Statistica Sinica **23** (2013), 1257-1272
doi:http://dx.doi.org/10.5705/ss.2011.287

PROBABILITY AND MOMENT INEQUALITIES UNDER DEPENDENCE

Weidong Liu, Han Xiao and Wei Biao Wu

Shanghai Jiao Tong University, Rutgers University and The University of Chicago

Abstract: We establish Nagaev and Rosenthal-type inequalities for dependent random variables. The imposed dependence conditions, which are expressed in terms of functional dependence measures, are directly related to the physical mechanisms of the underlying processes and are easy to work with. Our results are applied to nonlinear time series and kernel density estimates of linear processes.

Key words and phrases: m-dependence approximation, Nagaev inequality, nonlinear process, non-stationary process, physical dependence measure, Rosenthal inequality.

1. Introduction

Probability and moment inequalities play an important role in the study of properties of sums of random variables. A number of inequalities have been derived for independent random variables; see the recent collection by Lin and Bai (2010). The celebrated Nagaev and Rosenthal inequalities are two useful ones. We first start with the Nagaev inequality. Let $X_1, \ldots, X_n$ be mean 0 independent random variables and $S_n = \sum_{i=1}^n X_i$. Further assume that for all i, $\|X_i\|_p := (\mathbb{E}|X_i|^p)^{1/p} < \infty$, $p > 2$. By Corollary 1.7 in Nagaev (1979), for a positive number x, one has

$$\begin{aligned}\mathbb{P}(S_n \geq x) \leq \sum_{i=1}^n \mathbb{P}(X_i \geq y_i) + \exp\left\{-\frac{a_p x^2}{\sum_{i=1}^n \mathbb{E}(X_i^2 \mathbf{1}_{X_i \leq y_i})}\right\} \\ + \left(\frac{\sum_{i=1}^n \mathbb{E}(X_i^p \mathbf{1}_{0 \leq X_i \leq y_i})}{\beta x y^{p-1}}\right)^{\beta x/y}, \qquad (1.1)\end{aligned}$$

where $y_1, \ldots, y_n > 0$, $y = \max_i\{y_i, 1 \leq i \leq n\}$, $\beta = p/(p+2)$ and $a_p = 2e^{-p}(p+2)^{-2}$. With

$$\mu_{n,p} = \sum_{i=1}^n \mathbb{E}(|X_i|^p)$$

and $y_i = x\beta$ for all $1 \le i \le n$, one obtains from (1.1) that

$$\mathbb{P}(|S_n| \ge x) \le \left(1 + \frac{2}{p}\right)^p \frac{\mu_{n,p}}{x^p} + 2\exp\left(-\frac{a_p x^2}{\mu_{n,2}}\right); \tag{1.2}$$

see Corollary 1.8 in Nagaev (1979). If the random variables X_i, $i \in$, are independent and identically distributed (i.i.d.), then (1.1) implies

$$\mathbb{P}(|S_n| \ge x) \le \left(1 + \frac{2}{p}\right)^p \frac{n\|X_0\|_p^p}{x^p} + 2\exp\left(-\frac{a_p x^2}{n\|X_0\|_2^2}\right). \tag{1.3}$$

Inequalities of this type have applications in insurance and risk management. For example, for a small level $\alpha \in (0,1)$, if $x = x_\alpha$ is such that the right hand side of (1.2) is α, then the α-quantile or value-at-risk of S_n is bounded by x_α since $\mathbb{P}(S_n \ge x_\alpha) \le \alpha$. Inequality (1.2) suggests two types of bounds for the tail probability $\mathbb{P}(S_n \ge x)$: if x^2 is around the variance $\mu_{n,2} = \mathrm{var}(S_n)$, then one can use the Gaussian-type tail $\exp(-a_p x^2/\mu_{n,2})$. If x is larger, the algebraic decay tail $\mu_{n,p}/x^p$ is needed.

In dealing with temporal or time series data, the X_i are often dependent. Then the problem naturally arises on how to generalize the Nagaev inequality to dependent random variables. The latter problem is quite challenging and very few results have been obtained. Under some boundedness conditions on conditional expectations, Basu (1985) derived a similar result. However, the imposed conditions there appear too restrictive and they exclude many commonly used time series models. Nagaev (2001) considered uniformly mixing processes, a very strong type of dependence condition. In Nagaev (2007) he considered martingales. Bertail and Clémençon (2010) dealt with functionals of positive recurrent geometrically ergodic Markov chains; see also Rio (2000).

The Rosenthal inequality provides a bound for the moment $\mathbb{E}(|S_n|^p)$. Rosenthal (1970) proved that if X_i are i.i.d., then there exists a constant B_p such that

$$\mathbb{E}(|S_n|^p) \le B_p \max(\mu_{n,p},\, \mu_{n,2}^{p/2}). \tag{1.4}$$

The calculation of the constant B_p has been extensively discussed in the literature; see Pinelis and Utev (1984), Johnson, Schechtman, and Zinn (1985), Ibragimov and Sharakhmetov (2002), among others. Hitczenko (1990) obtained the best constant for a martingale version of the Rosenthal inequality. Under various types of strong mixing conditions, the Rosenthal type inequalities have been obtained for dependent random variables; see Shao (1988, 1995, 2000), Peligrad (1985, 1989), Utev and Peligrad (2003), and Rio (2000). Rio (2009) and Merlevède and Peligrad (2011) used projections and conditional expetations. Pinelis (2006) applied domination technique to obtain moment inequalities for supermartingales.

In this paper we establish Nagaev and Rosenthal-type inequalities for dependent random variables under easily verifiable dependence conditions. We assume that (X_n) is a stationary causal process of the form

$$X_i = g(\cdots, \varepsilon_{i-1}, \varepsilon_i), \tag{1.5}$$

where $\varepsilon_i, i \in \mathbb{Z}$, are i.i.d. random variables. We adopt the functional dependence measure introduced by Wu (2005). Let $\varepsilon_i, \varepsilon'_j, i, j \in \mathbb{Z}$, be i.i.d. random variables; let $\mathcal{F}_n = (\cdots, \varepsilon_{n-1}, \varepsilon_n)$ and $X'_n = g(\mathcal{F}_{-1}, \varepsilon'_0, \varepsilon_1, \ldots, \varepsilon_n)$. Define the dependence measure

$$\theta_{n,p} = \|X_n - X'_n\|_p$$

and the tail sum

$$\Theta_{m,p} = \sum_{i=m}^{\infty} \theta_{i,p}.$$

Assume throughout the paper that the short-range dependence condition $\Theta_{0,p} < \infty$ holds. Let $S_n = \sum_{k=1}^n X_k$ and $S_n^* = \max_{1 \le i \le n} |S_i|$. Our Rosenthal and Nagaev-type inequalities are expressed in terms of $\theta_{n,p}$ and $\Theta_{m,p}$, and are presented in Sections 2 and 3, respectively. Those inequalities are applied to nonlinear time series and kernel density estimates of linear processes. Section 4 provides an extension to non-stationary processes.

2. A Rosenthal-type Inequality

Throughout the paper we let c_p denote a constant that only depends on p, and whose values may change from place to place. Theorem 1 provides a Rosenthal-type inequality for the maximum partial sum S_n^*. Peligrad, Utev, and Wu (2007) proved that, for $p \ge 2$,

$$\|S_n^*\|_p \le c_p n^{1/2} \Big[\|X_1\|_p + \sum_{j=1}^{n} j^{-3/2} \|\mathbb{E}(S_j | \mathcal{F}_0)\|_p \Big].$$

This inequality can be viewed as a generalization of the Burkholder (1973, 1988) inequality to stationary processes. The Rosenthal inequality has a different flavor in that it relates higher moments of S_n to its variance. Rio (2009) showed a Rosenthal-type inequality for stationary processes: for $2 < p \le 3$, one has

$$\|S_n\|_p \le c_p n^{1/2} \sigma_N + c_p n^{1/p} (\|X_0\|_p + \Delta_N^{1/2} + D_N), \tag{2.1}$$

where $N = \min\{i : 2^i \ge n\}$ and

$$\sigma_N = \|X_0\|_2 + \frac{1}{2} \sum_{l=0}^{N-1} 2^{-l/2} \|\mathbb{E}(S_{2^l} | \mathcal{F}_0)\|_2,$$

$$\Delta_N = \sum_{l=0}^{N-1} 2^{-2l/p} \|\mathbb{E}(S_{2^l}^2|\mathcal{F}_0) - \mathbb{E}(S_{2^l}^2)\|_{p/2},$$
$$D_N = \sum_{l=0}^{N-1} 2^{-l/p} \|\mathbb{E}(S_{2^l}|\mathcal{F}_0)\|_p.$$

Merlevède and Peligrad (2011) obtained the following: for all $p > 2$,

$$\|S_n^*\|_p \le c_p n^{1/p} \left[\|X_0\|_p + \sum_{k=1}^n \frac{\|\mathbb{E}(S_k|\mathcal{F}_0)\|_p}{k^{1+1/p}} + \left(\sum_{k=1}^n \frac{\|\mathbb{E}(S_k^2|\mathcal{F}_0)\|_{p/2}^\delta}{k^{1+2\delta/p}} \right)^{1/(2\delta)} \right], \quad (2.2)$$

where $\delta = \min(1, (p-2)^{-1})$. For $2 < p \le 3$, (2.2) is a maximal version of Rio's inequality (2.1). If the X_i are independent, then $\mathbb{E}(S_k^2|\mathcal{F}_0) = \mathbb{E}(S_k^2) = k\|X_0\|_2^2$ and (2.2) reduces to (1.4). A key step in applying (2.2) is to deal with the quantity

$$\sum_{k=1}^n \frac{\|\mathbb{E}(S_k^2|\mathcal{F}_0)\|_{p/2}^\delta}{k^{1+2\delta/p}} \le \sum_{k=1}^n \frac{\|\mathbb{E}(S_k^2|\mathcal{F}_0) - \mathbb{E}(S_k^2)\|_{p/2}^\delta}{k^{1+2\delta/p}} + \sum_{k=1}^n \frac{[\mathbb{E}(S_k^2)]^\delta}{k^{1+2\delta/p}}.$$

In doing so, one needs to control $\|\mathbb{E}(S_k|\mathcal{F}_0)\|_p$ and $\|\mathbb{E}(S_k^2|\mathcal{F}_0) - \mathbb{E}(S_k^2)\|_{p/2}$. The computation of the latter can be quite involved. Merlevède and Peligrad (2011) provided an inequality in terms of individual summands that involve terms such as $\mathbb{E}(X_i X_j|\mathcal{F}_0)$ and $\mathbb{E}(X_j|\mathcal{F}_0)$. The latter quantities can be controlled by using various mixing coefficients in Bradley (2007), Rio (2000), and Dedecker et al. (2007).

Our Theorem 1 provides an upper bound for $\|S_n^*\|_p$ using the functional dependence measure $\theta_{n,p}$ which is easily computable in many applications; see Wu (2011). We do not need to deal with the quantity $\|\mathbb{E}(S_k^2|\mathcal{F}_0) - \mathbb{E}(S_k^2)\|_{p/2}$. Our inequality is powerful enough so that the behavior of $\|S_n^*\|_p$ for p near boundary can also be depicted; see Example 1.

In order to provide explicit constants in our Rosenthal-type inequality, we need the following version of the Rosenthal inequality for independent variables taken from Johnson, Schechtman, and Zinn (1985)

$$\Big\| \sum_{i=1}^n X_i \Big\|_p \le \frac{14.5p}{\log p} \left(\mu_{n,2}^{1/2} + \mu_{n,p}^{1/p} \right). \quad (2.3)$$

We also need a version of the Burkholder inequality due to Rio (2009): if $X_1, \ldots, X_n$ are martingale differences and $p \ge 2$, then

$$\Big\| \sum_{i=1}^n X_i \Big\|_p^2 \le (p-1) \sum_{i=1}^n \|X_i\|_p^2. \quad (2.4)$$

Theorem 1. *Assume* $\mathbb{E}X_1 = 0$ *and* $\mathbb{E}(|X_1|^p) < \infty$, $p > 2$. *Then*

$$\|S_n^*\|_p \le n^{1/2}\Big[\frac{87p}{\log p}\sum_{j=1}^{n}\theta_{j,2} + 3(p-1)^{1/2}\sum_{j=n+1}^{\infty}\theta_{j,p} + \frac{29p}{\log p}\|X_1\|_2\Big] + n^{1/p}\Big[\frac{87p(p-1)^{1/2}}{\log p}\sum_{j=1}^{n} j^{1/2-1/p}\theta_{j,p} + \frac{29p}{\log p}\|X_1\|_p\Big]. \tag{2.5}$$

Note that, since $p > 2$, we have $\theta_{j,2} \le \theta_{j,p}$. Then the term $\sum_{j=1}^{n}\theta_{j,2} + \sum_{j=n+1}^{\infty}\theta_{j,p}$ in (2.5) can be equivalently replaced by $\Theta_{1,2}+\Theta_{n+1,p}$. Hence we can rewrite (2.5) as

$$\|S_n^*\|_p \le c_p n^{1/2}(\Theta_{1,2} + \|X_1\|_2) + c_p n^{1/p}\Big[\sum_{j=1}^{\infty}\min(j,n)^{1/2-1/p}\theta_{j,p} + \|X_1\|_p\Big].$$

If the X_i are independent, then $\theta_{j,2} = 0$ and $\theta_{j,p} = 0$ for all $j \ge 1$, and (2.5) reduces to the traditional Rosenthal inequality (1.4). The presence of $\Theta_{1,2}$ and $\sum_{j=1}^{\infty}\min(j,n)^{1/2-1/p}\theta_{j,p}$ is due to dependence. It is generally convenient to apply Theorem 1 since the functional dependence measure $\theta_{j,p}$ is directly related to the data-generating mechanism of the underlying processes, and in many cases it can be easily computed; see Wu (2011) for examples of linear and nonlinear processes.

Proof of Theorem 1. For $i \ge 0$ and $j \ge 0$ let

$$S_{i,j} = \sum_{k=1}^{i} X_{k,j}, \text{ where } X_{k,j} = \mathbb{E}(X_k|\varepsilon_{k-j},\ldots,\varepsilon_k).$$

Note that $X_{k,j}$, $k \in \mathbb{Z}$, is j-dependent. Namely $X_{k,j}$ and $X_{k',j}$ are independent if $|k-k'| > j$. We write X_k as

$$X_k = X_k - X_{k,n} + \sum_{j=1}^{n}(X_{k,j} - X_{k,j-1}) + X_{k,0}. \tag{2.6}$$

Then

$$\|S_n^*\|_p \le \Big\|\max_{1\le i\le n}|S_i - S_{i,n}|\Big\|_p + \sum_{j=1}^{n}\Big\|\max_{1\le i\le n}|S_{i,j}-S_{i,j-1}|\Big\|_p + \Big\|\max_{1\le i\le n}|S_{i,0}|\Big\|_p. \tag{2.7}$$

For the second term in (2.7),

$$\Big\|\max_{1\le i\le n}|S_{i,j} - S_{i,j-1}|\Big\|_p \le \|S_{n,j} - S_{n,j-1}\|_p + \Big\|\max_{0\le i\le n-1}\Big|\sum_{k=n-i}^{n}(X_{k,j} - X_{k,j-1})\Big|\Big\|_p. \tag{2.8}$$

Note that $\{X_{k,j} - X_{k,j-1}, 1 \le k \le n\}$ are also j-dependent. Moreover, $X_{n-k,j} - X_{n-k,j-1}$, $0 \le k \le n-1$, are martingale differences with respect to $\sigma(\varepsilon_{n-k-j}, \varepsilon_{n-k-j+1}, \ldots)$. Thus, $\{|\sum_{k=n-i}^{n}(X_{k,j} - X_{k,j-1})|, 0 \le i \le n-1\}$ is a nonnegative submartingale with respect to $\sigma(\varepsilon_{n-i-j}, \varepsilon_{n-i-j+1}, \ldots)$. By the Doob inequality, we have

$$\Big\|\max_{0\le i\le n-1}\Big|\sum_{k=n-i}^{n}(X_{k,j} - X_{k,j-1})\Big|\Big\|_p \le p/(p-1)\cdot\|S_{n,j} - S_{n,j-1}\|_p. \tag{2.9}$$

Write $Y_{i,j} = \sum_{k=1+(i-1)j}^{(ij)\wedge n}(X_{k,j} - X_{k,j-1})$, where $a \wedge b := \min(a,b)$ for two real numbers a and b. With $l = \lfloor n/j \rfloor + 1$, we have

$$|S_{n,j} - S_{n,j-1}| = \Big|\sum_{i=1}^{l} Y_{i,j}\Big|. \tag{2.10}$$

Observe that $Y_{1,j}, Y_{3,j}, \ldots$ are independent and $Y_{2,j}, Y_{4,j}, \ldots$ are also independent. By (2.3),

$$\begin{aligned}\|S_{n,j} - S_{n,j-1}\|_p \le \frac{14.5p}{\log p}\Big[&\Big\|\sum_{i \text{ is odd}} Y_{i,j}\Big\|_2 + \Big\|\sum_{i \text{ is even}} Y_{i,j}\Big\|_2 \\ &+\Big(\sum_{i \text{ is odd}}\|Y_{i,j}\|_p^p\Big)^{1/p} + \Big(\sum_{i \text{ is even}}\|Y_{i,j}\|_p^p\Big)^{1/p}\Big]. \end{aligned}\tag{2.11}$$

By (2.4) we have, for $1 \le i \le l$,

$$\begin{aligned}\|Y_{i,j}\|_p &\le (p-1)^{1/2}[(ij)\wedge n - (i-1)j]^{1/2}\|X_{1,j} - X_{1,j-1}\|_p \\ &\le (p-1)^{1/2}[(ij)\wedge n - (i-1)j]^{1/2}\theta_{j,p}; \\ \|Y_{i,j}\|_2 &\le [(ij)\wedge n - (i-1)j]^{1/2}\theta_{j,2}. \end{aligned}\tag{2.12}$$

Thus (2.11) implies that for $1 \le j \le n$

$$\|S_{n,j} - S_{n,j-1}\|_p \le \frac{29p}{\log p}\left(\sqrt{n}\theta_{j,2} + (p-1)^{1/2}n^{1/p}j^{1/2-1/p}\theta_{j,p}\right). \tag{2.13}$$

By (2.8)-(2.13) and noting that $p/(p-1) \le 2$ when $p \ge 2$, we obtain

$$\begin{aligned}&\sum_{j=1}^{n}\Big\|\max_{1\le i\le n}|S_{i,j} - S_{i,j-1}|\Big\|_p \\ &\quad\le \frac{87p}{\log p}\Big(\sqrt{n}\sum_{j=1}^{n}\theta_{j,2} + (p-1)^{1/2}n^{1/p}\sum_{j=1}^{n}j^{1/2-1/p}\theta_{j,p}\Big). \end{aligned}\tag{2.14}$$

For the first term in (2.7), using the Burkholder inequality (2.4) and a similar argument as (2.8) and (2.9), we have

$$\Big\| \max_{1\le i\le n} |S_i - S_{i,n}| \Big\|_p \le 3(p-1)^{1/2} n^{1/2} \sum_{j=n+1}^{\infty} \theta_{j,p}. \tag{2.15}$$

For the third term in (2.7), noting that $X_{k,0}, k \in \mathbb{Z}$, are independent, again by (2.3) and the Doob inequality, we have

$$\Big\| \max_{1\le i\le n} |S_{i,0}| \Big\|_p \le \frac{29p}{\log p} \left(n^{1/2} \|X_1\|_2 + n^{1/p} \|X_1\|_p \right).$$

This, together with (2.7), (2.15), and (2.14), implies Theorem 1.

Example 1. Consider the nonlinear time series that is expressed in the form of iterated random functions (see for example Diaconis and Freedman (1999)):

$$X_i = F(X_{i-1}, \varepsilon_i) = F_{\varepsilon_i}(X_{i-1}), \tag{2.16}$$

where F is a bivariate measurable function and ε_i, $i \in \mathbb{Z}$, are i.i.d. innovations. Assume that there exist x_0 and $p > 2$ such that

$$\kappa_p := \|x_0 - F_{\varepsilon_0}(x_0)\|_p < \infty \tag{2.17}$$

and the Lipschitz constant

$$L_p := \sup_{x \ne x'} \frac{\|F_{\varepsilon_0}(x) - F_{\varepsilon_0}(x')\|_p}{|x - x'|} < 1. \tag{2.18}$$

By Theorem 2 in Wu and Shao (2004), conditions (2.17) and (2.18) imply that (2.16) has a stationary ergodic solution with $\|X_0\|_p \le |x_0| + \kappa_p/(1 - L_p) =: K_p$. Also the functional dependence measure $\theta_{i,p} \le L_p^i \|X_0 - X_0'\|_p \le 2K_p L_p^i$. We now apply Theorem 1 to the process (X_i). Assume $\mathbb{E}(X_i) = 0$ and let $A = 1/2 - 1/p$. Then

$$\begin{aligned} \frac{\sum_{j=n+1}^{\infty} \theta_{j,p} + n^{-A} \sum_{j=1}^{n} j^A \theta_{j,p}}{2K_p} &\le \sum_{j=n+1}^{\infty} L_p^j + \sum_{j=1}^{n} \left(\frac{j}{n}\right)^A L_p^j \\ &\le \min\left(\sum_{j=1}^{\infty} L_p^j,\ n^{-A} \sum_{j=1}^{\infty} j^A L_p^j \right). \end{aligned} \tag{2.19}$$

Elementary manipulations show that there exists a constant C_A such that, if $L_p \ge 1 - 1/n$, the right hand side of (2.19) is less than $C_A/(1 - L_p)$, while if $L_p \le 1 - 1/n$, it is less than $C_A/(n^A (1 - L_p)^{1+A})$. Combining these two cases, we obtain an upper bound for (2.19) as $C_A \min(1/(1-L_p),\, 1/(n^A(1-L_p)^{1+A}))$. Hence by Theorem 1, we obtain

$$\begin{aligned}\|S_n^*\|_p &\le c_p n^{1/2}\sum_{j=1}^n \theta_{j,2} + c_p n^{1/2}K_2 + c_p n^{1/p}K_p \\ &\quad + n^{1/2}K_p C_A \min\Big(\frac{1}{1-L_p},\, \frac{1}{n^A(1-L_p)^{1+A}}\Big) \\ &\le \frac{2K_2 c_p n^{1/2}}{1-L_2} + \frac{c_p n^{1/2}K_p}{1-L_p}\min(1,\, n^{-A}(1-L_p)^{-A}). \end{aligned} \tag{2.20}$$

If there exists a positive constant λ_0 such that $L_p < 1-\lambda_0$, then the second term in (2.20) has magnitude $O(n^{1/p})$, which together with the first term mimic the classical Rosenthal inequality (1.4). If this well-separateness condition is violated, then we can have a quite interesting behavior. Let $L_p = 1-r_n$ with $r_n \to 0$. If $r_n \ge 1/n$, then the second term in (2.20) has order $O(n^{1/p}r_n^{1/p-5/2})$. If $r_n \le 1/n$, then the order becomes $O(n^{1/2}r_n^{-2})$. As a specific example, consider the ARCH model with $F_{\varepsilon_i}(x) = \varepsilon_i(a^2+b^2x^2)^{1/2}$, where the ε_i are i.i.d. standard normal and $a, b > 0$ are parameters. Then $L_p = \|b\varepsilon_0\|_p$. Choose p_0 such that $L_{p_0} = 1$. Note that $\mathbb{E}(|X_0|^{p_0}) = \infty$ since, for some $C > 0$, $\mathbb{P}(X_0 \ge x) \sim Cx^{-p_0}$ as $x \to \infty$ (see Goldie (1991)). Since $L_p = L_{p_0} + O(|p-p_0|)$, if $p - p_0 = O(r_n)$, $L_p - 1 = O(r_n)$. Overall, as $r_n \downarrow 0$, the second term in (2.20) has bound $n^{1/2}r_n^{-2}\min(1,\, (nr_n)^{-A})$.

3. A Nagaev-type Inequality

Nagaev-type inequalities under dependence have been much less studied than Rosenthal-type inequalities for dependent random variables. If we just apply the Markov inequality and (2.5), we only obtain that

$$\mathbb{P}(S_n^* \ge x) \le \frac{\|S_n^*\|_p^p}{x^p} = O\Big(\frac{n^{p/2}}{x^p}\Big).$$

In comparison, the bound $O(n/x^p)$ in (1.3) is much sharper. We also observe that, according to Borovkov (1972), the Nagaev inequality (1.2) also holds for S_n^*, the maximum of absolute partial sums, when $X_1, \ldots, X_n$ are mean zero independent variables:

$$\mathbb{P}(S_n^* \ge x) \le \Big(1+\frac{2}{p}\Big)^p \frac{\mu_{n,p}}{x^p} + 2\exp\Big(-\frac{2x^2}{e^p(p+2)^2\mu_{n,2}}\Big). \tag{3.1}$$

We will need the Gaussian-like tail function

$$G_q(y) = \sum_{j=1}^\infty e^{-j^q y^2}, \quad y > 0,\ q > 0.$$

Note that $\sup_{y\geq 1} G_q(y)e^{y^2} = G_q(1)e$. Hence if $y \geq 1$, $G_q(y) \leq G_q(1)ee^{-y^2}$. In this section $C, C_1, \ldots$ denote constants that do not depend on x and n.

Theorem 2. (i) *Assume that*

$$\nu := \sum_{j=1}^{\infty} \mu_j < \infty, \ \textit{where } \mu_j = (j^{p/2-1}\theta_{j,p}^p)^{1/(p+1)}. \tag{3.2}$$

Then for all $x > 0$,

$$\begin{aligned}\mathbb{P}(S_n^* \geq x) &\leq c_p \frac{n}{x^p}\left(\nu^{p+1} + \|X_1\|_p^p\right) \\ &\quad + 4\sum_{j=1}^{\infty} \exp\left(-\frac{c_p \mu_j^2 x^2}{n\nu^2\theta_{j,2}^2}\right) + 2\exp\left(-\frac{c_p x^2}{n\|X_1\|_2^2}\right).\end{aligned} \tag{3.3}$$

(ii) *Assume that $\Theta_{m,p} = O(m^{-\alpha})$, $\alpha > 1/2 - 1/p$. Then there exist positive constants C_1, C_2 such that for all $x > 0$,*

$$\mathbb{P}(S_n^* \geq x) \leq \frac{C_1\Theta_{0,p}^p\, n}{x^p} + 4G_{1-2/p}\left(\frac{C_2 x}{\sqrt{n}\Theta_{0,p}}\right). \tag{3.4}$$

(iii) *If $\Theta_{m,p} = O(m^{-\alpha})$, $\alpha < 1/2 - 1/p$, then a variant of* (3.4) *holds:*

$$\mathbb{P}(S_n^* \geq x) \leq \frac{C_1\Theta_{0,p}^p\, n^{p(1/2-\alpha)}}{x^p} + 4G_{(p-2)/(p+1)}\left(\frac{C_2 x}{n^{(2p-1-2\alpha p)/(2+2p)}\Theta_{0,p}}\right). \tag{3.5}$$

If the X_i are independent, then $\theta_{j,2} = \theta_{j,p} = 0$ for all $j \geq 1$, and hence (3.3) reduces to the traditional Nagaev inequality (3.1).

We remark that those inequalities are non-asymptotic and they hold for any n and x. The exponential term in (3.3) decays to zero very quickly as $j \to \infty$. If $x = \sqrt{n}\nu^{1+1/p}y$ with $y > 0$, then $\mu_j^2 x^2/(n\nu^2\theta_{j,2}^2) \geq j^{1-2/p}y^2$ and

$$\sum_{j=1}^{\infty} \exp\Big(-\frac{c_p\mu_j^2 x^2}{n\nu^2\theta_{j,2}^2}\Big) \leq \sum_{j=1}^{\infty} \exp(-c_p j^{1-2/p} y^2)$$

is an upper bound for the second term in (3.3). Consider the two cases $y \geq 1$ and $y < 1$ separately, we conclude that there exists constants c_p and c_p' such that the second term in (3.3) is bounded by $c_p' \exp[-c_p x^2/(n\nu^{2+2/p})]$.

We now compare conditions on dependence in (i) and (ii). Consider the special case $\theta_{j,p} = j^{-\beta}$. Then (3.2) requires $\beta > 3/2$, and (ii) only requires $\beta > 3/2 - 1/p$. On the other hand, (3.2) implies $\Theta_{m,p} = o(m^{1/p-1/2})$ since $\sum_{j=m}^{2m-1} \theta_{j,p} \leq (\sum_{j=m}^{2m-1} \theta_{j,p}^{1/q})^q$, where $q = 1+1/p$, and $m^{(p/2-1)/(1+p)} \sum_{j=m}^{2m-1} \theta_{j,p}^{1/q} \leq$

$\sum_{j=m}^{2m-1} \mu_j = o(1)$. In case (iii) the dependence is stronger; as compensation, we need a larger numerator $n^{p(1/2-\alpha)}$ than n, and the term $n^{(2p-1-2\alpha p)/(2+2p)}$ in (3.5) is also larger than $\sqrt{n}$.

Proof of Theorem 2. (i) We use the decomposition (2.6). Let λ_j, $j = 1, \ldots, n$ be a positive sequence such that $\sum_{j=1}^n \lambda_j \le 1$. For $i \in \mathbb{Z}$ and $\ell \in \mathbb{N}$, write $\lfloor i \rfloor_\ell := \lfloor i/\ell \rfloor \ell$. Define

$$M_{i,j} = \sum_{k=1}^{i} (X_{k,j} - X_{k,j-1}) \text{ and } M_{n,j}^* = \max_{1 \le i \le n} |M_{i,j}|. \tag{3.6}$$

Then $S_{n,n} - S_{n,0} = \sum_{k=1}^n (X_{k,n} - X_{k,0}) = \sum_{j=1}^n M_{n,j}$. For each $1 \le j \le n$, as in (2.10), let $Y_{i,j} = \sum_{k=1+(i-1)j}^{(ij)\wedge n} (X_{k,j} - X_{k,j-1})$, $1 \le i \le l$, where $l = \lfloor n/j \rfloor + 1$. Define $W_{s,j}^e = \sum_{i=1}^s (1+(-1)^i)/2 \cdot Y_{i,j}$ and $W_{s,j}^o = \sum_{i=1}^s (1-(-1)^i)/2 \cdot Y_{i,j}$. Then

$$\begin{aligned} \mathbb{P}(M_{n,j}^* \ge 3\lambda_j x) &\le \mathbb{P}\left(\max_{i \le n} |M_{\lfloor i \rfloor_j, j}| \ge 2\lambda_j x\right) + \mathbb{P}\left(\max_{i \le n} |M_{\lfloor i \rfloor_j, j} - M_{i,j}| \ge \lambda_j x\right) \\ &\le \mathbb{P}\left(\max_{s \le l} |W_{l,j}^e| \ge \lambda_j x\right) + \mathbb{P}\left(\max_{s \le l} |W_{l,j}^o| \ge \lambda_j x\right) \\ &\quad + \frac{n}{j} \mathbb{P}\Big(\max_{i \le j} |M_{i,j}| \ge \lambda_j x\Big). \end{aligned} \tag{3.7}$$

Since $Y_{2,j}, Y_{4,j}, \ldots,$ are independent, from (3.1) and (2.12) we obtain

$$\begin{aligned} \mathbb{P}(\max_{s \le l} |W_{l,j}^e| \ge \lambda_j x) &\le c_p \frac{(n/j)\mathbb{E}(|Y_{2,j}|^p)}{(\lambda_j x)^p} + 2\exp\Big(-c_p \frac{(\lambda_j x)^2}{n\theta_{j,2}^2}\Big) \\ &\le c_p \frac{n}{x^p} \frac{j^{p/2-1}\theta_{j,p}^p}{\lambda_j^p} + 2\exp\Big(-c_p \frac{(\lambda_j x)^2}{n\theta_{j,2}^2}\Big). \end{aligned} \tag{3.8}$$

A similar inequality holds for $W_{s,j}^o$. For the last term in (3.7), noting that, by (2.4) and the Doob inequality,

$$\mathbb{E}(\max_{1 \le i \le j} |M_{i,j}|^p) \le 2^{p-1} \mathbb{E}\Big(|M_{j,j}|^p + \max_{1 \le i \le j} \Big| \sum_{k=i}^{j} (X_{k,j} - X_{k,j-1}) \Big|^p\Big) \le c_p j^{p/2} \theta_{j,p}^p,$$

we have

$$\mathbb{P}(M_{n,j}^* \ge 3\lambda_j x) \le c_p \frac{n}{x^p} \frac{j^{p/2-1}\theta_{j,p}^p}{\lambda_j^p} + 4\exp\Big(-c_p \frac{(\lambda_j x)^2}{n\theta_{j,2}^2}\Big). \tag{3.9}$$

Since $\|X_{1,0}\| \le \|X_1\|_2$ and $\|X_{1,0}\|_p \le \|X_1\|_p$, by (3.1), we have

$$P\left(\max_{1 \le i \le n} |S_{i,0}| \ge x\right) \le c_p \frac{n\|X_1\|_p^p}{x^p} + 2\exp\left(-\frac{c_p x^2}{n\|X_1\|_2^2}\right).$$

Choose $\lambda_j = \mu_j/\nu$. By (2.6) and (2.15), we obtain (3.3) in view of

$$\Theta_{n+1,p}^{p/(p+1)} \le \sum_{l=n+1}^{\infty} \theta_{l,p}^{p/(p+1)} \le \sum_{l=n+1}^{\infty} \left(\frac{l}{n}\right)^{(p/2-1)/(p+1)} \theta_{l,p}^{p/(p+1)},$$

$$\begin{aligned}\mathbb{P}(S_n^* \ge 5x) &\le \sum_{j=1}^{n} \mathbb{P}(M_{n,j}^* \ge 3\lambda_j x)\\ &\quad + P\left(\max_{1\le i\le n} |S_i - S_{i,n}| \ge x\right) + P\left(\max_{1\le i\le n} |S_{i,0}| \ge x\right).\end{aligned}$$

(ii) Let $0 = \tau_0 < \tau_1 < \ldots < \tau_L = n$ be a sequence of integers. As in (3.6), write

$$S_{i,n} - S_{i,0} = \sum_{l=1}^{L} \breve{M}_{i,l}, \text{ where } \breve{M}_{i,l} = \sum_{k=1}^{i} (X_{k,\tau_l} - X_{k,\tau_{l-1}}).$$

Then there exists a constant $c_p > 0$ such that

$$\frac{\|\breve{M}_{i,l}\|_p}{\sqrt{i}} \le c_p \sum_{i=1+\tau_{l-1}}^{\tau_l} \theta_{i,p} =: c_p \breve{\theta}_{l,p} \text{ and } \frac{\|\breve{M}_{i,l}\|_2}{\sqrt{i}} \le \sum_{i=1+\tau_{l-1}}^{\tau_l} \theta_{i,2} := \breve{\theta}_{l,2}.$$

Let $\breve{M}_{n,l}^* = \max_{i\le n} |\breve{M}_{i,l}|$. Again let $\breve{\lambda}_1, \ldots \breve{\lambda}_L$ be a positive sequence with $\sum_{l=1}^{L} \breve{\lambda}_l \le 1$. With the argument in (3.7), (3.8), and (3.9), we similarly obtain

$$\mathbb{P}(\breve{M}_{n,l}^* \ge 3\breve{\lambda}_l x) \le c_p \frac{n}{x^p} \frac{\tau_l^{p/2-1} \breve{\theta}_{l,p}^p}{\breve{\lambda}_l^p} + 4 \exp\Big(-c_p \frac{(\breve{\lambda}_l x)^2}{n \breve{\theta}_{l,2}^2}\Big).$$

Let $\breve{\mu}_l = (\tau_l^{p/2-1} \breve{\theta}_{l,p}^p \Theta_{0,p}^{-p})^{1/(p+1)}$, $\breve{\nu}_L = \sum_{l=1}^{L} \breve{\mu}_l$, and $\breve{\lambda}_l = \breve{\mu}_l/\breve{\nu}_L$. Using (3.10), we have

$$\begin{aligned}\mathbb{P}(S_n^* \ge 5x) &\le \sum_{l=1}^{L} \mathbb{P}(\breve{M}_{n,l}^* \ge 3\breve{\lambda}_l x)\\ &\quad + P\left(\max_{1\le i\le n} |S_i - S_{i,n}| \ge x\right) + P\left(\max_{1\le i\le n} |S_{i,0}| \ge x\right)\\ &\le c_p \frac{n}{x^p} \nu_L^{p+1} + 4 \sum_{l=1}^{L} \exp(-c_p \frac{(\breve{\lambda}_l x)^2}{n \breve{\theta}_{l,2}^2})\\ &\quad + c_p \frac{n^{p/2} \Theta_{n+1,p}^p}{x^p} + c_p \frac{n \|X_0\|_p^p}{x^p} + 2 \exp\left(-\frac{c_p x^2}{n \|X_0\|^2}\right). \quad (3.10)\end{aligned}$$

We now show that the above relation implies (3.4). Let $A = (1/2 - 1/p)p/(1+p)$ and $B = \alpha p/(1+p)$. Since $\breve{\theta}_{l,p} \le \Theta_{\tau_{l-1}+1,p}$, we have

$$\breve{\nu}_L = \sum_{l=1}^{L} (\tau_l^{p/2-1} \breve{\theta}_{l,p}^p)^{1/(p+1)}$$

$$= O(1) \sum_{l=1}^{L} (\tau_l^{1/2-1/p} \tau_{l-1}^{-\alpha})^{p/(1+p)}$$
$$= O(1) \sum_{l=1}^{L} \frac{\tau_l^A}{\tau_{l-1}^B}.$$

If $A < B$, choose $\rho \in (A/B, 1)$, $L = 1 + \lfloor (\log\log n)/(\log \rho^{-1}) \rfloor$, and $\tau_l = \lfloor n^{\rho^{L-l}} \rfloor$, $1 \le l \le L$. Since $A < \rho B$ and $\tau_l^A/\tau_{l-1}^B \sim n^{\rho^{L-l}(A-\rho B)}$, elementary calculation shows that $\sum_{l=1}^{L} \tau_l^A/\tau_{l-1}^B = O(1)$. Let $x = \sqrt{n}\Theta_{0,p}\breve{\nu}_L^{1+1/p} y$, then

$$\frac{(\breve{\lambda}_l x)^2}{n\breve{\theta}_{l,2}^2} = \frac{\breve{\mu}_l^2 \breve{\nu}_L^{2/p} \Theta_{0,p}^2 y^2}{\breve{\theta}_{l,2}^2} \ge \frac{\breve{\mu}_l^{2+2/p} \Theta_{0,p}^2 y^2}{\breve{\theta}_{l,2}^2} = \tau_l^{1-2/p} y^2,$$

and the second term on the right hand side of (3.10) is bounded by $\sum_{l=1}^{\infty} 4\exp(-c_p l^{1-2/p} y^2)$, which implies (3.4).

If $A > B$, let $r = (A/B)^{1/(A-B)}$, $\tau_l = \lfloor n/r^{L-l} \rfloor$, and $L = 1 + \lfloor (\log n - 1)/(\log r) \rfloor$. Then $\breve{\nu}_L = O(n^{A-B})$. Since $A - B = (1/2 - 1/p - \alpha)p/(1+p)$, we have, by (3.10),

$$\mathbb{P}(|S_n^*| \ge 5x) = \frac{O(n^{p(1/2-\alpha)})\Theta_{0,p}^p}{x^p} + 4\sum_{l=1}^{L} \exp\Big(-c_p \frac{(\breve{\lambda}_l x)^2}{n\breve{\theta}_{l,2}^2}\Big) + 2\exp\Big(-\frac{c_p x^2}{n\|X_0\|^2}\Big).$$

Let $x = \sqrt{n}\Theta_{0,p}\breve{\nu}_L y$, then

$$\frac{(\breve{\lambda}_l x)^2}{n\breve{\theta}_{l,2}^2} = \frac{\breve{\mu}_l^2 \Theta_{0,p}^2 y^2}{\breve{\theta}_{l,2}^2} = \frac{(\breve{\theta}_{l,p}/\Theta_{0,p})^{2p/(p+1)} \tau_l^{(p-2)/(p+1)} y^2}{(\breve{\theta}_{l,2}/\Theta_{0,p})^2} \ge \tau_l^{(p-2)/(p+1)} y^2,$$

and (3.5) follows.

Remark 1. We consider the boundary case of Theorem 2 with $\alpha = 1/2 - 1/p$. Recall the proof of the Theorem 2 for the definitions of A, B, and $\breve{\nu}_L$. Now we have $A = B$. Let $\tau_l = 2^l$ for $1 \le l < L$, where $L = \lfloor (\log n)/(\log 2) \rfloor$, we have $\nu_L = O(\log n)$. Then the argument there implies the following upper bound: there exist positive constants C_1 and C_2 such that for all $x > 0$,

$$\mathbb{P}(|S_n^*| \ge x) \le \frac{C_1 \Theta_{0,p}^p\, n(\log n)^{p+1}}{x^p} + 4G_{(p-2)/(p+1)}\Big[\frac{C_2 x}{\Theta_{0,p}\sqrt{n}\log n}\Big].$$

Example 2. (Kernel Density Estimation) Consider estimating the marginal density of the linear process $Y_i = \sum_{j=0}^{\infty} a_j \varepsilon_{i-j}$, where $(a_j)_{j=0}^{\infty}$ are real coefficients and the ε_j are i.i.d. innovations with density f_ε satisfying $f_* := \sup_u [f_\varepsilon(u) + |f'_\varepsilon(u)|] < \infty$. Based on the data $Y_1, \ldots, Y_n$, we estimate the marginal density f of Y_i by

$$\hat{f}_n(u) = \frac{1}{nb_n} \sum_{i=1}^{n} K\Big(\frac{u - Y_i}{b_n}\Big),$$

where b_n is the bandwidth sequence with $b_n \to 0$ and $nb_n \to \infty$, and K is a bounded kernel function with support $[-1, 1]$. We want an upper bound for the tail probability $\mathbb{P}(|\hat{f}_n(u) - \mathbb{E}\hat{f}_n(u)| \geq x)$. The latter problem has been studied for i.i.d. random variables; see Louani (1998), Gao (2003), and Joutard (2006), among others. However, the case of dependent random variables has been largely untouched.

Assume $a_0 = 1$. Let $W_{i-1} = \sum_{j=1}^{\infty} a_j \varepsilon_{i-j} = Y_i - \varepsilon_i$, $\mathcal{F}_{i-1} = (\ldots, \epsilon_{i-2}, \epsilon_{i-1})$, and

$$\hat{f}_n^{\diamond}(u) = \frac{1}{nb_n} \sum_{i=1}^{n} \mathbb{E}\Big[K\Big(\frac{u - Y_i}{b_n}\Big)\Big|\mathcal{F}_{i-1}\Big] = \frac{1}{n} \sum_{i=1}^{n} X_i,$$

where

$$X_i = \int_{-1}^{1} K(v) f_\varepsilon(u - W_{i-1} - vb_n)dv.$$

We compute the functional dependence measure of X_i. Let $W'_{i-1} = W_{i-1} - a_i\varepsilon_0 + a_i\varepsilon'_0$, where $\varepsilon'_0, \varepsilon_l, l \in \mathbb{Z}$, are i.i.d. Then $|X_i - X'_i| \leq |a_i||\varepsilon_0 - \varepsilon'_0|\kappa$, where $\kappa = f_* \int_{-1}^{1} |K(v)|dv$. Assume $\mathbb{E}(|\varepsilon_0|^q) < \infty$, $q > 0$. Since $|X_i| \leq \kappa$ and $|X'_i| \leq \kappa$, we obtain

$$\begin{aligned}\mathbb{E}(|X_i - X'_i|^p) &\leq \mathbb{E}[\min(2\kappa, |a_i||\varepsilon_0 - \varepsilon'_0|\kappa)^p]\\ &\leq (2\kappa)^p \mathbb{E}[\min(1, |a_i||\varepsilon_0 - \varepsilon'_0|)^{\min(p,q)}]\\ &\leq (2\kappa)^p (|a_i|^{\min(p,q)}) \mathbb{E}[|\varepsilon_0 - \varepsilon'_0|^{\min(p,q)}].\end{aligned}$$

Hence the the functional dependence measure of X_i, $\theta_{i,p} = O(|a_i|^{\min(1,q/p)})$. Assume that

$$\sum_{i=1}^{\infty} (i^{p/2-1}|a_i|^{\min(q,p)})^{1/(p+1)} < \infty.$$

By Theorem 2(i), there exists constants $C_1, C_2, C_3 > 0$ such that, for all $y > 0$,

$$\begin{aligned}\mathbb{P}[|\hat{f}_n^{\diamond}(u) - \mathbb{E}\hat{f}_n^{\diamond}(u)| \geq y] &= \mathbb{P}(|X_1 + \cdots + X_n - n\mathbb{E}X_1| \geq ny)\\ &\leq \frac{C_1 n}{(ny)^p} + C_3 \exp\Big(-\frac{C_2(ny)^2}{n}\Big). \qquad (3.11)\end{aligned}$$

Since $D_i := K((u - Y_i)/b_n) - \mathbb{E}[K((u - Y_i)/b_n)|\mathcal{F}_{i-1}]$, $i = 1, \ldots, n$, are martingale differences bounded by $K_2 = 2\sup_u |K(u)|$ and $\mathbb{E}(D_i^2|\mathcal{F}_{i-1}) \leq b_n K_3$, where $K_3 = f_* \int_{-1}^{1} K^2(v)dv$, by Freedman (1975)'s martingale exponential inequality, we obtain

$$\begin{aligned}\mathbb{P}[|f_n(u) - \hat{f}_n^{\diamond}(u)| \geq y] &\leq 2\exp\Big[-\frac{(nb_n y)^2}{2nb_n y K_2 + 2nb_n K_3}\Big]\\ &= 2\exp\Big[-\frac{nb_n y^2}{2yK_2 + 2K_3}\Big]. \qquad (3.12)\end{aligned}$$

For sufficiently large n, $b_n \leq K_3$. So by (3.11) and (3.12), we have the upper bound

$$\mathbb{P}[|f_n(u) - \mathbb{E}f_n(u)| \geq 2y] \leq 3\exp\Big[-\frac{nb_n y^2}{2yK_2 + 2K_3}\Big] + \frac{C_1 n}{(ny)^p}.$$

Hence, if $y \geq (\log n)/\sqrt{nb_n}$, the tail probability $\mathbb{P}[|f_n(u) - \mathbb{E}f_n(u)| \geq 2y]$ has an upper bound with order $n/(ny)^p = n^{1-p}y^{-p}$.

4. Extension to Non-stationary Processes

The inequalities for the stationary case can be generalized to causal non-stationary processes without essential difficulties. Consider the non-stationary process

$$X_i = g_i(\cdots, \varepsilon_{i-1}, \varepsilon_i), \tag{4.1}$$

where $\varepsilon_i, i \in \mathbb{Z}$, are i.i.d. and the g_i are measurable functions. If g_i does not depend on i, then (4.1) reduces to the stationary process (1.5). For any random vector $(X_1, \ldots, X_n)$, one can always find $g_1, \ldots, g_n$ and independent random variables ε_i uniformly distributed over $[0, 1]$ such that $(X_i)_{i=1}^n$ and $(g_i(\varepsilon_1, \ldots, \varepsilon_i))_{i=1}^n$ have the same distribution (see for example Rosenblatt (1952) and Wu and Mielniczuk (2010)). We define a uniform functional dependence measure. Again let $\varepsilon_i, \varepsilon'_j, i, j \in \mathbb{Z}$, be i.i.d. and assume for all i that $\mathbb{E}(|X_i|^p) < \infty$, $p > 2$, and $\mathbb{E}(X_i) = 0$. For $m \geq 0$ let

$$\theta^*_{m,p} = \sup_i \|X_i - g_i(\cdots, \varepsilon_{i-m-1}, \varepsilon'_{i-m}, \varepsilon_{i-m+1}, \ldots, \varepsilon_i)\|_p,$$

and define the tail sum $\Theta^*_{m,p} = \sum_{j=m}^{\infty} \theta^*_{j,p}$. A careful check of the proofs of Theorems 1 and 2 suggest that they remain valid if we instead use the uniform functional dependence measure $\theta^*_{m,p}$. The details are omitted.

Acknowledgements

We are grateful to Dr. Xiaohui Chen and two referees for their helpful comments and suggestions that resulted in a much improved version. Weidong Liu's research was supported by NSFC, Grant No.11201298, the Program for Professor of Special Appointment (Eastern Scholar) at Shanghai Institutions of Higher Learning, Foundation for the Author of National Excellent Doctoral Dissertation of PR China and the startup fund from SJTU. Han Xiao's research was supported in part by the US National Science Foundation (DMS–1209091). Wei Biao Wu's research was supported in part by the US National Science Foundation (DMS–0906073 and DMS–1106970).

References

Basu, A. K. (1985). Probability inequalities for sums of dependent random vector and applications. *Sankhyā Ser. A* **47**, 333-349.

Bertail, P. and Clémençon, S. (2010). Sharp bounds for the tails of functionals of Markov chains. *Theory Probab. Appl.* **54**, 505-515.

Borovkov, A. A. (1972). Notes on inequalities for sums of independent variables. *Theory Probab. Appl.* **17**, 556-557.

Bradley, R. C. (2007). *Introduction to Strong Mixing Conditions.* Vol. 1. Kendrick Press, Heber City, UT.

Burkholder, D. L. (1973). Distribution function inequalities for martingales. *Ann. Probab.* **1**,19-42.

Burkholder, D. L. (1988). Sharp inequalities for martingales and stochastic integrals. *Astérisque* **157-158**, 75-94.

Dedecker, J., Doukhan, P., Lang, G., León, J., Louhichi, S. and Prieur, C. (2007). *Weak Dependence: with Examples and Applications*, Volume 190 of *Lecture Notes in Statistics*. Springer, New York.

Diaconis, P. and Freedman, D. (1999). Iterated random functions. *SIAM Rev.* **41**, 45-76.

Freedman, D. (1975). On tail probabilities for martingales. *Ann. Probab.* **3**,100-118.

Gao, F. (2003). Moderate deviations and large deviations for kernel density estimators. *J. Theoret. Probab.* **16**, 401-418.

Goldie, C. (1991). Implicit renewal theory and tails of solutions of random equations. *Ann. Appl. Probab.* **1**, 126-166.

Hitczenko, P. (1990). Best constants in martingale version of Rosenthal's inequality. *Ann. Probab.* **18**, 1656-1668.

Ibragimov, R. and Sharakhmetov, S. (2002). The exact constant in the Rosenthal inequality for random variables with mean zero. *Theory Probab. Appl.* **46**, 127-131.

Johnson, W. B., Schechtman, G. and Zinn, J. (1985). Best constants in moment inequalities for linear combinations of independent and exchangeable random variables. *Ann. Probab.* **13**, 234-253.

Joutard, C. (2006). Sharp large deviations in nonparametric estimation. *J. Nonparametr. Stat.* **18**, 293-306.

Lin, Z. and Bai, Z. D. (2010). *Probability Inequalities.* Science Press Beijing, Beijing.

Louani, D. (1998). Large deviations limit theorems for the kernel density estimator. *Scand. J. Statist.* **25**, 243-253.

Merlevède, F. and Peligrad, M. (2011). Rosenthal-type inequalities for the maximum of partial sums of stationary processes and examples. *Ann. Probab.* To appear.

Nagaev, S. V. (1979). Large deviations of sums of independent random variables. *Ann. Probab.* **7**, 745-789.

Nagaev, S. V. (2001). Probabilistic and moment inequalities for dependent random variables. *Theory Probab. Appl.* **45**, 152-160.

Nagaev, S. V. (2007). On probability and moment inequalities for supermartingales and martingales. *Theory Probab. Appl.* **51**, 367-377.

Peligrad, M. (1985). Convergence rates of the strong law for stationary mixing sequences. *Z. Wahrsch. Verw. Gebiete* **70**, 307-314.

Peligrad, M. (1989). The r-quick version of the strong law for stationary φ-mixing sequences. In *Almost Everywhere Convergence (Columbus, OH(1988)*, pages 335-348. Academic Press, Boston, MA.

Peligrad, M., Utev, S. and Wu, W. B. (2007). A maximal $\mathbb{L}_p$-inequality for stationary sequences and its applications. *Proc. Amer. Math. Soc.* **135**, 541-550.

Pinelis, I. (2006). On normal domination of (super)martingales. *Electron. J. Probab.* **11**, 1049-1070.

Pinelis, I. F. and Utev, S. A. (1984). Estimates of moments of sums of independent random variables. *Teor. Veroyatnost. i Primenen.* **29**, 554-557.

Rio, E. (2000). *Théorie Asymptotique Des Processus Aléatoires Faiblement Dépendants*, volume 31 of *Mathématiques & Applications* (Berlin) [Mathematics & Applications]. Springer-Verlag, Berlin.

Rio, E. (2009). Moment inequalities for sums of dependent random variables under projective conditions. *J. Theoret. Probab.* **22**, 146-163.

Rosenblatt, M. (1952). Remarks on a multivariate transformation. *Ann. Math. Statist.* **23**, 470-472.

Rosenthal, H. P. (1970). On the subspaces of L^p $(p > 2)$ spanned by sequences of independent random variables. *Israel J. Math.* **8**, 273-303.

Shao, Q. M. (1988). A moment inequality and its applications. *Acta Math. Sinica* **31**, 736-747.

Shao, Q. M. (1995). Maximal inequalities for partial sums of ρ-mixing sequences. *Ann. Probab.* **23**, 948-965.

Shao, Q.-M. (2000). A comparison theorem on moment inequalities between negatively associated and independent random variables. *J. Theoret. Probab.* **13**, 343-356.

Utev, S. and Peligrad, M. (2003). Maximal inequalities and an invariance principle for a class of weakly dependent random variables. *J. Theoret. Probab.* **16**, 101-115.

Wu, W. B. (2005). Nonlinear system theory: another look at dependence. *Proc. Natl. Acad. Sci. USA* **102**, 14150-14154.

Wu, W. B. (2011). Asymptotic theory for stationary processes. *Stat. Interface* **4**(207-226.

Wu, W. B. and Mielniczuk, J. (2010). A new look at measuring dependence. In *Dependence in probability and statistics*, Volume 200 of *Lecture Notes in Statistics*, 123-142. Springer, Berlin.

Wu, W. B. and Shao, X. (2004). Limit theorems for iterated random functions. *J. Appl. Probab.* **41**, 425-436.

Department of Mathematics, Institute of Natural Sciences and MOE-LSC, Shanghai Jiao Tong University, Shanghai, China.

E-mail: liuweidong99@gmail.com

Department of Statistics and Biostatistics, Rutgers University, Piscataway, NJ 08854, U.S.A.

E-mail: hxiao@stat.rutgers.edu

Department of Statistics, The University of Chicago, Chicago, IL 60637, U.S.A.

E-mail: wbwu@galton.uchicago.edu

(Received November 2011; accepted November 2012)