ANNALES DE L'I. H. P., SECTION B

VICTOR H. DE LA PEÑA

# A bound on the moment generating function of a sum of dependent variables with an application to simple random sampling without replacement

*Annales de l'I. H. P., section B*, tome 30, n° 2 (1994), p. 197-211

<http://www.numdam.org/item?id=AIHPB_1994__30_2_197_0>

*Ann. Inst. Henri Poincaré,*

Vol. 30, n° 2, 1994, p. 197-211. *Probabilités et Statistiques*

# A bound on the moment generating function of a sum of dependent variables with an application to simple random sampling without replacement

by

**Victor H. DE LA PEÑA** (1)

Department of Statistics,
Columbia University NY, NY 10027, U.S.A.

ABSTRACT. – In this paper we prove the following inequality: Let $\{x_i\}$ be an arbitrary sequence of random variables. Then there exists a $\sigma$-field $\mathcal{G}$, and a $\mathcal{G}$-*conditionally independent* sequence $\{y_i\}$ *tangent* to $\{x_i\}$ (in particular, $y_i$ has the same distribution as $x_i$ for all $i$) such that for all $\lambda$,

$$(*) \qquad \mathrm{E}\exp\left(\lambda\sum_{i=1}^{n} x_i\right)\leqq\sqrt{\mathrm{E}\exp\left(2\lambda\sum_{i=1}^{n} y_i\right)}.$$

As application of the above we show that the tail behaviour of $\sum_{i=1}^{n} y_i$ controls the tail behaviour of $\sum_{i=1}^{n} x_i$ whenever the *conditionally independent* sequence is sub-Gaussian. Furthermore, by considering $\lambda$ as a random variable independent of $\{x_i\}$, $\{y_i\}$ we show that $(*)$ implies several new decoupling inequalities including a new result for $l_2$ valued random variables. Making the theory of *decoupling* available to the mainstream of statistics we give several examples where the *conditionally independent*

*A.M.S. Classification:* 60 E 15, 60G 40, 60 G 50, 60 G 42.

(1) Supported in part by N.S.F. Grants DMS-9108006, 9002252.

sequence can be identified, and introduce *conditionally independent* sampling as an alternative sampling scheme to sampling without replacement from finite populations.

*Key words :* Decoupling, tangent sequences, moment generating function, Laplace transform, sampling without replacement, martingales.

RÉSUMÉ. – Nous établissons l'inégalité suivante : Soit $\{x_i\}$ une suite arbitraire de variables aléatoires. Alors, il existe une tribu $\mathcal{G}$ et une suite de variables aléatoires $\{y_i\}$ $\mathcal{G}$-conditionnellement indépendante, tangente à $\{x_i\}$ (en particulier, $y_i$ a la même loi que $x_i$ pour tout $i$) telle que pour tout $\lambda$ :

$$(*)\qquad \mathrm{E}\exp\left(\lambda\sum_{i=1}^{n} x_i\right)\leqq\sqrt{\mathrm{E}\exp\left(2\lambda\sum_{i=1}^{n} y_i\right)}.$$

Comme application de (*), nous montrons que dès que $\{y_i\}$ est « sous-gaussienne », le comportement de la queue de $\sum_{i=1}^{n} y_i$ contrôle celui de $\sum_{i=1}^{n} x_i$. Par ailleurs, nous obtenons grace à (*) plusieurs nouvelles inégalités de découplage et notamment un nouveau résultat pour des variables aléatoires aux valeurs de $l_2$. Afin de rendre la théorie du découplage utile pour les statistiques, nous donnons plusieurs exemples où la suite conditionnellement indépendante peut être identifiée et nous introduisons les échantillons conditionnellement indépendants comme une alternative possible des échantillons sans remise à partir d'une population finie.

## 0. INTRODUCTION

In this paper we introduce several inequalities comparing functionals of arbitrary random variables to functionals of conditionally independent random variables. In particular, we present a new inequality for the moment generating function of an arbitrary sum of real dependent variables. We use this inequality to make a study on the tail behaviour of sums of arbitrary random variables. We also show how to turn the above inequality into a machine for generating decoupling inequalities (to be defined later) by the use of the Laplace transform and other transforms.

As a statistical application, we introduce a new sampling scheme which we call conditionally independent sampling as an alternate sampling scheme to simple random sampling without replacement.

A *decoupling inequality* typically aims at making comparisons between two processes one of which may have a simpler dependency structure. This type of inequality has proven to be a useful tool in solving important problems from various areas of probability and statistics. An early result of D. L. Burkholder and T. R. McConnell (*see* Lemma 1 of Burkholder [2]) dealing with problems involving martingale transforms of a Rademacher sequence has been important in the study of integral operators on Lebesgue-Bochner spaces. A second early paper of Jacod [11] dealt with the use of *decoupling* in extending the theory of semimartingales. Of special importance is the work on multinilear forms of independent random variables of McConnell and Taqqu [21] which inspired much of the initial work on *decoupling inequalities*. In the area of stochastic integration the results in McConnell [23], Krakowiak and Szulga [16], Kwapień and Woyczyński [18] and Kallenberg and Szulga [13] have had an important influence. The theory of sequential analysis has benefited from the works of Klass [14], [15] and de la Peña [6]. In those papers, they provide generalizations of Wald's lemma for randomly stopped sums of independent variables. The general theory of martingales was advanced as a result of the development of best possible $L_p$ bounds for martingales as given in de la Peña [4]. In the area of theoretical statistics, de la Peña [5] contains a decoupling result for U-statistics (more generally U-processes) which as an application gives a new *symmetrization* inequality for U-processes. This *symmetrization* inequality was used by Arcones and Giné [1] as a key tool in developing the general theory of U-processes. They used it (among other things) to obtain exponential and Bernstein's type inequalities for U-statistics.

The results on decoupling inequalities in Zinn [24] and Hitczenko [7] provided a solid foundation for the theory of *decoupling* by introducing fairly general results that are readily applicable. Their results consist of *decoupling* inequalities showing that the $L_p$ norms of two *tangent* processes are equivalent in several instances (*see* Definition 1 below ). The theory was further advanced by the tail probability comparison results for *tangent* sequences of Kwapień and Woyczyński [18], [19], later generalized by Kwapień and Woyczyński [20] and de la Peña [6]. These results allow for tail comparisons such as Lenglart-type and Good-Lambda inequalities. Kwapień [17] and Kwapień and Woyczyński [20] include several general results including strict decoupling inequalities for the tail probabilities of quadratic forms of independent symmetric variables and certain multilinear forms.

It is to be remarked that the known results (previous to our own) did not include general exponential decoupling inequalities, which as in the case of [1] have been shown to be useful tools in applications. As a continuation of our previous work this paper presents new exponential *decoupling inequalities* for sums of arbitrarily dependent random variables. It is important to observe that "The Principle of Conditioning" introduced in Jakubowski [12] is closely related to decoupling ideas and has inspired our present work on decoupling inequalities.

We start by introducing some relevant examples in particular an example dealing with simple random sampling without replacement.

## 1. DEFINITIONS AND EXAMPLES

DEFINITION 1. — *Let* $\{x_i\}$, $\{y_i\}$ *be two sequences of random variables adapted to an increasing sequence of* $\sigma$*-fields* $\{\mathscr{F}_i\}$. *Then* $\{x_i\}$ *is said to be* tangent *to* $\{y_i\}$ *if for all i*, $\mathscr{L}(x_i|\mathscr{F}_{i-1})=\mathscr{L}(y_i|\mathscr{F}_{i-1})$, *were* $\mathscr{L}(x_i)$ *stands for the probability law of* $x_i$.

DEFINITION 2. — *A sequence of random variables* $\{x_i\}$ *is said to be* conditionally symmetric *if* $x_i$ *is* tangent *to* $-x_i$ *for all i.*

DEFINITION 3. — *A sequence* $\{y_i\}$ *of random variables adapted to an increasing sequence of* $\sigma$*-field* $\{\mathscr{F}_i\}$ *contained in* $\mathscr{F}$ *is said to be* conditionally independent (CI) *is there exists a* $\sigma$*-algebra* $\mathscr{G}$ *contained in* $\mathscr{F}$ *such that* $\{y_i\}$ *is conditionally independent given* $\mathscr{G}$ *and* $\mathscr{L}(y_i|\mathscr{F}_{i-1})=\mathscr{L}(y_i|\mathscr{G})$.

DEFINITION 4. — *Let* $\{x_i\}$ *be an arbitrary sequence of random variables, then a* conditionally independent *sequence* $\{y_i\}$ *which is also* tangent *to* $\{x_i\}$ *will be called a* decoupled *version of* $\{x_i\}$.

A key result in the area of *decoupling inequalities*, which will be used extensively in this paper was introduced in Kwapień and Woyczyński [18] (*see* also Jakubowski [12]). We state it as a proposition.

PROPOSITION 1. — *For any sequence of random variables* $\{x_i\}$ *one can find a* decoupled sequence $\{y_i\}$ (*on a possibly enlarged probability space*) *which is* tangent *to the original sequence and in addition* conditionally independent *given a* master $\sigma$ field $\mathscr{G}$. *Frequently* $\mathscr{G}=\sigma(\{x_i\})$. *Details of this result may also be found in Hitczenko* [8].

One approach for constructing a *conditionally independent* sequence is to proceed sequencially. Assuming that at the *j*-th stage in the sampling process producing $\{x_i\}$ it is possible to obtain a conditionally independent copy $y_j$ of $x_j$ given $x_1, \ldots, x_{j-1}$. The following diagram illustrates the idea.

Let $y_1$ be an independent copy of $x_1$. For each $j \geqq 2$,

$$\begin{array}{cccccccc} x_1 \rightarrow & x_2 \rightarrow & x_3 \rightarrow & x_4 \rightarrow & \ldots \rightarrow & x_{j-1} \rightarrow & x_j \\ \searrow & \searrow & \searrow & \searrow & & \searrow & \searrow \\ y_1 & y_2 & y_3 & y_4 & \ldots & y_{j-1} & y_j. \end{array}$$

At the $j$-th stage, given $\{x_1, \ldots, x_{j-1}\}$, $x_j$ and $y_j$ are drawn from independent repetitions of the same random mechanism.

In what follows we present three examples identifying a *decoupled* version for several discrete time processes. In particular Example 1 shows how to obtain *conditionally independent* processes *tangent* to randomly stopped sums, U-statistics, quadratic forms and martingale transforms of independent variables.

*Example* 1. – Let $\{z_i\}$ be a sequence of independent variables. Let $\{\tilde{z}_i\}$ be an independent copy of $\{z_i\}$. Take measurable functions $\{f_j\}$ from $R^j \rightarrow R$. Then, $f_j(z_1, \ldots, z_{j-1}; z_j)$ is *tangent* to $f_j(z_1, \ldots, z_{j-1}; \tilde{z}_j)$ with $\mathcal{F}_n = \sigma(z_1, \ldots, z_n; \tilde{z}_1, \ldots, \tilde{z}_n)$. The results to be introduced in this paper, can be used to compare $\sum_{j=1}^{n} f_j(z_1, \ldots, z_{j-1}; z_j)$ to $\sum_{j=1}^{n} f_j(z_1, \ldots, z_{j-1}; \tilde{z}_j)$ which is a sum of *conditionally independent* variables given $\mathcal{G} = \sigma(z_1, z_2, \ldots)$. In particular, the case of U-statistics with kernel $f$ can be obtained by taking $f_j(z_1, \ldots, z_{j-1}; z_j) = \sum_{i=1}^{j-1} f(z_i, z_j)$. By letting $f_j(z_1, \ldots, z_{j-1}; z_j) = \sum_{i=1}^{j-1} a_{ij} z_i z_j$ for some constants $\{a_{ij}\}$ one gets quadratic forms. In this case it is possible to compare the quadratic form $\sum_{j=2}^{n} \left( \sum_{i=1}^{j-1} a_{ij} zi \right) z_j$ to $\sum_{j=2}^{n} \left( \sum_{i=1}^{j-1} a_{ij} z_i \right) \tilde{z}_j$. In the case of martingale transforms of independent variables take

$$f_j(z_1, \ldots, z_{j-1}; z_j) = v_{j-1}(z_1, \ldots, z_{j-1}) z_j,$$

for mean zero $z_j's$ and measurable $v_j's$. We can also deal with randomly stopped sums of independent variables by letting $f_j(z_1, \ldots, z_{j-1}; z_j) = z_j 1(T \geqq j)$ for T a stopping time on the sequence $\{z_i\}$. Then $\sum_{i=1}^{T} z_i = \sum_{i=1}^{\infty} z_i 1(T \geqq i)$, and by using *decoupling* one could compare $\sum_{i=1}^{T} z_i$ to $\sum_{i=1}^{T} \tilde{z}_i$.

The next example shows the importance of *decoupling* in the study of simple random sampling without replacement from finite populations. To do this *conditionally independent* sampling is introduced and compared to simple random sampling without replacement. For convenience, the notation for random variables and their associated sample values is the same.

*Example* 2. – Consider drawing a sample of size $n$ from a box with N balls $\{b_1, \ldots, b_N\}$, $n \leqq N < \infty$. The sequence $\{x_i\}_{i=1}^n$ will represent a sample without replacement $\{y_i\}_{i=1}^n$ a *conditionally independent* sample and $\{z_i\}_{i=1}^n$ a sample with replacement. In obtaining a *conditionally independent* sequence proceed as follows. At the $i$-th stage of a simple random sample without replacement both $x_i$ and $y_i$ are obtained by sampling uniformly from $\{b_1, \ldots, b_N\} \setminus \{x_1, \ldots, x_{i-1}\}$. This may be attained by randomization if N is known or by selectively returning balls to the box after sampling if N is unknown. More precisely, we can obtain both sequences in the following way: at the $i$-th stage at first we draw $y_i$, return the ball, then draw $x_i$ and put the ball aside. It is easy to see that the above procedure will make $\{y_i\}_{i=1}^n$ *tangent* to $\{x_i\}_{i=1}^n$ with $\mathcal{F}_n = \sigma(x_1, \ldots, x_n; y_1, \ldots, y_n)$ (*see* Definition 1). Moreover, $\{y_i\}_{i=1}^n$ is *conditionally independent* given $\mathcal{G} = \sigma(x_1, \ldots, x_N)$.

The results of this paper (Corollary 1) could be compared with those of Hoeffding [10]. In that paper it is proved that if $\{x_i\}$ is obtained from simple random sampling without replacement and $\{z_i\}$ from simple random sampling with replacement from the same finite population, then for any continuous and convex function $\Psi$, the inequality

$$E\Psi\left(\sum_{i=1}^n x_i\right) \leqq E\Psi\left(\sum_{i=1}^n z_i\right)$$

holds. This result has been extensively applied in making inferences about simple random sampling without replacement by using simple random sampling with replacement.

*Example* 3. – In this example we provide a *conditionally independent* sequence to an auto regressive model. Let $x_0 = 0$ and for all $i \geqq 1$, $x_i = \theta x_{i-1} + \varepsilon_i$ where $|\theta| < 1$ and $\varepsilon_i$ is a sequence of *i.i.d.*, mean zero random variables. Then, *a conditionally independent* sequence *tangent* to $\{x_i\}$ is $\{y_i\}$ where for each $i$, $y_i = \theta x_{i-1} + \tilde{\varepsilon}_i$ with $\{\tilde{\varepsilon}_i\}$ an independent copy of $\{\varepsilon_i\}$.

It follows from the theory of *decoupling* that *conditionally independent* sequences share several properties with the originating sequence. In particular, for any two *tangent* sequences $\{x_i\}$ and $\{y_i\}$ of non-negative or *conditionally symmetric* real random variables, there exists constants $0 < A_p$, $B_p < \infty$ (depending on $p$ only) such that

$$A_p E \max_{m \leqq n} |y_1 + \ldots + y_m|^p \leqq E \max_{m \leqq n} |x_1 + \ldots + x_m|^p \leqq B_p E \max_{m \leqq n} |y_1 + \ldots + y_n|^p$$

for all $p > 0$ (*see* Hitczenko [7]).

In Section 2 we introduce a series of new *decoupling* inequalities.

## 2. MAIN INEQUALITIES

The results to be introduced in this section are useful in comparing any two *tangent* sequences when one of them is *conditionally independent*. From Proposition 1, our comparisons serve as bounds for functionals of arbitrary real random variables. The following key result was developed with some input from M. Klass. Its proof involves the use of martingale theory and the properties of *conditionally independent sequences*. For the sake of clarity, we state the result in two parts A and B.

THEOREM 1 A. — *Let* $\{x_i\}$ *be a sequence of positive variables. Then, there exists a* $\sigma$*-field* $\mathcal{G}$ *and a* $\mathcal{G}$-conditionally independent *sequence* $\{y_i\}$, tangent *to* $\{x_i\}$ *such that,*

$$\mathrm{E}\left(\prod_{i=1}^{n} x_i\right)^{1/2} \leqq \left(\mathrm{E}\prod_{i=1}^{n} y_i\right)^{1/2}.$$

THEOREM 1 B. — *Let* $\{x_i\}$ *be a sequence of positive variables. Let* $\mathcal{G}$ *be a* $\sigma$*-field. Then, for any* $\mathcal{G}$-conditionally independent *sequence* $\{y_i\}$, tangent *to* $\{x_i\}$ *one has,*

$$\mathrm{E}\left(\prod_{i=1}^{n} x_i\right)^{1/2} \leqq \left(\mathrm{E}\prod_{i=1}^{n} y_i\right)^{1/2}.$$

[Recall that $\mathcal{G}$ may be taken to equal $\sigma(\{x_i\})$.]

*Proof.* — From Proposition 1, it follows that for any sequence $\{x_i\}$ one can find a tangent sequence $\{y_i\}$ where $\{y_i\}$ is *conditionally independent* given a master $\sigma$-field $\mathcal{G}$. Let $\mathcal{F}_i$ be the $\sigma$-field generated by $\{x_1, \ldots, x_i, y_1, \ldots, y_i\}$ it is easy to see that,

$$\mathrm{E}\,\frac{\prod_{i=1}^{n} x_i}{\prod_{i=1}^{n} \mathrm{E}(x_i \mid \mathcal{F}_{i-1})} = 1. \tag{2.1}$$

Also, since $\{y_i\}$ is tangent to $\{x_i\}$ and *conditionally independent* given $\mathcal{G}$,

$$\mathrm{E}(x_i \mid \mathcal{F}_{i-1}) = \mathrm{E}(y_i \mid \mathcal{F}_{i-1}) = \mathrm{E}(y_i \mid \mathcal{G}). \tag{2.2}$$

Hence,

$$\begin{aligned}
&\mathrm{E}\left(\prod_{i=1}^{n} x_i\right)^{1/2}\\
&=\mathrm{E}\left(\prod_{i=1}^{n} x_i\right)^{1/2} \times \frac{\left(\prod_{i=1}^{n} \mathrm{E}(x_i \mid \mathcal{F}_{i-1})\right)^{1/2}}{\left(\prod_{i=1}^{n} \mathrm{E}(x_i \mid \mathcal{F}_{i-1})\right)^{1/2}}\\
&=\mathrm{E}\left(\frac{\prod_{i=1}^{n} x_i}{\prod_{i=1}^{n} \mathrm{E}(x_i \mid \mathcal{F}_{i-1})}\right)^{1/2} \times \left(\prod_{i=1}^{n} \mathrm{E}(x_i \mid \mathcal{F}_{i-1})\right)^{1/2}\\
&\leqq \left(\mathrm{E}\,\frac{\prod_{i=1}^{n} x_i}{\prod_{i=1}^{n} \mathrm{E}(x_i \mid \mathcal{F}_{i-1})}\right)^{1/2}\\
&\quad \times \left(\mathrm{E}\prod_{i=1}^{n} \mathrm{E}(x_i \mid \mathcal{F}_{i-1})\right)^{1/2} \quad \textit{(by Hölder's Inequality)}\\
&=\left(\mathrm{E}\prod_{i=1}^{n} \mathrm{E}(x_i \mid \mathcal{F}_{i-1})\right)^{1/2} \quad (\textit{from } (2.1))\\
&=\left(\mathrm{E}\prod_{i=1}^{n} \mathrm{E}(y_i \mid \mathcal{G})\right)^{1/2} \quad (\textit{from } (2.2))\\
&=\left(\mathrm{E}\left(\mathrm{E}\left(\prod_{i=1}^{n} y_i \mid \mathcal{G}\right)\right)\right)^{1/2}\\
&\quad (\textit{since } \{y_i\} \textit{ is } \mathcal{G} \textit{ conditionally independant})\\
&=\left(\mathrm{E}\prod_{i=1}^{n} y_i\right)^{1/2}.
\end{aligned}$$

The above result is sharp. To see this, take $x$, $\tilde{x}$ to be *i.i.d.* with $\mathrm{P}(x=1)=\mathrm{P}(x=0)=1/2$. Then, one can see that for $x_1=x$, $x_2=x$ and $y_1=\tilde{x}_1$, $y_2=x$ one has $\mathrm{E}\,x=\mathrm{E}\,x_1 x_2 \leqq \sqrt{\mathrm{E}\,(y_1 y_2)^2}=\mathrm{E}\,x$ by independence. This example is due to P. Hitczenko [9].

In the sequel, most of the results could be stated in two parts. For conciseness we state them in the form of Theorem 1 A.

Among the corollaries to Theorem 1, we start with the following exponential inequality comparing the moment generating function of an arbitrary sequence to that of a *conditionally independent* sequence.

COROLLARY 1. – *Let* $\{x_i\}$ *be an arbitrary sequence of random variables. Then, there exists a* $\sigma$*-field* $\mathcal{G}$ *and a* $\mathcal{G}$-conditionally independent *sequence* $\{y_i\}$, tangent *to* $\{x_i\}$ *such that, for all finite* $\lambda$,

$$\mathrm{E}\exp\left(\lambda\sum_{i=1}^{n} x_i\right)\leqq\sqrt{\mathrm{E}\exp\left(2\lambda\sum_{i=1}^{n} y_i\right)}. \tag{2.3}$$

*Note that if the* $y_i^{\prime s}$ *are mean zero, the* $\sqrt{.}$ *symbol may be removed.*

COROLLARY 2. – *Under the assumptions of Corollary* 1,

$$\mathrm{E}\exp\left(\lambda\left|\sum_{i=1}^{n} x_i\right|\right)\leqq 2\sqrt{\mathrm{E}\exp\left(2|\lambda|\left|\sum_{i=1}^{n} y_i\right|\right)}. \tag{2.4}$$

Tail probability comparisons of minimums are possible as shown next.

COROLLARY 3. – *Let* $\{\mathrm{D}_i\}$ *be a sequence of arbitrary sets adapted to an increasing sequence of* $\sigma$*-fields,* $\mathcal{F}_i$. *Then, there exists a sequence* $\{\mathrm{E}_i\}$ *of sets tangent to* $\{\mathrm{D}_i\}$ *moreover,* $\{\mathrm{E}_i\}$ *is* $\mathcal{G}$-conditionally independent *given a* $\sigma$*-field* $\mathcal{G}$ *and the following inequality holds,*

$$\mathrm{P}\left(\bigcap_{i=1}^{n}\mathrm{D}_i\right)\leqq\mathrm{P}\left(\bigcap_{i=1}^{n}\mathrm{E}_i\right)^{1/2}.$$

By letting $\mathrm{D}_i=\{x_i>t\}$ and $\mathrm{E}_i=\{y_i>t\}$ above, we get

$$\mathrm{P}(\min_{1\leqq i\leqq n} x_i>t)\leqq\mathrm{P}(\min_{1\leqq i\leqq n} y_i>t)^{1/2}, \tag{2.5}$$

for all $t$.

The previous result provides a complement to the following result from Kwapień and Woyczyński [18] and Hitczenko [7].

LEMMA 1. – *Let* $\{u_n\}$, $\{v_n\}$ *be tangent sequences, adapted to an increasing sequence of* $\sigma$*-fields* $\mathcal{F}_n$. *Then,*

$$\mathrm{P}(\sup_{j\leqq\infty}|u_j|>t)\leqq 2\,\mathrm{P}(\sup_{j\leqq\infty}|v_j|>t). \tag{2.6}$$

For convenience to the reader we prove it next.

*Proof.* – Let $\sigma=\inf k:|v_k|>t$, $u^*=\sup_{j\leqq\infty}|u_j|$, $v^*=\sup_{j\leqq\infty}|v_j|$. Then,

$$\mathrm{P}(u^*>t)\leqq\mathrm{E}\sum_{i=1}^{\sigma}1(|u_i|>t)+\mathrm{P}(\sigma<\infty)$$

$$=\mathrm{E}\sum_{i=1}^{\sigma}1(|v_i|>t)+\mathrm{P}(\sigma<\infty)=2\,\mathrm{P}(v^*>t).$$

In fact, one can use Corollary 2 to improve the above result when the sequence $\{u_i\}$ is *conditionally independent* by using the inequality

$1-(1-x)^{1/2} \geqq x/2$. Corollary 4 gives a *symmetrization* inequality for martingales which is related to the one used by Arcones and Giné [1].

COROLLARY 4. — *Let* $\{x_i\}$ *be a mean-zero martingale difference sequence. Then there exists a* $\sigma$*-field* $\mathcal{G}$ *and a* $\mathcal{G}$-conditionally independent *sequence* $\{y_i\}$, tangent *to* $\{x_i\}$ *such that for all finite* $\lambda$,

$$\mathrm{E} \exp\left(\lambda \sum_{i=1}^{n} x_i\right) \leqq \sqrt{\mathrm{E} \exp\left(4\lambda \sum_{i=1}^{n} y_i r_i\right)}, \tag{2.7}$$

*for a sequence* $\{r_i\}$ *of independent variables with*

$$\mathrm{P}(r_i=1)=\mathrm{P}(r_i=-1)=1/2,$$

*independent of* $\{x_i\}$ *and* $\{y_i\}$. *The following corollary provides a new decoupling inequality for* U*-statistics and multilinear forms* (see *Example* 1). *It can be obtained by a repeated application of Corollary* 1.

COROLLARY 5. — *Let* $\{z_i^1\}$, $\{z_i^2\}$, . . ., $\{z_i^k\}$ *independent copies of a sequence of independent random variables* $\{z_i\}$. *For* $k \geqq 2$ *let* $f_{i_1, i_2, \ldots, i_k}$ *be an array of functions from* $\mathrm{R}^k$ *into* R. *Then for every finite* $t$,

$$\mathrm{E} \exp\left\{t \sum f_{i_1, i_2, \ldots, i_k}(z_{i_1}, z_{i_2}, \ldots, z_{i_k})\right\}$$
$$\leqq \left\{\mathrm{E} \exp\left\{2^{k-1} t \sum f_{i_1, i_2, \ldots, i_k}(z_{i_1}^1, z_{i_2}^2, \ldots, z_{i_k}^k)\right\}\right\}^{1/2^{k-1}},$$

where $\sum_{1 \leqq i_1 < i_2 < \ldots < i_k \leqq n} = \sum$.

In Theorem 2 we show how the tail behaviour of the *conditionally independent* sequence (CI) controls the tail behaviour of the original sequence whenever the CI sequence is sub-Gaussian.

THEOREM 2. — *Let* $\{x_i\}$ *be an arbitrary sequence of random variables. Assume that for a* $\sigma$*-field* $\mathcal{G}$, $\{y_i\}$ *is a* $\mathcal{G}$-conditionally independent *sequence* tangent *to* $\{x_i\}$. *Let* Z *be a normal random variable with mean zero and variance* 1. *Then, the inequality*

$$\mathrm{P}\left(\left|\sum_{i=1}^{n} y_i\right| \geqq x\right) \leqq \mathrm{AP}(|\mathrm{Z}| \geqq x)$$

*for some universal constant* $0<\mathrm{A}<\infty$, *and all* $x>0$, *implies*

$$\mathrm{P}\left(\left|\sum_{i=1}^{n} x_i\right| \geqq x\right) \leqq 2\sqrt{\mathrm{A}}\,\mathrm{P}(\sqrt{2}|\mathrm{Z}| \geqq x),$$

*for all* $x \geqq 1$.

*Proof.* — It is well known (*see* Chow and Teicher [3]) that the tail probability of the Gaussian random variable satisfies the following,

$$\frac{2x}{1+x^2}\exp\left(-\frac{x^2}{2}\right)<P(|Z|\geqq x)<\frac{2}{x}\exp\left(\frac{-x^2}{2}\right),$$

for all $x\geqq 0$. Furthermore, the condition of the tail behaviour of $\sum_{i=1}^{n} y_i$ implies that for all $t>0$,

$$E\exp\left(t\left|\sum_{i=1}^{n} y_i\right|\right)\leqq A\exp\left(\frac{t^2}{2}\right).$$

Hence,

$$\begin{aligned}
P\left(\left|\sum_{i=1}^{n} x_i\right|\geqq x\right) &\leqq \inf_{t>0}\left\{\exp(-tx).E\exp\left(t\left|\sum_{i=1}^{n} x_i\right|\right)\right\}\\
&\leqq \inf_{t>0}\left\{\exp(-tx).2\sqrt{E\exp\left(2t\left|\sum_{i=1}^{n} y_i\right|\right)}\right\} \text{ (from Corollary 2)}\\
&=2\sqrt{\inf_{t>0}\left\{\exp(-tx).E\exp\left(t\left|\sum_{i=1}^{n} y_i\right|\right)\right\}}\\
&\leqq 2\sqrt{A\inf_{t>0}\exp\left(-tx+\frac{t^2}{2}\right)}\\
&=2\sqrt{A}\exp\left(-\frac{x^2}{4}\right)\\
&\leqq 2\sqrt{A}\,\frac{x/\sqrt{8}}{1+(x^2/8)}\exp\left(-\frac{x^2}{8}\right)\\
&\leqq 2\sqrt{A}\,P\left(|Z|\geqq\frac{x}{\sqrt{2}}\right).
\end{aligned}$$

*Note.* — To avoid unnecessary repetitions, in Section 3 we assume that all the objects in question are both measurable and integrable.

## 3. A GENERATING FUNCTION OF DECOUPLING INEQUALITIES

In this section we show how to exploit the properties of Corollary 1 to produce a wealth of new results. The idea consists of integrating both sides of (2.3) against non-negative functions $g(\lambda)$ with total Lebesgue integral equal 1 (densities) followed by a use of Fubini's theorem and Jensen's inequality (taking into account the concavity of the $\sqrt{\cdot}$ function) to obtain new inequalities comparing, for example, the Laplace transform of $\sum_{i=1}^{n} x_i$ to a re-scaled version of the Laplace transform of $\sum_{i=1}^{n} y_i$. A related approach consists of considering $\lambda$ as a random variable independent of both $\{x_i\}$ and $\{y_i\}$. Conditioning on these two sequences in (2.3) and varying the distribution of $\lambda$ we obtain a multitude of new results. We have just proved the following.

THEOREM 3. – *Let* $\{x_i\}$ *be an arbitrary sequence of random variables. Then, there exists a* $\sigma$*-field* $\mathcal{G}$ *and a* $\mathcal{G}$-conditionally independent *sequence* $\{y_i\}$, tangent *to* $\{x_i\}$ *such that for all densities* $g(\lambda)$,

$$E\int_{-\infty}^{+\infty} \exp\left(\lambda \sum_{i=1}^{n} x_i\right) g(\lambda)\, d\lambda \leqq \sqrt{E\int_{-\infty}^{+\infty} \exp\left(2\lambda \sum_{i=1}^{n} y_i\right) g(\lambda)\, d\lambda}. \quad (3.1)$$

THEOREM 4. – *Let* $\{x_i\}$ *be an arbitrary sequence of random variables. Then, there exists a* $\sigma$*-field* $\mathcal{G}$ *and a* $\mathcal{G}$-conditionally independent *sequence* $\{y_i\}$, tangent *to* $\{x_i\}$ *such that, for all random variables t independent of both* $\{x_i\}$ and $\{y_i\}$ *we get,*

$$E\exp\left(t \sum_{i=1}^{n} x_i\right) \leqq \sqrt{E\exp\left(2t \sum_{i=1}^{n} y_i\right)}. \quad (3.2)$$

The following result provides an interesting application of Theorem 4.

THEOREM 5. – *Let* $\{X_i\}$ *be a sequence of random vectors in* $l_2$, *with* $\|\cdot\|$ *representing the Euclidian norm in* $l_2$. *Then, there exists a* $\sigma$*-field and a* $\mathcal{G}$-conditionally independent *sequence* $\{Y_i\}$ tangent *to* $\{X_i\}$ *for which,*

$$E\exp\left\{\left\|\sum_{i=1}^{n} X_i\right\|^2\right\} \leqq \sqrt{E\exp\left\{4\left\|\sum_{i=1}^{n} Y_i\right\|^2\right\}}. \quad (3.3)$$

*Proof.* – We prove the case of $R^k$ valued random variables. The general case follows by an easy limiting argument. Let $t_1, \ldots, t_k$ be an *i.i.d.* sequence of normal random variables with mean zero and variance one, independent of $\{X_i\}$ and $\{Y_i\}$. From the assumptions, $X_i = (x_i^1, \ldots, x_i^k)$ for real $x'^s$. Since $X_i$ is *tangent to* $Y_i$, $\sum_{j=1}^{k} t_j x_i^j$ is also *tangent* to $\sum_{j=1}^{k} t_j y_i^j$. Therefore, we can apply Theorem 4 to get,

$$E\exp\left\{\sum_{i=1}^{n}\sum_{j=1}^{k} t_j x_i^j\right\} \leqq \sqrt{E\exp\left\{2\sum_{i=1}^{n}\sum_{j=1}^{k} t_j y_i^j\right\}}.$$

Conditioning on $\{x_i^j\}$, $\{y_i^j\}$ and $t_1, \ldots, t_{k-1}$ we get

$$\mathrm{E}\exp\left\{\sum_{i=1}^{n}\sum_{j=1}^{k-1} t_j x_i^j + \frac{\left(\sum_{i=1}^{n} x_i^k\right)^2}{2}\right\} \leqq \sqrt{\mathrm{E}\exp\left\{2\left(\sum_{i=1}^{n}\sum_{j=1}^{k-1} t_j y_i^j\right) + \frac{\left(2\sum_{i=1}^{n} y_i^k\right)^2}{2}\right\}}.$$

Proceeding in the same fashion for $j=k-1, \ldots, 1$ we get

$$\mathrm{E}\exp\left\{\sum_{j=1}^{k}\frac{\left(\sum_{i=1}^{n} x_i^j\right)^2}{2}\right\} \leqq \sqrt{\mathrm{E}\exp\left\{\sum_{j=1}^{k}\frac{\left(2\sum_{i=1}^{n} y_i^j\right)^2}{2}\right\}},$$

which gives the desired result.

*Remark.* — Taking $t$ to be the product of independent random variables in Theorem 4 would be helpful in obtained new results by integrating separately each random variable.

Several other inequalities can be obtained by following our general approach.

*Example* 4. — By letting $t$ have a Poisson distribution with parameter $\mu>0$ in Theorem 4 we get, for a sequence $\{x_i\}$ and its *decoupled* counterpart $\{y_i\}$,

$$\mathrm{E}\, e^{\mu\left(e^{\sum_{i=1}^{n} x_i}-1\right)} \leqq \sqrt{\mathrm{E}\, e^{\mu\left(e^{2\sum_{i=1}^{n} y_i}-1\right)}}. \tag{3.4}$$

Corollary 2 can be useful when the interest is on results for the absolute value of sums. For example, by letting $t$ have a uniform distribution on $(a, b)$ in a variation of Theorem 4 we get for mean zero $x'^s$.

*Example* 5. — For a sequence of mean zero random variables $\{x_i\}$, let $\{y_i\}$ be its *decoupled* counterpart. Then,

$$\mathrm{E}\,\frac{e^{\left|\sum_{i=1}^{n} x_i\right| b} - e^{\left|\sum_{i=1}^{n} x_i\right| a}}{\left|\sum_{i=1}^{n} x_i\right|} \leqq \mathrm{E}\,\frac{e^{2\left|\sum_{i=1}^{n} y_i\right| b} - e^{2\left|\sum_{i=1}^{n} y_i\right| a}}{\left|\sum_{i=1}^{n} y_i\right|}, \tag{3.5}$$

in the case the variables satisfy $\mathrm{P}\left(\left|\sum_{i=1}^{n} x_i\right|=0\right)=\mathrm{P}\left(\left|\sum_{i=1}^{n} y_i\right|=0\right)=0.$

ACKNOWLEDGEMENTS

This paper was completed while the author was on leave from Columbia University. The author wishes to thank the following individuals and institutions for their support. D. L. Burkholder, P. Deheuvels, J. Jacod, M. J. Klass, H. Levene, V. Perez-Abreu and E. Schuster. The "Centro de Investigaciones en Mathematicas" of Guanajuato México, Departments of Mathematics of The University of Texas at El Paso, Case Western Reserve University, The Mathematical Sciences Research Institute of Berkeley California, The Department of Statistics of UC Berkeley and the Department of Mathematics of the University of Illinois at Urbana-Champaign. The Laboratoire de Probabilités and Laboratoire de Statistique Théorique et Appliquée de Paris-VI. The author benefited from conversations with Lou Gordon, P. Hitczenko, J. Jacod and D. Nolan. Finally, the referee provided helpful remarks and comments.

## REFERENCES

[1] M. ARCONES and E. GINÉ, Limit Theorems for U-processes, 1990, To appear in *Ann. of Probab.*

[2] D. L. BURKHOLDER, A Geometric Condition that Implies the Existence of Certain Singular Integrals of Banach-Space-Valued Functions, *Conference on Harmonic Analysis in Honor of Antoni Zygmund*, Chicago, 1981, edited by William Beckner, Alberto P. Calderón, Robert Fefferman and Peter W. Jones, Wadsworth, Belmont, California, 1983.

[3] Y. S. CHOW and H. TEICHER, *Probability Theory*, Springer-Verlag, New York, 1978.

[4] V. H. DE LA PEÑA, Bounds on the Expectation of Functions of Martingales and Sums of Positive rv's in Terms of Norms of Sums of Independent Random Variables, *Proccedings Amer. Math. Soc.*, Vol. **108**, No. 1, pp. 233-239.

[5] V. H. DE LA PEÑA, Decoupling and Khintchine's Inequalities for U-Statistics, *Ann. Probab.*, Vol. **20**, No. 4, 1992, pp. 1877-1892.

[6] V. H. DE LA PEÑA, Inequalities for Tails of Adapted Processes with an Application to Wald's Lemma, *Journal of Theoretical Probability*, Vol. **6**, No. 2, 1993.

[7] P. HITCZENKO, Comparison of Moments of Tangent Sequences of Random Variables, *Prob. Th. Rel. Fields*, Vol. **78**, No., 2, 1988, pp. 223-230.

[8] P. HITCZENKO, Best Constants in Martingale Version of Rosenthal's Inequality, *Ann. of Probab.*, Vol. **18**, No. 4, 1990, pp. 1656-1668.

[9] P. HITCZENKO, *Personal communication*, 1991.

[10] W. HOEFFDING, Probability Inequalities for Sums of Bounded Random Variables, *Journal of the American Statistical Association*, Vol. **58**, 1963, pp. 13-30.

[11] J. JACOD, Une géneralisation des semimartingales : Les processus admettant un processus à accroissements independants tangent, *Lecture Notes in Math*, No. **1247**, 1984, pp. 479-514.

[12] A. JAKUBOWSKI, Principle of Conditioning in Limit Theorems for Sums of Random Variables, *Ann. Probab.*, Vol. **14**, No 3, pp. 902-915.

[13] O. KALLENBERG and J. SZULGA, Multiple Integration with Respect to Poisson and Lévy Processes, *Probab. Th. Rel. Fields*, Vol. **83**, 1989, pp. 101-134.

[14] M. J. KLASS, A Best Improvement of Wald's Lemma, *Ann. of Probab.*, Vol. **16**, No. 2, 1988, pp. 840-853.

[15] M. J. KLASS, Uniform Lower Bounds for Randomly Stopped Banach Space-Valued Random Sums, *Ann. of Probab.*, Vol. **18**, No. 2, 1990, pp. 790-809.

[16] W. KRAKOWIAK and J. SZULGA, On a $p$-Stable Multiple Integral, II, *Probab. Th. Rel. Fields*, Vol. **78**., No. 3, 1988, pp. 449-453.

[17] S. KWAPIEŃ, Decoupling Inequalities for Polynomial Chaos, *Ann. of Probab.*, Vol. **15**, 1987, pp. 1062-1072.

[18] S. KWAPIEŃ and W. A. WOYCZYŃSKI, *Semimartingale Integrals via Decoupling Inequalities and Tangent processes*, Preprint No. 88-93, Department of Mathematics and Statistics, Case Western Reserve University, 1988.

[19] S. KWAPIEŃ and W. A. WOYCZYŃSKI, Tangent Sequences of Random Variables: Basic Inequalities and their Applications, *Proceedings of Conference on Almost Everywhere Convergence in Probability and Ergodic Theory*, pp. 237-265, Columbus Ohio, June 1988. G. A. Edgar and L. Sucheston, Editors, Academic Press, New York, 1989.

[20] S. KWAPIEŃ and W. A. WOYCZYŃSKI, *Random Series and Stochastic Integrals. Single and Multiple*, Birkhäuser, Boston, 1992.

[21] T. R. MCCONNELL, and M. TAQQU, Decoupling Inequalities for Multilinear Forms in Independent Symmetric Random Variables, *Ann. of Probab.*, Vol. **14**, No. 3, 1986, pp. 943-954.

[22] T. R. MCCONNELL and M. TAQQU, Decoupling Inequalities for Banach-Valued Multilinear Forms in Independent Symmetric Banach-Valued Random Variables, *Probab. Th. Rel. Fields*, Vol. **75**, 1987, pp. 499-507.

[23] T. R. MCCONNELL, Decoupling and Stochastic Integration in UMD Banach Spaces, *Prob. and Math. Stat.*, Vol. **10**, No. 2, 1989, pp. 283-295.

[24] J. ZINN, Comparison of Martingale Difference Sequences. Proceedings of Conference Probability in Banach Spaces V, *Lecture Notes in Mathematics*, No. **1153**, 1975, pp. 453-457, Springer-Verlag, Berlin.

(*Manuscript received May* 27, 1992;
*revised March* 2, 1993.)