

CONCENTRATION INEQUALITIES FOR SEMI-BOUNDED MARTINGALES

YU MIAO¹

Abstract. In this paper, we apply the technique of decoupling to obtain some exponential inequalities for semi-bounded martingale, which extend the results of de la Peña, *Ann. probab.* **27** (1999) 537–564.

Mathematics Subject Classification. 60E15, 60G42.

Received April 21, 2006.

1. INTRODUCTION

In this paper we extend the results of de la Peña [3]. The main method that we use is the theory of decoupling, which has been developed in de la Peña [2] and [3]. Decoupling theory provides a general framework for analyzing problems involving dependent random variables as if they were independent. We will apply the theory of decoupling as in de la Peña [3] and some new inequalities for independent random variables motivated by Maurer [1] to obtaining extensions of exponential inequalities of de la Peña [3].

The paper is organized as follows. In Section 2, by usual methods, we present some new exponential inequalities for random variables and discrete time martingales. A brief introduction to the theory of decoupling is presented in Section 3. In Section 4, we use the theory of decoupling to establish some further results which refine those in Section 2 and extend de la Peña's exponential inequalities from bounded martingale difference sequence to semi-bounded ones.

2. GENERAL EXPONENTIAL INEQUALITIES

We will begin with the following basic inequalities for random variable, whose ideas stem from Maurer [1].

Lemma 2.1. *Let X be a random variable such that $\mathbb{E}X^2 < \infty$ and $\mathbb{P}(X \leq C) = 1$, where $0 < C < \infty$. If $\mathbb{E}X = 0$, then for any $\lambda > 0$,*

$$\mathbb{E}(\exp\{\lambda X\}) \leq \exp\left\{\frac{\lambda^2}{2}(C^2 + \mathbb{E}X^2)\right\}. \quad (2.1)$$

More generally, if $\mathbb{E}X \leq 0$, then for any $\lambda > 0$, we have

$$\mathbb{E}(\exp\{\lambda X\}) \leq \exp\{\lambda \mathbb{E}X + \lambda^2(C^2 + \mathbb{E}X^2)\}. \quad (2.2)$$

Keywords and phrases. Decoupling, exponential inequalities, martingale, conditionally symmetric variables.

¹ College of Mathematics and Information Science, Henan Normal University, 453007 Henan, China and School of Mathematics and Statistics, Wuhan University, 430072 Hubei, China; yumiao728@yahoo.com.cn

Proof. (following Maurer [1]) For $\lambda > 0$, using $e^{-x} \leq 1 - x + \frac{1}{2}x^2$ for any $x \geq 0$, and $e^x \geq 1 + x$ for all $x \in \mathbb{R}$, we have

$$\begin{aligned} \mathbb{E}(\exp\{\lambda X\}) &= \mathbb{E}(\exp\{\lambda C\} \exp\{-\lambda(C - X)\}) \\ &\leq \exp\{\lambda C\} \mathbb{E}\left(1 - \lambda(C - X) + \frac{\lambda^2}{2}(C - X)^2\right) \\ &= \exp\{\lambda C\} \mathbb{E}\left(1 - \lambda(C - X) + \frac{\lambda^2}{2}(C^2 + X^2) - \lambda^2 CX\right) \\ &\leq \exp\left\{\lambda \mathbb{E}X + \frac{\lambda^2}{2}(C^2 + \mathbb{E}X^2) - \lambda^2 C \mathbb{E}X\right\} \end{aligned}$$

which gives (2.1) if $\mathbb{E}X = 0$. If $\mathbb{E}X \leq 0$, (2.2) follows by the inequality: $2C\mathbb{E}X \leq C^2 + (\mathbb{E}X)^2$. \square

Lemma 2.2. *Let X be a random variable such that $\mathbb{E}X^2 < \infty$ and $\mathbb{P}(X \leq C) = 1$, where $0 < C < \infty$. Then for any $\lambda > 0$, we have*

$$\mathbb{E}(\exp\{\lambda X\}) \leq \exp\left\{\lambda \mathbb{E}X + \frac{\lambda^2}{2}e^C \mathbb{E}X^2\right\}. \quad (2.3)$$

In particular if $\mathbb{E}X = 0$, then for any $\lambda > 0$,

$$\mathbb{E}(\exp\{\lambda X\}) \leq \exp\left\{\frac{\lambda^2}{2}e^C \mathbb{E}X^2\right\}. \quad (2.4)$$

Proof. For any $\lambda > 0$, using the inequality $e^x \leq 1 + x + \frac{1}{2}x^2 e^C$ for any $x \leq C$, we have

$$\mathbb{E}(\exp\{\lambda X\}) \leq \mathbb{E}\left(1 + \lambda X + \frac{\lambda^2}{2}e^C X^2\right) \leq \exp\left\{\lambda \mathbb{E}X + \frac{\lambda^2}{2}e^C \mathbb{E}X^2\right\}$$

which is (2.3). That implies (2.4) if $\mathbb{E}X = 0$. \square

The following theorem treats the case of martingale difference sequence.

Theorem 2.3. *Let $\{d_i; \mathcal{F}_i\}$ be a martingale difference sequence with respect to the filtration $(\mathcal{F}_i)_{i \geq 0}$ (i.e., d_i is \mathcal{F}_i -measurable and $\mathbb{E}(d_i | \mathcal{F}_{i-1}) = 0$), such that $\mathbb{E}(d_i^2 | \mathcal{F}_{i-1}) \leq \sigma_i^2$, where $\sigma_i^2 \geq 0$ is constant. Furthermore, assume that $\mathbb{P}(d_i \leq c_i | \mathcal{F}_{i-1}) = 1$, where $c_i > 0$, $\sigma_i^2 > 0$ are constants. Then for any $\lambda > 0$,*

$$\mathbb{E} \exp\left\{\lambda \sum_{i=1}^n d_i\right\} \leq \exp\left\{\frac{\lambda^2}{2} \sum_{i=1}^n (c_i^2 + \sigma_i^2)\right\}. \quad (2.5)$$

In particular, for any $r > 0$,

$$\mathbb{P}\left(\sum_{i=1}^n d_i \geq r\right) \leq \exp\left\{-\frac{r^2}{2 \sum_{i=1}^n (c_i^2 + \sigma_i^2)}\right\}. \quad (2.6)$$

More generally, if $\{d_i; \mathcal{F}_i\}$ is a super-martingale difference sequence (i.e., $\mathbb{E}(d_i | \mathcal{F}_{i-1}) \leq 0$) with $\mathbb{E}(d_i^2 | \mathcal{F}_{i-1}) \leq \sigma_i^2$ and $\mathbb{P}(d_i \leq c_i | \mathcal{F}_{i-1}) = 1$, where $c_i, \sigma_i^2 \geq 0$ are constants. Then, for any $\lambda > 0$,

$$\mathbb{E} \exp\left\{\lambda \sum_{i=1}^n d_i\right\} \leq \exp\left\{\sum_{i=1}^n \lambda^2 (c_i^2 + \sigma_i^2)\right\}. \quad (2.7)$$

In particular, for any $r > 0$,

$$\mathbb{P}\left(\sum_{i=1}^n d_i \geq r\right) \leq \exp\left\{-\frac{r^2}{4\sum_{i=1}^n (c_i^2 + \sigma_i^2)}\right\}. \quad (2.8)$$

Proof. Here we only consider the case of martingale difference sequence, and the other case of super-martingale difference sequence is similar. By Lemma 2.1 or as the proof of Lemma 2.1, for any $\lambda > 0$ and d_i , we have:

$$\mathbb{E}\left(\exp\{\lambda d_i\} \mid \mathcal{F}_{i-1}\right) \leq \exp\left\{\frac{\lambda^2}{2}(b_i^2 + c_i^2)\right\}.$$

Thus

$$\begin{aligned} \mathbb{E} \exp\left\{\lambda \sum_{i=1}^n d_i\right\} &= \mathbb{E}\left[\exp\left\{\lambda \sum_{i=1}^{n-1} d_i\right\} \mathbb{E}\left(\exp\{\lambda d_n\} \mid \mathcal{F}_{n-1}\right)\right] \\ &\leq \exp\left\{\frac{\lambda^2}{2}(b_n^2 + \sigma_n^2)\right\} \mathbb{E} \exp\left\{\lambda \sum_{i=1}^{n-1} d_i\right\}, \end{aligned}$$

which yields to (2.5) by induction. For any $r > 0$, from (2.5) and Chebychev's inequality, we have for any $\lambda \geq 0$,

$$\mathbb{P}\left(\sum_{i=1}^n d_i \geq r\right) \leq \exp\{-\lambda r\} \mathbb{E} \exp\left\{\lambda \sum_{i=1}^n d_i\right\} \leq \exp\left\{-\lambda r + \frac{\lambda^2}{2} \sum_{i=1}^n (c_i^2 + \sigma_i^2)\right\}$$

where (2.6) follows by optimization over $\lambda \geq 0$. □

As the proof of Theorem 2.3, from Lemma 2.2, we derive

Theorem 2.4. *Under the assumptions of Theorem 2.3, let $c = \max_i c_i$ and if $\{d_i; \mathcal{F}_i\}$ is a martingale difference sequence, then for any $\lambda > 0$, we have*

$$\mathbb{E} \exp\left\{\lambda \sum_{i=1}^n d_i\right\} \leq \exp\left\{\frac{\lambda^2}{2} \sum_{i=1}^n e^{c_i} \sigma_i^2\right\} \leq \exp\left\{\frac{\lambda^2}{2} e^c \sum_{i=1}^n \sigma_i^2\right\}. \quad (2.9)$$

3. GENERAL THEORY OF DECOUPLING

In this section, we will recall general theory of decoupling (*cf.* de la Peña [2] and [3]). Decoupling theory provides a general framework for analyzing problems involving dependent random variables as if they were independent. The theory will play a key role in improving the results in Section 2 and extending the works of de la Peña [3].

Definition 3.1. Let $(X_i)_{i \in \mathbb{N}}$, $(Y_i)_{i \in \mathbb{N}}$ be two sequences of random variables adapted to an increasing sequence of σ -fields $\{\mathcal{F}_i\}$. Then $\{X_i\}$ is said to be tangent to $\{Y_i\}$ if for all i , $\mathcal{L}(X_i | \mathcal{F}_{i-1}) = \mathcal{L}(Y_i | \mathcal{F}_{i-1})$, where $\mathcal{L}(X_i | \mathcal{F}_{i-1})$ is the conditional distribution of X_i knowing \mathcal{F}_{i-1} .

Definition 3.2. A sequence of random variables $\{X_i\}$ is said to be conditionally symmetric if $\{X_i\}$ is tangent to $\{-X_i\}$.

Definition 3.3. A sequence $\{Y_i\}$ of random variables adapted to an increasing sequence of sub- σ -field \mathcal{F}_i (of \mathcal{F}) is said to be conditionally independent (CI) if there exists a σ -algebra \mathcal{G} contained in \mathcal{F} such that $\{Y_i\}$ is conditionally independent given \mathcal{G} and $\mathcal{L}(Y_i | \mathcal{F}_{i-1}) = \mathcal{L}(Y_i | \mathcal{G})$.

Definition 3.4. 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 Jakubowski [4] (see also de la Peña [2]). We state it as a proposition.

Proposition 3.1. *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 σ -field \mathcal{G} .*

Remarks 3.2. Frequently, one can take $\mathcal{G} = \sigma(\{X_i, i \in \mathbb{N}\})$ (cf. Kwapien and Woyczyński [5,6]) or de la Peña [3]). The general strategy that we will follow consists of applying this result conditionally on \mathcal{G} and use known inequalities for sums of independent random variables.

Next we will mention the following decoupling inequality (see de la Peña [3]), which is an important tool in proving the results in Section 4.

Theorem 3.3. *Let $\{X_i, i \in \mathbb{N}\}$ be a sequence of nonnegative, non-degenerate random variables. Then there exists a σ -field \mathcal{G} and a \mathcal{G} -conditionally independent sequence $\{Y_i\}$, tangent to $\{X_i\}$, such that for all random variables $g \geq 0$ measurable with respect to \mathcal{G} ,*

$$\mathbb{E} \left(g \prod_{i=1}^n X_i \right)^{1/2} \leq \left(\mathbb{E} g \prod_{i=1}^n Y_i \right)^{1/2}. \quad (3.1)$$

[Recall that \mathcal{G} may be taken to be $\sigma(\{X_i, i \in \mathbb{N}\})$.]

Corollary 3.4. *Let $\{X_i\}$, $\{Y_i\}$ be \mathcal{F}_i -tangent. Assume that $\{Y_i\}$ is decoupled (CI). Let $g \geq 0$ be any random variable measurable with respect to $\sigma(\{X_i\}_{i=1}^\infty)$. Then for all finite t ,*

$$\mathbb{E} \left(g \exp \left\{ t \sum_{i=1}^n X_i \right\} \right) \leq \sqrt{\mathbb{E} \left(g^2 \exp \left\{ 2t \sum_{i=1}^n Y_i \right\} \right)}. \quad (3.2)$$

4. APPLICATIONS OF THE THEORY OF DECOUPLING

In this section, we extend exponential inequalities in de la Peña [3] from bounded martingale differences to semi-bounded ones. As in de la Peña [3], in order to avoid eventual problems with the definition of conditional expectation, in this paper we will use the following notation. Let X be a positive random variable on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let A be an \mathcal{F} -measurable set with $\mathbb{P}(A) > 0$, then $\mathbb{E}(X|A) = \int_A X d\mathbb{P}/\mathbb{P}(A)$.

In what follows, we will use the notation $\mathcal{G} = \sigma\{d_i, i \in \mathbb{N}\}$.

Theorem 4.1. *Let $\{d_i\}$ be a martingale difference sequence (i.e., $\mathbb{E}(d_i|d_1, \dots, d_{i-1}) = 0$), $M_n := \sum_{i=1}^n d_i$, $\langle M \rangle_n = \sum_{i=1}^n \mathbb{E}(d_i^2|d_1, \dots, d_{i-1})$. Assume that $\mathbb{P}(d_i \leq c_i) = 1$ where $0 < c_i < \infty$. Then, for all \mathcal{G} measurable sets A , $x > 0$, $\lambda > 0$,*

$$\mathbb{P} \left(\frac{M_n}{\langle M \rangle_n} \geq x, A \right) \leq \mathbb{E} \left(\exp \left\{ -\lambda \langle M \rangle_n x + \frac{\lambda^2}{2} \left(\sum_{i=1}^n c_i^2 + \langle M \rangle_n \right) \right\} \middle| \frac{M_n}{\langle M \rangle_n} \geq x, A \right) \quad (4.1)$$

and

$$\mathbb{P} \left(\frac{M_n}{\langle M \rangle_n} \geq x, A \right) \leq \sqrt{\mathbb{E} \left(\exp \left\{ -\lambda \langle M \rangle_n x + \frac{\lambda^2}{2} \left(\sum_{i=1}^n c_i^2 + \langle M \rangle_n \right) \right\} 1_A \right)}. \quad (4.2)$$

In particular, under the assumption that $\sum_{i=1}^{\infty} c_i^2 < \infty$, for any $y > 0$, we have

$$\mathbb{P} \left(\frac{M_n}{\langle M \rangle_n} \geq x, \frac{1}{\langle M \rangle_n} \leq y \text{ for some } n \right) \leq \exp \left\{ -\frac{x^2}{2(y^2 \sum_{i=1}^{\infty} c_i^2 + y)} \right\}. \quad (4.3)$$

Proof. Consider the sequence $\{e_i\}$ which is tangent to $\{d_i\}$ and is conditionally independent given \mathcal{G} (maybe defined on some enlarged probability space, see Prop. 3.1 and Rem. 3.2). Applying Chebychev's inequality first, followed by use of Corollary 3.4 with $g = \exp\{-(\lambda/2)\langle M \rangle_n x\} 1_{\{M_n/\langle M \rangle_n \geq x, A\}}$, we obtain

$$\begin{aligned} \mathbb{P}(M_n \geq \langle M \rangle_n x, A) &\leq \mathbb{E} \left(\exp \left\{ \frac{\lambda}{2} (M_n - \langle M \rangle_n x) \right\} 1_{\{M_n \geq \langle M \rangle_n x, A\}} \right) \\ &\leq \sqrt{\mathbb{E} \left(\exp \left\{ \lambda \left(\sum_{i=1}^n e_i - \langle M \rangle_n x \right) \right\} 1_{\{M_n/\langle M \rangle_n \geq x, A\}} \right)} \\ &= \sqrt{\mathbb{E} \left(1_{\{M_n/\langle M \rangle_n \geq x, A\}} \exp \{-\lambda \langle M \rangle_n x\} \mathbb{E} \left(\exp \left\{ \lambda \left(\sum_{i=1}^n e_i \right) \right\} \middle| \mathcal{G} \right) \right)}, \end{aligned}$$

where the last equality follows for the variables outside the conditional expectation are \mathcal{G} -measurable. Since $\{d_i\}$ and $\{e_i\}$ are tangent and $\{e_i\}$ is conditionally independent given \mathcal{G} , the moment assumptions on the distribution of d_i transfer to (e_i) , we can apply Lemma 2.1 to get

$$\mathbb{E} \left(\exp \left\{ \lambda \sum_{i=1}^n e_i \right\} \middle| \mathcal{G} \right) \leq \exp \left\{ \frac{\lambda^2}{2} \left(\sum_{i=1}^n c_i^2 + \langle M \rangle_n \right) \right\}. \quad (4.4)$$

Plugging this estimate into the previous inequality we obtain

$$\begin{aligned} &\mathbb{P}(M_n \geq \langle M \rangle_n x, A) \\ &\leq \sqrt{\mathbb{E} \left(1_{\{M_n \geq \langle M \rangle_n x, A\}} \exp \left\{ -\lambda \langle M \rangle_n x + \frac{\lambda^2}{2} \left(\sum_{i=1}^n c_i^2 + \langle M \rangle_n \right) \right\} \right)}. \end{aligned}$$

Dividing both sides by $\sqrt{\mathbb{P}(M_n \geq \langle M \rangle_n^2 x, A)}$ gives (4.1). (4.2) follows easily from (4.1).

We now turn to prove (4.3). Let $A = \{M_n/\langle M \rangle_n \geq x, 1/\langle M \rangle_n \leq y \text{ for some } n\}$, and

$$\tau = \inf \left\{ n \geq 1; \frac{M_n}{\langle M \rangle_n} \geq x, 1/\langle M \rangle_n \leq y \right\}, \quad (4.5)$$

with $\inf \emptyset = \infty$. Note that on A , we have that $\tau < \infty$, $\frac{M_\tau}{\langle M \rangle_\tau} \geq x$ and $\langle M \rangle_\tau \geq 1/y$, so $1_A = 1_A 1_{\{\tau < \infty\}} 1_{\{M_\tau/\langle M \rangle_\tau \geq x\}}$ and $\mathbb{P}(A) = \mathbb{P}(M_\tau/\langle M \rangle_\tau \geq x, A)$. Applying Chebychev's inequality first, followed

by Fatou's lemma (valid since $\tau < \infty$ on A) and a use of Corollary 3.4, we have for any $0 < \lambda < 2x$,

$$\begin{aligned}
\mathbb{P}(A) &= \mathbb{P}(M_\tau / \langle M \rangle_\tau \geq x, A) \\
&\leq \mathbb{E} \left(\exp \left\{ \frac{\lambda}{2} (M_\tau - \langle M \rangle_\tau x) \right\} 1_{\{M_\tau / \langle M \rangle_\tau \geq x, A\}} \right) \\
&\leq \liminf_{n \rightarrow \infty} \mathbb{E} \left(\exp \left\{ \frac{\lambda}{2} (M_{\tau \wedge n} - \langle M \rangle_{\tau \wedge n} x) \right\} 1_{\{M_\tau / \langle M \rangle_\tau \geq x, A\}} \right) \\
&\leq \liminf_{n \rightarrow \infty} \sqrt{\mathbb{E} \left(\exp \left\{ \lambda \left(\sum_{i=1}^{\tau \wedge n} e_i - \langle M \rangle_{\tau \wedge n} x \right) \right\} 1_{\{M_\tau / \langle M \rangle_\tau \geq x, A\}} \right)} \\
&= \liminf_{n \rightarrow \infty} \sqrt{\mathbb{E} \left[1_{\{M_\tau / \langle M \rangle_\tau \geq x, A\}} \exp \{-\lambda \langle M \rangle_{\tau \wedge n} x\} \mathbb{E} \left(\exp \left\{ \lambda \sum_{i=1}^{\tau \wedge n} e_i \right\} \middle| \mathcal{G} \right) \right]},
\end{aligned}$$

where the last equality follows since the variables outside the conditional expectation are \mathcal{G} measurable. By Lemma 2.1 we obtain

$$\mathbb{E} \left(\exp \left\{ \lambda \sum_{i=1}^{\tau \wedge n} e_i \right\} \middle| \mathcal{G} \right) \leq \exp \left\{ \frac{\lambda^2}{2} \left(\sum_{i=1}^{\tau \wedge n} c_i^2 + \langle M \rangle_{\tau \wedge n} \right) \right\}. \quad (4.6)$$

Substituting this into the above bound, we get

$$\mathbb{P} \left(\frac{M_\tau}{\langle M \rangle_\tau} \geq x, A \right) \leq \liminf_{n \rightarrow \infty} \sqrt{\mathbb{E} \left[1_{\left\{ \frac{M_{n \wedge \tau}}{\langle M \rangle_{n \wedge \tau}} \geq x, A \right\}} \exp \left\{ -\lambda \langle M \rangle_{\tau \wedge n} x + \frac{\lambda^2}{2} \left(\sum_{i=1}^{\tau \wedge n} c_i^2 + \langle M \rangle_{\tau \wedge n} \right) \right\} \right]}.$$

Since the variable inside the expectation is dominated by $\exp \left(\frac{\lambda^2}{2} \sum_{i=1}^{\infty} c_i^2 \right)$ (for $\lambda < 2x$), and converges to (as n goes to infinity)

$$1_{\left\{ \frac{M_\tau}{\langle M \rangle_\tau} \geq x, A \right\}} \exp \left\{ - \left(\lambda x - \frac{\lambda^2}{2} \right) \langle M \rangle_\tau + \frac{\lambda^2}{2} \sum_{i=1}^{\tau} c_i^2 \right\}$$

using the dominated convergence theorem, we get

$$\mathbb{P} \left(\frac{M_\tau}{\langle M \rangle_\tau} \geq x, A \right) \leq \sqrt{\mathbb{E} \left[1_{\{M_\tau / \langle M \rangle_\tau \geq x, A\}} \exp \left\{ - \left(\lambda x - \frac{\lambda^2}{2} \right) \langle M \rangle_\tau + \frac{\lambda^2}{2} \sum_{i=1}^{\tau} c_i^2 \right\} \right]}.$$

Dividing both sides by $\sqrt{\mathbb{P} \left(\frac{M_\tau}{\langle M \rangle_\tau} \geq x, A \right)}$ gives

$$\mathbb{P} \left(\frac{M_\tau}{\langle M \rangle_\tau} \geq x, A \right) \leq \mathbb{E} \left[\exp \left\{ - \left(\lambda x - \frac{\lambda^2}{2} \right) \langle M \rangle_\tau + \frac{\lambda^2}{2} \sum_{i=1}^{\tau} c_i^2 \right\} \middle| \left(\frac{M_\tau}{\langle M \rangle_\tau} \geq x, A \right) \right].$$

Then, since $\tau < \infty$, $M_\tau / \langle M \rangle_\tau \geq x$ and $\langle M \rangle_\tau \geq 1/y$ on A , we have

$$\mathbb{P} \left(\frac{M_n}{\langle M \rangle_n} \geq x, \frac{1}{\langle M \rangle_n} \leq y \quad \text{for some } n \right) \leq \exp \left\{ - \left(\lambda x - \frac{\lambda^2}{2} \right) \frac{1}{y} + \frac{\lambda^2}{2} \sum_{i=1}^{\infty} c_i^2 \right\}.$$

Minimizing the above expression at $\lambda = \frac{x}{y \sum_{i=1}^{\infty} c_i^2 + 1} < x < 2x$, we have

$$\mathbb{P} \left(\frac{M_n}{\langle M \rangle_n} \geq x, \frac{1}{\langle M \rangle_n} \leq y \quad \text{for some } n \right) \leq \exp \left\{ -\frac{x^2}{2y(y \sum_{i=1}^{\infty} c_i^2 + 1)} \right\}.$$

□

Using Lemma 2.2 or Theorem 2.4 and following the proof of Theorem 4.1, we obtain the following

Theorem 4.2. *Let $\{d_i\}$ be a martingale difference sequence, $M_n := \sum_{i=1}^n d_i$, $\langle M \rangle_n = \sum_{i=1}^n \mathbb{E}(d_i^2 | d_1, \dots, d_{i-1})$. Assume that $\mathbb{P}(d_i \leq c_i) = 1$ where $0 < c_i < \infty$ and let $c = \max_i c_i$. Then, for all $\mathcal{G} := \sigma(d_i, i \geq 1)$ measurable sets A , $x > 0, \lambda > 0$,*

$$\mathbb{P} \left(\frac{M_n}{\langle M \rangle_n} \geq x, A \right) \leq \mathbb{E} \left(\exp \left\{ -\lambda \langle M \rangle_n x + \frac{\lambda^2}{2} e^c \langle M \rangle_n \right\} \mid \frac{M_n}{\langle M \rangle_n} \geq x, A \right) \tag{4.7}$$

and

$$\mathbb{P} \left(\frac{M_n}{\langle M \rangle_n} \geq x, A \right) \leq \sqrt{\mathbb{E} \left(\exp \left\{ -\lambda \langle M \rangle_n x + \frac{\lambda^2}{2} e^c \langle M \rangle_n \right\} 1_A \right)}. \tag{4.8}$$

Minimizing the above expressions (4.7) and (4.8) at $\lambda = x/e^c$, we have

$$\mathbb{P} \left(\frac{M_n}{\langle M \rangle_n} \geq x, A \right) \leq \mathbb{E} \left(\exp \left\{ -\frac{x^2}{2e^c} \langle M \rangle_n \right\} \mid \frac{M_n}{\langle M \rangle_n} \geq x, A \right)$$

and

$$\mathbb{P} \left(\frac{M_n}{\langle M \rangle_n} \geq x, A \right) \leq \sqrt{\mathbb{E} \left(\exp \left\{ -\frac{x^2}{2e^c} \langle M \rangle_n \right\} 1_A \right)}.$$

Furthermore, for any $y > 0$, we have

$$\mathbb{P} \left(\frac{M_n}{\langle M \rangle_n} \geq x, \frac{1}{\langle M \rangle_n} \leq y \quad \text{for some } n \right) \leq \exp \left\{ -\frac{x^2}{2e^c y} \right\}. \tag{4.9}$$

Acknowledgements. Here the author wishes to thank his supervisor–Professor Liming Wu of Université Blaise Pascal and Wuhan University–for many suggestions and helpful discussions. Furthermore, the author is very grateful to the referee for his valuable report.

REFERENCES

- [1] A. Maurer, Abound on the deviation probability for sums of non-negative random variables. *J. Inequa. Pure Appl. Math.* **4** (2003) Article 15.
- [2] V.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. *Ann. Inst. H. Poincaré Probab. Statist.* **30** (1994) 197–211.
- [3] V.H. De La Peña, A general class of exponential inequalities for martingales and ratios. *Ann. Probab.* **27** (1999) 537–564.
- [4] A. Jakubowski, Principle of conditioning in limit theorems for sums of random variables. *Ann. Probab.* **14** (1986) 902–915.
- [5] S. Kwapien and W.A. Woyczyński, Tangent sequences of random variables: basic inequalities and their applications, in *Proceeding of Conference on Almost Everywhere Convergence in Probability and Ergodic Theory*, G.A. Edgar and L. Sucheston Eds., Academic Press, New York (1989) 237–265.
- [6] S. Kwapien and W.A. Woyczyński, *Random series and Stochastic Integrals: Single and Multiple*. Birkhäuser, Boston (1992).
- [7] I. Pinelis, Optimum bounds for the distributions of martingales in Banach space. *Ann. Probab.* **22** (1994) 1679–1706.
- [8] G.L. Wise and E.B. Hall, *Counterexamples in probability and real analysis*. Oxford Univ. Press, New York.(1993).