $\widetilde{\Pi}$ and $\xi$ are independent. It is shown that for a broad class of convex functions $\Psi$ on $\mathbb{R}^{\mathcal{T}}$ the following symmetrization and decoupling inequalities hold:

$$\mathbb{E}\,\Psi(S - \mathbb{E}\,S) \ \le \ \begin{cases} \mathbb{E}\,\Psi(\kappa S^o), \\ \mathbb{E}\,\Psi(\gamma(\widetilde{S} - \mathbb{E}\,S)), \end{cases}$$

where $\kappa, \gamma > 0$ are universal constants.

**Keywords:** exponential inequality, linear rank statistic, permutation bridge, random permutation

**Correspondence to:** Lutz Duembgen, Institute of Mathematical Statistics and Actuarial Science, University of Bern, Sidlerstrasse 5, CH-3012 Bern, Switzerland;

e-mail: `duembgen@stat.unibe.ch`

*Research supported in part by European Union Human Capital and Mobility Program ERB CHRX-CT 940693.

## 1 Introduction and statement of results

Symmetrization and decoupling are a powerful tool in order to obtain moment and tail inequalities for certain stochastic processes. Very informative references for the merits of symmetrizing empirical processes or, more generally, sums of independent processes are Pollard (1990) and van der Vaart and Wellner (1996). Another important application of such tools are U-processes. We refer to Arcones and Giné (1993) who utilized symmetrization and decoupling inequalities of de la Peña (1992).

The present note treats stochastic processes of the following form: Let $\Phi = (\phi_{ij})_{1\le i,j\le n}$ be a random matrix whose components $\phi_{ij}$ are independent stochastic processes on some set $\mathcal{T}$ (or just real random variables), where $n \ge 2$. Let $S = \sum_{i=1}^n \phi_{i\Pi(i)}$, where $\Pi$ is a random permutation of $\{1, 2, \ldots, n\}$, independent from $\Phi$. Processes of this type appear in different statistical applications involving randomization. An early reference is Hoeffding (1951) who considered a fixed matrix $\Phi \in \mathbb{R}^{n\times n}$ and proved a central limit theorem for $S$. General processes and examples are treated in Dümbgen (1994). We mention just three:

Example 1 (Sampling without replacement from a finite population). For $1 \le m < n$ let $G_1, G_2, \ldots, G_m$ be a random sample without replacement from a fixed finite collection $f_1, f_2, \ldots, f_n$ of functions on $\mathcal{T}$. If we define $\phi_{ij} := 1\{i \le m\} f_j$, then the sum $\sum_{i=1}^m G_i$ is distributed as $S$. For this special case Hoeffding (1963) obtained the following decoupling inequality (see also LeCam 1986, Lemma 16.7.2): Let $\widetilde{G}_1, \widetilde{G}_2, \ldots, \widetilde{G}_m$ be a random sample *with* replacement from $f_1, f_2, \ldots, f_n$. Then for any convex function $\Psi : \mathbb{R}^{\mathcal{T}} \to \mathbb{R}$,

$$(1) \qquad \mathbb{E}\,\Psi\Big(\sum_{i=1}^m G_i\Big) \;\le\; \mathbb{E}\,\Psi\Big(\sum_{i=1}^m \widetilde{G}_i\Big).$$

Example 2 (Linear rank statistics). In order to test exchangeability of a random vector $X \in \mathbb{R}^n$ one may employ a linear rank statistic $\sum_{i=1}^n a_i b_{R(i)}$ with fixed vectors $a, b \in \mathbb{R}^n$, where $R(i)$ is the rank of $X_i$ in the sample $X$. Under the hypothesis of exchangeability of $X$, the rank statistic is distributed as $S$ if $\phi_{ij} := a_i b_j$. It will be shown below that this random variable $S$ satisfies a sub-Gaussian tail inequality involving only

$$\|a\|_2 \;:=\; \Big(\sum_{i=1}^n a_i^2\Big)^{1/2} \quad \text{and} \quad \|b\|_\infty \;:=\; \max_{1\le i\le n} |b_i|.$$

Such an inequality has applications in nonparametric regression; cf. Dümbgen (1998).

EXAMPLE 3 ("Permutation bridges"). Let $a \in \mathbb{R}^n$ be a fixed vector such that

$$a_+ := \sum_{i=1}^n a_i \ = \ 0 \quad \text{and} \quad \|a\|_2 \ = \ 1. \tag{2}$$

For $t \in [0,1]$ define

$$B(t) \ := \ \sum_{i=1}^n 1\{i \le nt\} a_{\Pi(i)}.$$

For large $n$ this process has approximately the same covariance function as a Brownian bridge. In fact, Theorem 24.1 of Billingsley (1968) states that $B$ converges in distribution in $\mathcal{D}[0,1]$ to a Brownian bridge as $\|a\|_\infty \to 0$. It will be shown below that the supremum-norm of $B$ has sub-Gaussian tails, *uniformly* for all $a$ satisfying Condition (**??**).

The random element $S$ is compared with its symmetrized version $S^o := \sum_{i=1}^n \xi_i \phi_{i\Pi(i)}$ and its decoupled version $\widetilde{S} := \sum_{i=1}^n \phi_{i\widetilde{\Pi}(i)}$. Here $\xi = (\xi_i)_{1 \le i \le n}$ is a Rademacher sequence, i.e. uniformly distributed on $\{-1,1\}^n$, and $\widetilde{\Pi}$ is uniformly distributed on $\{1,2,\ldots,n\}^n$. Moreover, $\Phi$, $\Pi$, $\widetilde{\Pi}$ and $\xi$ are independent. Throughout we assume that

$$\mu_{ij} \ := \ \mathbb{E}\, \phi_{ij}$$

exists in $\mathbb{R}^{\mathcal{T}}$ for all $i,j$. Then

$$\mathbb{E}\, S \ = \ \mathbb{E}\, \widetilde{S} \ = \ \frac{1}{n} \sum_{i,j=1}^n \mu_{ij} \quad \text{and} \quad \mathbb{E}\, S^o \ \equiv \ 0.$$

The goal is to compare moments of $\|S - \mathbb{E}\, S\|$ with moments of $\|S^o\|$ and $\|\widetilde{S} - \mathbb{E}\, S\|$, where

$$\|x\| \ := \ \sup_{t \in \mathcal{T}_o} |x(t)|$$

for any $x : \mathcal{T} \to \mathbb{R}$ with some countable subset $\mathcal{T}_o$ of $\mathcal{T}$. More generally, in what follows let $\Psi : \mathbb{R}^{\mathcal{T}} \to ]-\infty, \infty]$ be a convex function of the form

$$\Psi(x) \ := \ \sup_{t \in \mathcal{T}_o} \psi_t(x(t))$$

with convex functions $\psi_t : \mathbb{R} \to ]-\infty,\infty]$. A special example is given by $\Psi(x) = \psi(\|x\|)$ with a convex, nondecreasing function $\psi$ on $[0,\infty[$. Here is our symmetrization inequality for $S$:

**Theorem 1.**

$$\mathbb{E}\, \Psi(S - \mathbb{E}\, S) \ \le \ \mathbb{E}\, \Psi(\kappa S^o),$$

*where* $8 \ge \kappa := 2(n/\lfloor n/2 \rfloor + 1) \to 6$ *as* $n \to \infty$.

Here $\lfloor r \rfloor$ denotes the largest integer not exceeding $r$. The proof of Theorem **??** is an extension of a symmetrization argument introduced in Dümbgen (1994). In fact, Theorem **??** can be used to derive the main results of the latter paper under slightly stronger moment conditions. Other applications are sub-Gaussian inequalities for Examples 2 and 3.

**Corollary 1.** **(a)** *Let $a, b \in \mathbb{R}^n$ such that $a_+ = 0$ or $b_+ = 0$. Then for all $\eta > 0$,*

$$\mathbb{P}\Big\{\sum_{i=1}^{n} a_i b_{\Pi(i)} \geq \eta\Big\} \;\leq\; \exp\Big(-\frac{\eta^2}{2\kappa^2 \|a\|_2^2 \|b\|_\infty^2}\Big).$$

**(b)** *Let $B = B(\cdot \mid a)$ be defined as in Example 3, where $a \in \mathbb{R}^n$ satisfies Condition (**??**). Then for all $\eta > 0$,*

$$\mathbb{P}\Big\{\sup_{t\in[0,1]} |B(t)| \geq \eta\Big\} \;\leq\; 4\exp\Big(-\frac{\eta^2}{2\kappa^2}\Big).$$

Example 1 and intuition suggest that $\|S - \mathbb{E}\,S\|$ is dominated in some sense by $\|\widetilde{S} - \mathbb{E}\,S\|$. Starting from Theorem **??** we shall deduce the following decoupling inequality:

**Theorem 2.**

$$\mathbb{E}\,\Psi(S - \mathbb{E}\,S) \;\leq\; \mathbb{E}\,\Psi_{\mathrm{s}}(\gamma(\widetilde{S} - \mathbb{E}\,S)),$$

*where $\Psi_{\mathrm{s}}(x) := (\Psi(x) + \Psi(-x))/2$ and $16/(1 - e^{-1}) \geq \gamma := 2\kappa/(1 - (1 - 1/n)^n) \to 12/(1 - e^{-1})$ as $n \to \infty$.*

Presumably the constant $\gamma$ is suboptimal. In view of (**??**) we conjecture that Theorem **??** is true with $\gamma = 1$.

## 2 Proofs

In what follows we'll frequently use a classical symmetrization inequality which is essentially an application of Jensen's inequality; see the derivation of Pollard (1990, Theorem 2.2).

**Lemma 1.** *Let $Z$ and $Z'$ be independent stochastic processes on $\mathcal{T}$ such that $\mathbb{E}\,Z'$ is well-defined in $\mathbb{R}^{\mathcal{T}}$. Then*

$$\mathbb{E}\,\Psi(Z - \mathbb{E}\,Z') \;\leq\; \mathbb{E}\,\Psi(Z - Z'). \qquad \Box$$

**Proof of Theorem ??.** *First step:* In order to get the Rademacher variables $\xi_i$ into play we replace the whole sum $S$ with $S_J := \sum_{i\in J} \phi_{i\Pi(i)}$, where $J := \{1, 2, \ldots, m\}$ and

$m := \lfloor n/2 \rfloor$. The latter sum is compared with

$$S_J^* \;:=\; \sum_{i \in J} \phi_{i\Pi(m+i)}.$$

These processes $S_J$ and $S_J^*$ are identically distributed but possibly dependent. Conditionally on $\Pi(J)$ they are independent but may have different distributions. Note that

$$S_J - S_J^* \;=\; \sum_{i \in J} (\phi_{i\Pi(i)} - \phi_{i\Pi(m+i)}) \;=_{\mathcal{L}}\; \sum_{i \in J} \xi_i (\phi_{i\Pi(i)} - \phi_{i\Pi(m+i)}),$$

which is easily verified by conditioning on the two-point sets $\{\Pi(i), \Pi(m+i)\}$, $1 \le i \le m$. Hence with $\mathbb{E}_o := \mathbb{E}(\cdot \,|\, \Phi, \Pi)$ one can deduce that for arbitrary $c > 0$,

$$\begin{aligned}
\mathbb{E}\,\Psi(c(S_J - S_J^*)) &\le \mathbb{E}\Big(2^{-1}\Psi\Big(2c\sum_{i \in J}\xi_i\phi_{i\Pi(i)}\Big) + 2^{-1}\Psi\Big(-2c\sum_{i \in J}\xi_i\phi_{i\Pi(m+i)}\Big)\Big) \\
&= \mathbb{E}\,\Psi\Big(2c\sum_{i \in J}\xi_i\phi_{i\Pi(i)}\Big) \\
&= \mathbb{E}\,\mathbb{E}_o\,\Psi\Big(2c\sum_{i \in J}\xi_i\phi_{i\Pi(i)} + \mathbb{E}_o\sum_{i \notin J}\xi_i\phi_{i\Pi(i)}\Big) \\
&\le \mathbb{E}\,\mathbb{E}_o\,\Psi\Big(2c\sum_{i=1}^{n}\xi_i\phi_{i\Pi(i)}\Big) \quad \text{[Lemma ??]} \\
&= \mathbb{E}\,\Psi(2cS^o). \qquad (3)
\end{aligned}$$

*Second step:* Now we compare $S_J - \mathbb{E}\,S_J$ with $S_J - S_J^*$. Letting $\mathbb{E}_o := \mathbb{E}(\cdot \,|\, \Pi(J))$ one may write

$$\begin{aligned}
\mathbb{E}_o\,S_J &= \frac{1}{m}\sum_{i \in J}\sum_{j \in \Pi(J)}\mu_{ij}, \\
\mathbb{E}_o\,S_J^* &= \frac{1}{n-m}\sum_{i \in J}\sum_{j \notin \Pi(J)}\mu_{ij}, \\
\mathbb{E}\,S_J &= \frac{1}{n}\sum_{i \in J}\sum_{j=1}^{n}\mu_{ij} \;=\; \frac{m}{n}\,\mathbb{E}_o\,S_J + \frac{n-m}{n}\,\mathbb{E}_o\,S_J^*.
\end{aligned}$$

If we define $Z := S_J - \mathbb{E}\,S_J$ and $Z^* := S_J^* - \mathbb{E}\,S_J$ it follows that

$$\mathbb{E}_o\,Z^* \;=\; -\gamma\,\mathbb{E}_o\,Z \quad \text{with } \gamma := \frac{m}{n-m} \in \,]0,1].$$

Given $\Pi(J)$, the processes $Z$ and $Z^*$ are independent. Hence for arbitrary $d > 0$ and $\lambda \in \,]0,1[$,

$$\begin{aligned}
&\mathbb{E}_o\,\Psi(d(S_J - \mathbb{E}\,S_J)) = \mathbb{E}_o\,\Psi(dZ) \\
&\quad= \mathbb{E}_o\,\Psi(d(Z - \mathbb{E}_o\,Z^* - \gamma\,\mathbb{E}_o\,Z))
\end{aligned}$$

$$
\begin{aligned}
&\le \; \mathbb{E}_o \Psi(d(Z - Z^*) - d\gamma\, \mathbb{E}_o Z)) \quad \text{[Lemma **??**]} \\
&= \; \mathbb{E}_o \Psi\Big(d(Z - Z^*) + d(m/n)\, \mathbb{E}_o(Z^* - Z)\Big) \\
&\le \; \lambda\, \mathbb{E}_o \Psi\Big(\frac{d}{\lambda}(Z - Z^*)\Big) + (1-\lambda)\Psi\Big(\frac{d(m/n)}{1-\lambda}\, \mathbb{E}_o(Z^* - Z)\Big) \\
&\le \; \lambda\, \mathbb{E}_o \Psi\Big(\frac{d}{\lambda}(Z - Z^*)\Big) + (1-\lambda)\, \mathbb{E}_o \Psi\Big(\frac{d(m/n)}{1-\lambda}(Z^* - Z)\Big) \quad \text{[Lemma **??**]} \\
&= \; \frac{n}{n+m}\, \mathbb{E}_o \Psi\Big(d(1+m/n)(S_J - S_J^*)\Big) + \frac{m}{n+m}\, \mathbb{E}_o \Psi\Big(d(1+m/n)(S_J^* - S_J)\Big)
\end{aligned}
$$

if $\lambda := n/(n+m)$. Integrating both sides with respect to the distribution of $\Pi(J)$ and plugging in (**??**) yields

$$
\mathbb{E}\, \Psi(d(S_J - \mathbb{E}\, S_J)) \;\le\; \mathbb{E}\, \Psi(2d(1+m/n)S^o). \tag{4}
$$

*Third and final step:* Inequality (**??**) remains valid if $S_J$ is replaced with $S_L := \sum_{i\in L} \phi_{i\Pi(i)}$, where $L$ is an arbitrary subset of $\{1, 2, \ldots, n\}$ containing $m$ points. One may write

$$
S \;=\; \sum_{i=1}^n \phi_{i\Pi(i)} \binom{n-1}{m-1}^{-1} \sum_{L:\#L=m} 1\{i \in L\} \;=\; \binom{n}{m}^{-1} \sum_{L:\#L=m} (n/m) S_L.
$$

Consequently, by convexity of $\Psi$ and (**??**),

$$
\mathbb{E}\, \Psi(S - \mathbb{E}\, S) \;\le\; \binom{n}{m}^{-1} \sum_{L:\#L=m} \mathbb{E}\, \Psi\Big((n/m)(S_L - \mathbb{E}\, S_L)\Big) \;\le\; \mathbb{E}\, \Psi(2(n/m+1)S^o). \;\Box
$$

**Proof of Corollary ??.** As for part (a), Theorem **??** entails that

$$
\begin{aligned}
\mathbb{E} \exp\Big(\lambda \sum_{i=1}^n a_i b_{\Pi(i)}\Big) \;&\le\; \mathbb{E} \exp\Big(\lambda\kappa \sum_{i=1}^n \xi_i a_i b_{\Pi(i)}\Big) \\
&=\; \mathbb{E}\, \mathbb{E}\Big(\exp\Big(\lambda\kappa \sum_{i=1}^n \xi_i a_i b_{\Pi(i)}\Big) \,\Big|\, \Pi\Big) \\
&=\; \mathbb{E} \prod_{i=1}^n \cosh(\lambda\kappa a_i b_{\Pi(i)}) \\
&\le\; \mathbb{E} \exp\Big(\lambda^2\kappa^2 \sum_{i=1}^n a_i^2 b_{\Pi(i)}^2/2\Big) \\
&\le\; \exp\Big(\lambda^2\kappa^2 \|a\|_2^2 \|b\|_\infty^2/2\Big).
\end{aligned}
$$

Thus, by Tshebyshev's inequality, $\mathbb{P}\Big\{\sum_{i=1}^n a_i b_{\Pi(i)} \ge \eta\Big\}$ is not greater than

$$
\inf_{\lambda>0} \mathbb{E} \exp\Big(\lambda \sum_{i=1}^n a_i b_{\Pi(i)}\Big) \exp(-\lambda\eta) \;\le\; \exp\Big(-\eta^2/(2\kappa^2 \|a\|_2^2 \|b\|_\infty^2)\Big).
$$

Part (b) is proved similarly: By Theorem **??**, $\mathbb{E}\exp(\lambda\|B\|) \le \mathbb{E}\exp(\lambda\kappa\|B^o\|)$ for any $\lambda > 0$, where $B^o(t) = \sum_{i=1}^n 1\{i \le nt\}\xi_i a_{\Pi(i)}$. Conditional on $\Pi$, the process $B^o$ has independent, symmetrically distributed increments. Consequently,

$$\begin{aligned}
\mathbb{E}\exp(\lambda\kappa\|B^o\|) &= \int_0^\infty \mathbb{P}\Big\{\|B^o\| > \log(r)/(\lambda\kappa)\Big\}\,dr \\
&\le 2\int_0^\infty \mathbb{P}\Big\{|B^o(1)| > \log(r)/(\lambda\kappa)\Big\}\,dr \quad \text{[Lévy's inequality]} \\
&\le 4\int_0^\infty \mathbb{P}\Big\{B^o(1) > \log(r)/(\lambda\kappa)\Big\}\,dr \\
&= 4\,\mathbb{E}\exp(\lambda\kappa B^o(1)) \\
&\le 4\exp(\lambda^2\kappa^2/2),
\end{aligned}$$

see also the proof of part (a). Again this inequality yields the assertion via Tshebyshev's inequality. □

**Proof of Theorem ??.** Since

$$S - \mathbb{E}\,S = \sum_{i=1}^n \Big(\phi_{i\Pi(i)} - \frac{1}{n}\sum_{j=1}^n \mu_{ij}\Big) \quad \text{and} \quad \widetilde{S} - \mathbb{E}\,S = \sum_{i=1}^n \Big(\phi_{i\widetilde{\Pi}(i)} - \frac{1}{n}\sum_{j=1}^n \mu_{ij}\Big),$$

we may and do assume that $\sum_{j=1}^n \mu_{ij} \equiv 0$ for $1 \le i \le n$, so that $\mathbb{E}\,S \equiv 0$. Let $K$ be a random subset of $\{1, 2, \ldots, n\}$ such that $\Phi$, $\Pi$, $(\widetilde{\Pi}, K)$ and $\xi$ are independent. Namely, let $\mathcal{L}(K \,|\, \widetilde{\Pi})$ be the uniform distribution on the set of all $L \subset \{1, 2, \ldots, n\}$ such that

$$\{\widetilde{\Pi}(i) : 1 \le i \le n\} = \{\widetilde{\Pi}(i) : i \in L\} \quad \text{and} \quad \#\{\widetilde{\Pi}(i) : 1 \le i \le n\} = \#L.$$

Symmetry considerations show that

$$\begin{aligned}
\mathbb{P}\{i \in K\} &= \mathbb{E}\,\#\{\widetilde{\Pi}(j) : 1 \le j \le n\}/n \\
&= 1 - \mathbb{P}\Big\{1 \notin \{\widetilde{\Pi}(j) : 1 \le j \le n\}\Big\} \\
&= 1 - (1 - 1/n)^n.
\end{aligned}$$

Hence

$$S^o = \lambda\,\mathbb{E}(S_K^o \,|\, \Phi, \Pi, \xi),$$

where $\lambda := (1 - (1 - 1/n)^n)^{-1}$ and $S_K^o := \sum_{i \in K} \xi_i \phi_{i\Pi(i)}$. Consequently,

$$\begin{aligned}
\mathbb{E}\,\Psi(S) &\le \mathbb{E}\,\Psi(\kappa S^o) \quad \text{[Theorem ??]} \\
&= \mathbb{E}\,\Psi\Big(\kappa\lambda\,\mathbb{E}(S_K^o \,|\, \Phi, \Pi, \xi)\Big) \\
&\le \mathbb{E}\,\Psi(\kappa\lambda S_K^o) \quad \text{[Lemma ??]}.
\end{aligned}$$

But one easily verifies that $(\Pi(i))_{i\in K} =_{\mathcal{L}} (\widetilde{\Pi}(i))_{i\in K}$, whence

$$S^o_K \ =_{\mathcal{L}} \ \widetilde{S}^o_K \ := \ \sum_{i\in K} \xi_i \phi_{i\widetilde{\Pi}(i)}.$$

Moreover,

$$\widetilde{S}^o_K \ = \ \mathbb{E}\Big(\widetilde{S}^o \,\Big|\, \Phi, \widetilde{\Pi}, K, (\xi_i)_{i\in K}\Big) \quad \text{with} \quad \widetilde{S}^o \ := \ \sum_{i=1}^n \xi_i \phi_{i\widetilde{\Pi}(i)}.$$

Therefore another application of Lemma **??**, this time to the conditional distribution given $(\Phi, \widetilde{\Pi}, K, (\xi_i)_{i\in K})$, yields

$$\mathbb{E}\,\Psi(S) \ \le \ \mathbb{E}\,\Psi(\kappa\lambda\widetilde{S}^o).$$

The final step is standard: If we define $D(s) := \{i : \xi_i = s\}$ and $\widetilde{S}_L := \sum_{i\in L} \phi_{i\widetilde{\Pi}(i)}$, then

$$\begin{aligned}
\mathbb{E}\,\Psi(\kappa\lambda\widetilde{S}^o) \ &= \ \mathbb{E}\,\Psi(\kappa\lambda(\widetilde{S}_{D(1)} - \widetilde{S}_{D(-1)})) \\
&\le \ \mathbb{E}\Big(2^{-1}\Psi(2\kappa\lambda\widetilde{S}_{D(1)}) + 2^{-1}\Psi(-2\kappa\lambda\widetilde{S}_{D(-1)})\Big) \\
&= \ \mathbb{E}\,\Psi_{\mathrm{s}}(2\kappa\lambda\widetilde{S}_{D(1)}) \\
&= \ \mathbb{E}\,\Psi_{\mathrm{s}}\Big(2\kappa\lambda\,\mathbb{E}(\widetilde{S}\,|\,\xi, \widetilde{S}_{D(1)})\Big) \\
&\le \ \mathbb{E}\,\Psi_{\mathrm{s}}(\gamma\widetilde{S}) \quad [\text{Lemma } \mathbf{??}]. \qquad \Box
\end{aligned}$$

**Acknowledgement.** Many thanks to a referee for constructive comments.

## References


Arcones, M.A. and E. Giné (1993). Limit Theorems for $U$-processes. *Ann. Prob.* **21**, 1494-1542

de la Peña, V.H. (1992). Decoupling and Khintchine's inequalities for $U$-Statistics. *Ann. Probab.* **20**, 1877-1892

Dümbgen, L. (1994). Combinatorial stochastic processes. *Stochastic Process. Appl.* **52**, 75-92

Dümbgen, L. (1998). Application of local rank tests to nonparametric regression. Preprint

Hoeffding, W. (1951). A combinatorial Central Limit Theorem. *Ann. Math. Statist.* **22**, 558-566

Hoeffding, W. (1963). Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.* **58**, 13-30

LeCam, L.M. (1986). *Asymptotic methods in statistical decision theory.* Springer, New York

Pollard, D. (1990). *Empirical Processes: Theory and Applications.* Regionial Conf. Series Prob. Statist. **2**, IMS

van der Vaart, A.W. and J.A. Wellner (1996). *Weak Convergence and Empirical Processes with Applications to Statistics.* Springer, New York