A Note on Quantile Coupling Inequalities and Their Applications

Harrison H. Zhou
Department of Statistics, Yale University, New Haven, CT 06520, USA.
E-mail:huibin.zhou@yale.edu

June 21, 2006

Abstract

A relationship between the large deviation and quantile coupling is studied. We apply this relationship to the coupling of the sum of n i.i.d. symmetric random variables with a normal random variable, improving the classical quantile coupling inequalities (the key part in the celebrated KMT constructions) with a rate $1/\sqrt{n}$ for random variables with continuous distributions, or the same rate modulo constants for the general case. Applications to the asymptotic equivalence theory and nonparametric function estimation are discussed.

Keywords: Quantile Coupling; Large deviation; KMT/Hungarian construction; Asymptotic equivalence; Function estimation

RUNNING TITLE: Quantile Coupling Inequalities.

1 Introduction

The KMT/Hungarian construction in Komlós, Major, and Tusnády (1975) is considered one of the most important statistics and probability results of the last forty years. It has been widely applied in many areas of statistics and probability (cf. Shorack and Wellner (1986)). The quantile coupling of the sum of i.i.d. Bernoulli(1/2) with a normal random variable lies at the heart of KMT/Hungarian construction for empirical process. In this paper, we study the coupling of the sum of n i.i.d. *symmetric* random variables with a normal random variable, and improve the classical quantile coupling bounds with a rate $1/\sqrt{n}$ for random variables whose distributions are absolutely continuous with respect to a Lebesgue measure, or the same rate modulo constants for the general case. This paper can be regarded as a generalization of Carter and Pollard (2004), which studied the coupling of Binomial$(n, 1/2)$ and a normal random variable and improved the classical quantile coupling bounds (called Tusnády's Lemma) with a rate $1/\sqrt{n}$ modulo constants.

The KMT construction played a key role in the progress of the asymptotic equivalence theory in the last decade. Nussbaum (1996), a breakthrough of asymptotic equivalence theory, established the asymptotic equivalence of density estimation and Gaussian white noise under a Hölder smoothness condition. A major step toward the proof of this equivalence result is the functional KMT construction for empirical process by Koltchinskii (1994), where lying at the heart of the construction is Tusnády's Lemma. The impact of this result is that an asymptotically optimal result in one of these nonparametric models automatically yields an analogous results in the other model. Starting from Donoho and Johnstone (1995), Besov smoothness constraint became a standard assumption in the nonparametric estimation. Recently, Brown, Carter, Low and Zhang (2004) extended the result of Nussbaum (1996) under a sharp Besov smoothness constraint via the improved Tusnády's inequality by Carter and Pollard (2004). This asymptotic equivalence result is considered an important progress in this area. It is might be worthwhile to mention that the classical Tusnády's inequality may not be sufficient to establish asymptotic equivalence under the conditions stated in the paper of Brown, carter, Low and Zhang (2004). General quantile coupling inequalities (see Sakhanenko (1984) and Komlós, Major, and Tusnády (1975)) led to an extension of asymptotic equivalence theory in Nussbaum (1996) to general nonparametric estimation models (see Grama and Nussbaum (1998, 2002a, 2002b)). Among those models an important one is the spectral density estimation model. In Zhou (2004) or Golubev, Nussbaum and Zhou (2005), we applied a sharp quantile coupling bound between a Beta and a normal random variable (a special case of general results in this paper) to establish the asymptotic equivalence of the spectral density estimation and Gaussian white noise under a Besov smoothness constraint.

One possibly interesting application of our result is coupling a median statistic with a normal random variable. We obtain a sharp quantile coupling inequality which also improves the classical quantile coupling bounds with a rate $1/\sqrt{n}$ under certain smoothness conditions for the distribution function (see section 5). It includes the Cauchy distribution as a special case. This coupling result may be of independent interest because of the fundamental role of median in statistics.

The paper is organized as follows. In section 2, we list basic results for the quantile coupling of the sum of n i.i.d. symmetric random variable. In section 3, we give a general assumption to obtain a quantile coupling inequality with an improved rate $1/\sqrt{n}$, which immediately implies a sharp quantile coupling result for the sum of n i.i.d. symmetric random variable with continuous distribution. Section 4 gives a general assumption to obtain a quantile coupling inequality with an improved rate modulo constants. Some applications of the coupling results are discussed in section 5.

2 Basic Results

The quantile coupling of the sum of i.i.d. Bernoulli(1/2) (or Binomial$(n, 1/2)$) with a normal random variable is a key step in KMT/Hungarian coupling of the empirical distribution with a Brownian bridge in Komlós, Major, and Tusnády (1975). The tight quantile coupling bound for Binomial$(n, 1/2)$ in Tusnády (1977) is formulated as follows: there is a random variable X distributed Binomial$(n, 1/2)$ and a $Y = n/2 + \sqrt{n}Z/2$ distributed $N(n/2, n/4)$ such that

$$|X - Y| \leq C + C\frac{|X|^2}{n}, \text{ when } |X| \leq \varepsilon\sqrt{n}$$

for some $C, \varepsilon > 0$. See Massart (2004) for possible explicit values of C and ε, although we don't need them in establishing asymptotic equivalence results. The proof of this bound was first sketched in Komlós, Major, and Tusnády (1975) and detailed in several papers, e.g., Mason and van Zwet (1987), Bretagnolle and Massart (1989), Dudley (2000), Major (2000), Mason (2001), Lawler and Trujillo Ferreras (2005), etc.

Carter and Pollard (2004) improved that classical quantile bounds for Binomial$(n, 1/2)$ with a rate $1/\sqrt{n}$ modulo constants. More specifically, they showed that for the coupling between an X distributed Binomial$(n, 1/2)$ and a $Y = n/2 + \sqrt{n}Z/2$ distributed

$N(n/2, n/4)$,

$$|X - Y| \leq C + C\frac{|X|^3}{n^2}, \text{ when } |X| \leq \varepsilon\sqrt{n}$$

for some $C, \varepsilon > 0$.

The coupling bounds for general random variables and the detailed proofs can be found in Sakhanenko (1984, 1996). In this section, we extend the result of Carter and Pollard (2004) to general symmetric random variables, i.e., sharpens the bound in Sakhanenko (1984, 1996) (or Komlós, Major, and Tusnády (1975)) for the sum of symmetric random variables.

The following proposition is the classical quantile coupling result (cf. Lemma 2 in Sakhanenko (1996) or Lemma 1 in Komlós, Major, and Tusnády (1975)).

Proposition 1 *Let $X_1, X_2, \ldots, X_n$ be i.i.d. random variables such that $EX_1 = 0$, $EX_1^2 = 1$, $E\exp\{t|X_1|\} < \infty$ for some $t > 0$. Let $S_n = \frac{1}{\sqrt{n}}\sum_{i=1}^n X_i$, and Z be a standard normal random variable. Then for every n, there is a random variable $\widetilde{S}_n$ with $\mathcal{L}\left(\widetilde{S}_n\right) = \mathcal{L}(S_n)$ such that*

$$\left|\widetilde{S}_n - Z\right| \leq \frac{C}{\sqrt{n}} + \frac{C}{\sqrt{n}}\left|\widetilde{S}_n\right|^2$$

for $\left|\widetilde{S}_n\right| \leq \varepsilon\sqrt{n}$, where C_1, $\varepsilon > 0$ do not depend on n.

In many practical situations, the random variables are symmetric. We have an improvement on the classical quantile coupling result with a rate $1/\sqrt{n}$ for random variables with continuous distributions.

Theorem 1 *In addition to the assumptions in Proposition 1 suppose that $EX_1^3 = 0$ and the characteristic function $v(t)$ satisfies $\limsup_{|t|\to\infty}|v(t)| < 1$. Then for every n, there is a random variable $\widetilde{S}_n$ with $\mathcal{L}\left(\widetilde{S}_n\right) = \mathcal{L}(S_n)$ such that*

$$\left|\widetilde{S}_n - Z\right| \leq \frac{C}{n} + \frac{C}{n}\left|\widetilde{S}_n\right|^3$$

for $\left|\widetilde{S}_n\right| \leq \varepsilon\sqrt{n}$, where C, $\varepsilon > 0$ do not depend on n.

If the absolutely continuous component of the random variable X_1 is nonzero, the assumption $\limsup_{|t|\to\infty}|v(t)| < 1$ in Theorem 1 is satisfied.

Without that assumption for the characteristic function $v(t)$, we have an improvement on the classical quantile coupling bound with a rate $1/\sqrt{n}$ modulo constants.

Theorem 2 *In addition to the assumptions in Proposition 1 suppose that $EX_1^3 = 0$. Then for every n, there is a random variable $\widetilde{S}_n$ with $\mathcal{L}\left(\widetilde{S}_n\right) = \mathcal{L}(S_n)$ such that*

$$\left|\widetilde{S}_n - Z\right| \leq \frac{C}{\sqrt{n}} + \frac{C}{n}\left|\widetilde{S}_n\right|^3$$

for $\left|\widetilde{S}_n\right| \leq \varepsilon\sqrt{n}$, where C, $\varepsilon > 0$ do not depend on n.

The assumptions of Theorem 2 are satisfied for X_1 =Bernoulli(1/2) − 1/2. Theorem 2 is then a natural extension of Carter and Pollard (2004).

3 Quantile Coupling for Continuous case

In this section, we give a general assumption to obtain a quantile coupling inequality with an improved rate. We then apply this inequality to the sum of independent random variables with vanishing third moment to obtain Theorem 1 which includes the coupling of the sum of symmetric random variables as a special case.

A basic inequality for Mill's ratio will be needed to derive the quantile coupling inequality.

Lemma 1 *For $x > 0$, we have*

$$\frac{\varphi(x)}{\bar{\Phi}(x)} > \min\left\{x, \frac{2}{\sqrt{2\pi}}\right\} \geq \frac{1}{2}\left(x + \frac{2}{\sqrt{2\pi}}\right).$$

The following theorem gives the relationship between the existence of a certain type of large deviation result and a sharp quantile coupling inequality. That type of large deviation is often called "Petrov expansion". Actually, the expansion we use in this paper is even more "precise" than that of Petrov (see Remark 2). Maybe it is better to call it Saulis expansion (see page 249 in Petrov (1975)). Theorem 1 is just an immediate consequence of the following theorem and Proposition 2.

In this paper, we use a notation $O(x)$, which means a value between $-Cx$ and Cx for some $C > 0$.

Theorem 3 *Let Z be a standard normal random variable. Let S_n be a random variable with a distribution function $G(x) = \mathbb{P}(S_n \leq x)$. Assume that there is a positive ε such that for all n,*

$$\begin{aligned}\mathbb{P}(S_n < -x) &= \Phi(-x)\exp\left(O\left(n^{-1}x^4 + n^{-1}\right)\right), \\ 1 - \mathbb{P}(S_n < x) &= \overline{\Phi}(x)\exp\left(O\left(n^{-1}x^4 + n^{-1}\right)\right),\end{aligned}$$

where $\overline{G}(x) = 1 - G(x)$, *and* $\overline{\Phi}(x) = 1 - \Phi(x)$, *and* $O(n^{-1}x^4 + n^{-1})$ *is uniform on the interval* $x \in [0, \varepsilon\sqrt{n}]$. *And the expansion above holds when "$<$" is replaced by "$\leq$". Then for every* n, *there is a random variable* $\widetilde{S}_n$ *with* $\mathcal{L}\left(\widetilde{S}_n\right) = \mathcal{L}(S_n)$ *such that*

$$\left|\widetilde{S}_n - Z\right| \leq \frac{C_1}{n} + \frac{C_1}{n}\left|\widetilde{S}_n\right|^3 \tag{1}$$

for $\left|\widetilde{S}_n\right| \leq \varepsilon_1\sqrt{n}$, *where* $C_1, \varepsilon_1 > 0$ *do not depend on* n.

Remark 1 *The definition of distribution function here is different from that in Petrov (1975), or Major (2000), or Mason (2001), etc. They define* $G(x) = \mathbb{P}(S_n < x)$. *But we use the more standard definition* $G(x) = \mathbb{P}(S_n \leq x)$.

Remark 2 *Let*

$$a(n, x) = n^{-1/2}x^3 + n^{-1/2}x + n^{-1/2}.$$

The Petrov expansion is replacing $O(n^{-1}x^4 + n^{-1})$ *in the Theorem by* $O(a(n, x))$ *(see Theorem 1 in Chapter VIII of Petrov (1975), or Theorem A in Komlós, Major, and Tusnády (1975)). But the corresponding coupling inequality will be*

$$\left|\widetilde{S}_n - Z\right| \leq \frac{C_1}{\sqrt{n}} + \frac{C_1}{\sqrt{n}}\left|\widetilde{S}_n\right|^2$$

(see Sakhanenko (1984, 1996)). The deviation term $O(n^{-1}x^4 + n^{-1})$ *improves* $O(a(n, x))$ *with a rate* $1/\sqrt{n}$ *for* x *in a constant level, so is the corresponding quantile coupling inequality.*

The following is a detailed proof of Theorem 3. It is a modification of the proof for the classical case, which was sketched in Komlós, Major, and Tusnády (1975).

Proof: Define

$$\widetilde{S}_n = G^{-1}\Phi(Z) \tag{2}$$

where

$$G^{-1}(x) = \inf\{u, G(u) \geq x\},$$

such that $\mathcal{L}\left(\widetilde{S}_n\right) = \mathcal{L}(S_n)$. Without loss of generality, we assume that $0 \leq \widetilde{S}_n \leq \varepsilon\sqrt{n}$, because the derivation for $-\varepsilon\sqrt{n} \leq \widetilde{S}_n \leq 0$ is similar. The equation (1) is equivalent to

$$-C_1\frac{1}{n}\left(1 + \left|\widetilde{S}_n\right|^3\right) \leq \widetilde{S}_n - Z \leq C_1\frac{1}{n}\left(1 + \left|\widetilde{S}_n\right|^3\right)$$

i.e.,

$$\Phi\left(\widetilde{S}_n - C_1\frac{1}{n}\left(1+\left|\widetilde{S}_n\right|^3\right)\right) \le \Phi(Z) \le \Phi\left(\widetilde{S}_n + C_1\frac{1}{n}\left(1+\left|\widetilde{S}_n\right|^3\right)\right).$$

Define $G\left(\widetilde{S}_n-\right) = \mathbb{P}(S_n < x)$. From the definition of $\widetilde{S}_n$ in (2) we have $G\left(\widetilde{S}_n-\right) \le \Phi(Z) \le G\left(\widetilde{S}_n\right)$, then we need only to show

$$\begin{aligned} &\Phi\left(\widetilde{S}_n - C_1\frac{1}{n}\left(1+\left|\widetilde{S}_n\right|^3\right)\right) \\ \le\ & G\left(\widetilde{S}_n-\right) \le G\left(\widetilde{S}_n\right) \le \Phi\left(\widetilde{S}_n + C_1\frac{1}{n}\left(1+\left|\widetilde{S}_n\right|^3\right)\right). \end{aligned}$$

i.e.

$$\begin{aligned} &\log\left(\frac{1-\Phi\left(x - C_1\frac{1}{n}\left(1+x^3\right)\right)}{1-\Phi(x)}\right) \\ \ge\ & \log\frac{1-G(x-)}{1-\Phi(x)} \ge \log\frac{1-G(x)}{1-\Phi(x)} \\ \ge\ & \log\left(\frac{1-\Phi\left(x + C_1\frac{1}{n}\left(1+x^3\right)\right)}{1-\Phi(x)}\right) \end{aligned}$$

when $0 \le x \le \varepsilon\sqrt{n}$. From the assumption in the theorem, we know

$$\max\left\{\left|\log\frac{1-G(x)}{1-\Phi(x)}\right|, \left|\log\frac{1-G(x-)}{1-\Phi(x)}\right|\right\} \le C\left(n^{-1}x^4 + n^{-1}\right)$$

for some $C > 0$. Thus it is enough to show there is $C_1 > 0$ such that

$$\begin{aligned} \log\left(\frac{1-\Phi\left(x - C_1\frac{1}{n}\left(1+x^3\right)\right)}{1-\Phi(x)}\right) \ &\ge\ C\left(n^{-1}x^4 + n^{-1}\right) \\ &\ge\ \log\left(\frac{1-\Phi\left(x + C_1\frac{1}{n}\left(1+x^3\right)\right)}{1-\Phi(x)}\right) \end{aligned} \tag{3}$$

We only show the first part of the inequality above due to the symmetry of the equation.

It is easy to see that the first part of the equation above is satisfied under the condition $x - C_1\frac{1}{2n}\left(1+|x|^3\right) \le 0$ (we will see later the value of C_1 can be specified as $18\sqrt{2\pi}C$). It implies $x \le C_1/n \le 1$ for n sufficiently large under an assumption that $C_1\varepsilon^2 \le 1$, which holds choosing sufficiently small ε. Then for $0 \le x \le C_1/n \le 1$ and n sufficiently large, we have

$$\begin{aligned}
\log\left(\frac{1-\Phi\left(x-C_1\frac{1}{n}\left(1+x^3\right)\right)}{1-\Phi\left(x\right)}\right) &\geq \log\left(\frac{1-\Phi\left(x-C_1\frac{1}{n}\left(1+x^3\right)\right)}{1-\Phi\left(0\right)}\right) \\
&= \log\left(1+\left[1-2\Phi\left(x-C_1\frac{1}{n}\left(1+x^3\right)\right)\right]\right) \\
&\geq \frac{1}{2}-\Phi\left(x-C_1\frac{1}{n}\left(1+x^3\right)\right),
\end{aligned}$$

where the last inequality follows from the fact $\log\left(1+y\right)\geq y/2$ for $0\leq y\leq 1$. Write

$$\frac{1}{2}-\Phi\left(x-C_1\frac{1}{n}\left(1+x^3\right)\right)=\Phi\left(C_1\frac{1}{n}\left(1+x^3\right)-x\right)-\Phi\left(0\right).$$

Since $\left|C_1\frac{1}{n}\left(1+x^3\right)\right|\leq 2$ and $\varphi\left(x\right)\geq\frac{1}{9\sqrt{2\pi}}$ for $0\leq x\leq 2$, the intermediate value theorem implies

$$\begin{aligned}
\Phi\left(C_1\frac{1}{n}\left(1+x^3\right)-x\right)-\Phi\left(0\right) &\geq \frac{1}{9\sqrt{2\pi}}\left(C_1\frac{1}{n}\left(1+x^3\right)-x\right) \\
&\geq \frac{1}{9\sqrt{2\pi}}\cdot\frac{C_1}{2n}\left(1+x^3\right)\geq\frac{C_1}{18\sqrt{2\pi}}\left(n^{-1}x^4+n^{-1}\right)
\end{aligned}$$

which is more than $C\left(n^{-1}x^4+n^{-1}\right)$ when $C_1\geq 18\sqrt{2\pi}C$. Thus the equation (3) is established in the case of $x-C_1\frac{1}{2n}\left(1+\left|x\right|^3\right)\leq 0$.

Now we consider the case $x-C_1\frac{1}{2n}\left(1+\left|x\right|^3\right)\geq 0$. The intermediate value theorem tells us there is a number ξ between x and $x-\frac{C_1}{4n}\left(1+x^3\right)$ such that

$$\begin{aligned}
&\log\left(\frac{1-\Phi\left(x-C_1\frac{1}{n}\left(1+x^3\right)\right)}{1-\Phi\left(x\right)}\right) \\
\geq\ &\log\left(\frac{1-\Phi\left(x-\frac{C_1}{4n}\left(1+x^3\right)\right)}{1-\Phi\left(x\right)}\right) \\
=\ &\frac{C_1}{4}\frac{1}{n}\left(1+x^3\right)\frac{\varphi\left(\xi\right)}{1-\Phi\left(\xi\right)}.
\end{aligned}$$

From the lemma (1), we have

$$\begin{aligned}
&\log\left(\frac{1-\Phi\left(x-\frac{C_1}{4n}\left(1+x^3\right)\right)}{1-\Phi\left(x\right)}\right) \\
\geq\ &\frac{C_1}{4n}\left(1+x^3\right)\cdot\frac{1}{2}\left(x-\frac{C_1}{4n}\left(1+x^3\right)+\frac{2}{\sqrt{2\pi}}\right) \\
\geq\ &\frac{C_1}{4}\frac{1}{n}\left(1+x^3\right)\cdot\frac{1}{2}\left(\frac{x}{2}+\frac{2}{\sqrt{2\pi}}\right) \\
\geq\ &\frac{C}{n}x^4+\frac{C}{n}.
\end{aligned}$$

when $C_1 \geq 16C$.

Putting all together, we establish (3) and prove the theorem. ∎

In some applications, it is more convenient to use the following corollary. The bound involves only the normal random variable. In Zhou (2004), we used the coupling of Beta distribution with a normal to establish asymptotic equivalence of Gaussian variance regression and Gaussian white noise with a drift, and we found that it was much easier to use the following bound in moments calculations.

Corollary 1 *Under the assumption of Theorem 3, for every n there is a random variable* $\widetilde{S}_n$ *with* $\mathcal{L}\left(\widetilde{S}_n\right) = \mathcal{L}(S_n)$ *such that*

$$\left|\widetilde{S}_n - Z\right| \leq \frac{C}{n}\left(1 + |Z|^3\right), \text{ when } |Z| \leq \varepsilon\sqrt{n}$$

for some $C, \varepsilon > 0$.

Proof: Obviously the inequality (1) still holds, when $\left|\widetilde{S}_n\right| \leq \varepsilon_1\sqrt{n}$ for $0 < \varepsilon_1 \leq \varepsilon$. Let's choose ε_1 small enough such that $C\varepsilon_1^2 < 1/2$. When $\left|\widetilde{S}_n\right| \leq \varepsilon_1\sqrt{n}$, we have

$$\left|\widetilde{S}_n - Z\right| \leq \frac{C}{n} + \frac{1}{2}\left|\widetilde{S}_n\right|,$$

from (1), which implies

$$\left|\widetilde{S}_n\right| - |Z| \leq \frac{C}{n} + \frac{1}{2}\left|\widetilde{S}_n\right|$$

by the triangle inequality, i.e.,

$$\left|\widetilde{S}_n\right| \leq \frac{2C}{n} + 2|Z|, \tag{4}$$

so we have

$$\left|\widetilde{S}_n - Z\right| \leq \frac{C}{n} + \frac{C}{n}\left(\frac{2C}{n} + 2|Z|\right)^3 \leq \frac{C_1}{n}\left(1 + |Z|^3\right)$$

for some $C_1 > 0$.

When $\widetilde{S}_n = \varepsilon_1\sqrt{n} > 0$ for any ε_1 with $0 < \varepsilon_1 \leq \varepsilon$, we know $Z \geq 0$ from the definition of quantile coupling, and from (4) we have

$$Z \geq \varepsilon_1\sqrt{n} - \frac{2C}{n}.$$

In the definition of quantile coupling, we see that $\widetilde{S}_n$ is an increasing function of Z. So we have $\widetilde{S}_n \leq \varepsilon_1 n$, when $Z \leq \varepsilon_1\sqrt{n} - \frac{2C}{n}$. Similarly we may show $\widetilde{S}_n \geq -\varepsilon_1 n$, when $Z \geq -\varepsilon_1\sqrt{n} + \frac{2C}{n}$. Thus we have

$$\left|\widetilde{S}_n\right| \leq \varepsilon_1\sqrt{n}, \text{ when } |Z| \leq \varepsilon_1 n - \frac{2C}{\sqrt{n}}. \tag{5}$$

Let $\varepsilon_2 = \varepsilon_1/2$. We have $\varepsilon_2\sqrt{n} < \varepsilon_1\sqrt{n} - \frac{2C}{n}$ for $n > \left(\frac{2C}{\varepsilon_2}\right)^{2/3}$, then we know

$$\left\{|Z| \le \varepsilon_2\sqrt{n}\right\} \subset \left\{ \ |Z| \le \varepsilon_1\sqrt{n} - \frac{2C}{n} \right\} \subset \left\{\left|\widetilde{S}_n\right| \le \varepsilon_1\sqrt{n}\right\}$$

from (5), so we have

$$\left|\widetilde{S}_n - Z\right| \le \frac{C}{n}\left(1 + |Z|^3\right), \text{ when } |Z| \le \varepsilon_2\sqrt{n} \text{ and } n > \left(\frac{2C}{\varepsilon_2}\right)^{2/3}.$$

Thus we have

$$\left|\widetilde{S}_n - Z\right| \le \frac{C}{n}\left(1 + |Z|^3\right), \text{ when } |Z| \le \varepsilon_2\sqrt{n}.\blacksquare$$

An application of Theorem 3 and Corollary 1 is the coupling of the sum of independent random variables with a normal random variable. Assume that those random variables have finite exponential moment and vanishing third moment (e.g. symmetric random variable). The following is the Saulis expansion (See page 249 in Petrov (1975), or page 188 in Saulis and Statulevicius (1991)), which gives a sharp approximation to the tail probability of the sum of those random variables. The proof of the this result can also be derived based on similar arguments in Section 8.2 in Petrov (1975).

Proposition 2 *Let $X_1, X_2, \ldots, X_n$ be i.i.d. random variables for which*

$$\begin{aligned} EX_1 &= 0, \ EX_1^2 = 1, \ EX_1^3 = 0, \\ E\exp\left(a\left|X_1\right|\right) &< \infty \text{ for some } a > 0, \end{aligned}$$

and

$$\limsup_{|t|\to\infty} \exp E\left(itX_1\right) < 1.$$

Define $\widetilde{S}_n = \frac{1}{\sqrt{n}}\sum_{i=1}^n X_i$. Then their exists positive constants C and η such that

$$\begin{aligned} \mathbb{P}\left(\widetilde{S}_n < -x\right) &= \Phi\left(-x\right)\exp\left(Cn^{-1}x^4 + Cn^{-1}\right) \\ 1 - \mathbb{P}\left(S_n < x\right) &= \overline{\Phi}\left(-x\right)\exp\left(-Cn^{-1}x^4 - Cn^{-1}\right) \end{aligned}$$

in the interval $0 \le x \le \eta$.

Note that the expansion above holds for $Y_i = -X_i$. This implies the expansion above holds when "$<$" is replaced by "$\le$". This proposition and Theorem 3 immediately imply the following corollary and Theorem 1.

Corollary 2 *Under the assumption in Proposition 2, for every* n*, there is a random variable* $\widetilde{S}_n$ *with* $\mathcal{L}\left(\widetilde{S}_n\right) = \mathcal{L}\left(\widetilde{S}_n\right)$ *such that*

$$\left|\widetilde{S}_n - Z\right| \le \min\left\{\frac{C}{n} + \frac{C}{n}\left|\widetilde{S}_n\right|^3, \frac{C}{n} + \frac{C}{n}|Z|^3\right\}$$

for $\left|\widetilde{S}_n\right| \le \varepsilon\sqrt{n}$ *or* $|Z| \le \varepsilon\sqrt{n}$ *, where* C*,* $\varepsilon > 0$ *do not depend on* n*.*

4 Quantile Coupling for General Case

In this section, we give a general assumption to obtain a quantile coupling inequality with an improved rate modulo constants. One application of the result is sharpening classical quantile coupling inequality with a rate modulo constants for the sum of independent symmetric random variables. So this result is a generalization of Carter and Pollard (2004), where they considered coupling for Binomial$(n, 1/2)$.

The following theorem and Lemma 2 imply Theorem 2.

Theorem 4 *Let* Z *be a standard normal random variable. Let* $\widetilde{S}_n$ *be a random variable with a distribution function* $G(x)$*. Assume that there is a positive* ε *such that for all* n*,*

$$\begin{aligned} \mathbb{P}(S_n < -x) &= \Phi(-x)\exp\left(O\left(n^{-1}x^4 + n^{-1/2}\right)\right), \\ 1 - \mathbb{P}(S_n < x) &= \overline{\Phi}(x)\exp\left(O\left(n^{-1}x^4 + n^{-1/2}\right)\right), \end{aligned}$$

where $\overline{G}(x) = 1 - G(x)$*, and* $\overline{\Phi}(x) = 1 - \Phi(x)$*, and* $O\left(n^{-1}x^4 + n^{-1/2}\right)$ *is uniform on the interval* $x \in [0, \varepsilon\sqrt{n}]$ *with* $\varepsilon > 0$*. And the expansion above holds when "$<$" is replaced by "$\le$". Then for every* n*, there is a random variable* $\widetilde{S}_n$ *with* $\mathcal{L}\left(\widetilde{S}_n\right) = \mathcal{L}(S_n)$ *such that*

$$\left|\widetilde{S}_n - Z\right| \le \frac{C_1}{\sqrt{n}} + \frac{C_1}{n}\left|\widetilde{S}_n\right|^3$$

for $\left|\widetilde{S}_n\right| \le \varepsilon_1\sqrt{n}$*, where* C_1*,* $\varepsilon_1 > 0$ *do not depend on* n*.*

The proof of Theorem 4 is similar to that of Theorem 3, so we skip the proof.

Similar to the proof of Corollary 1, we have

Corollary 3 *Under the assumption of Theorem 4, for every* n *there is a random variable* $\widetilde{S}_n$ *with* $\mathcal{L}\left(\widetilde{S}_n\right) = \mathcal{L}(S_n)$ *such that*

$$\left|\widetilde{S}_n - Z\right| \le \frac{C}{\sqrt{n}} + \frac{C}{n}|Z|^3, \text{ when } |Z| \le \varepsilon\sqrt{n}$$

where C*,* $\varepsilon > 0$ *do not depend on* n*.*

An application of Theorem 4 and Corollary 3 is the coupling of the sum of independent random variables with a normal random variable. Assume that those random variables have finite exponential moment and vanishing third moment (e.g. symmetric random variable). An approximation to the tail probability of the sum of those random variables is given in the following lemma. The proof of the approximation is based on similar arguments in Section 8.2 in Petrov (1975). It is an extension of Theorem 1 in Carter and Pollard (2004).

Lemma 2 *Let X_1, X_2, ..., X_n be i.i.d. random variables for which*

$$\begin{aligned} EX_1 &= 0,\ EX_1^2 = 1,\ EX_1^3 = 0, \\ E\exp\left(a\left|X_1\right|\right) &< \infty \text{ for some } a > 0. \end{aligned}$$

Then their exists positive constants ε such that

$$\mathbb{P}\left(S_n < -x\right) = \Phi\left(-x\right)\exp\left(O\left(n^{-1}x^4 + n^{-1/2}\right)\right) \tag{6}$$

$$1 - \mathbb{P}\left(S_n < x\right) = \overline{\Phi}\left(x\right)\exp\left(O\left(n^{-1}x^4 + n^{-1/2}\right)\right) \tag{7}$$

in the interval $0 \le x \le \varepsilon$, where

$$S_n = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} X_i,$$

and the expansion above holds when "$<$" is replaced by "$\le$".

Proof: From Theorem 2 of Section 8.2 in Petrov (1975), we know

$$\mathbb{P}\left(S_n < -x\right) = \Phi\left(-x\right)\exp\left(Cn^{-1}x^4 + C\left(x+1\right)n^{-1/2}\right)$$

(Our notation is different from that of Petrov. Our x here is their $z = x/\sqrt{n}$). Under the assumption of $EX_1^3 = 0$, we can replace the terms $1 + O\left(z\right)$ of equations (2.37) and (2.38) in Section 8.2 in Petrov (1975) by $1 + O\left(z^2\right)$. In the same section, from equation (2.35) we can replace the term $1 + O\left(z\right)$ of equation (2.40) by $1 + C/\sqrt{n}$. We keep everything else in the proof Theorem 2 of Section 8.2 in Petrov (1975). Then we establish the following approximation of the tail probability

$$\mathbb{P}\left(S_n < -x\right) = \Phi\left(-x\right)\exp\left(Cn^{-1}x^4 + C\left(x^2/n + n^{-1/2}\right)\right)$$

Note that $x^2/n \le \frac{1}{2}\left(n^{-1}x^4 + n^{-1}\right) \le \frac{1}{2}\left(n^{-1}x^4 + n^{-1/2}\right)$. Then we obtain equation (6). The argument for equation (7) is similar.

The expansion holds if replacing X_i by $-X_i$. This implies the expansion above holds when "$<$" is replaced by "$\le$".∎

Remark 3 *In Lemma 2, we assume that those random variables are identically distributed, but it can be extended to non-identically case similar to Theorem 2 of Section 8.2 in Petrov (1975).*

Theorem 4 and Lemma 2 imply the following corollary (or basically Theorem 2). It extends Theorem 2 of Carter and Pollard (2004).

Corollary 4 *Under the assumption in Lemma 2, for every n, there is a random variable $\widetilde{S}_n$ with $\mathcal{L}\left(\widetilde{S}_n\right) = \mathcal{L}\left(S_n\right)$ such that*

$$\left|\widetilde{S}_n - Z\right| \leq \min\left\{\frac{C}{\sqrt{n}} + \frac{C}{n}\left|\widetilde{S}_n\right|^3, \frac{C}{\sqrt{n}} + \frac{C}{n}\left|Z\right|^3\right\}$$

for $\left|\widetilde{S}_n\right| \leq \varepsilon\sqrt{n}$ or $|Z| \leq \varepsilon\sqrt{n}$, where C, $\varepsilon > 0$ do not depend on n.

5 Some Examples

In this section, we discuss some applications of results in previous sections to asymptotic equivalence theory and nonparametric function estimation.

Example 1: Asymptotic equivalence of density estimation and Gaussian white noise:

$$\begin{aligned}\mathbb{E}_n &: y(1), ..., y(n), \text{ i.i.d. with density } f \text{ on } [0,1] \\ \mathbb{F}_n &: dy_t = f^{1/2}(t)\,dt + \frac{1}{2}n^{-1/2}dW_t\end{aligned}$$

The asymptotic equivalence result above was established in Brown, Low, Carter and Zhang (2004) under a Besov smoothness constraint. The key idea of that paper is applying the classical KMT construction. We then need a coupling for Binomial random variable and a normal random variable. Let X_1, X_2, $\ldots$, X_n be i.i.d. Bernoulli(1/2). Then Corollary 4 tells us for every n there is a random variable $\widetilde{S}_n$ with $L\left(\widetilde{S}_n\right) = L\left(\widetilde{S}_n\right)$ such that

$$\left|\widetilde{S}_n - Z\right| \leq \min\left\{\frac{C}{\sqrt{n}} + \frac{C}{n}\left|\widetilde{S}_n\right|^3, \frac{C}{\sqrt{n}} + \frac{C}{n}\left|Z\right|^3\right\}$$

for $\left|\widetilde{S}_n\right| \leq \varepsilon\sqrt{n}$ or $|Z| \leq \varepsilon\sqrt{n}$, where C, $\varepsilon > 0$ do not depend on n (see also Carter and Pollard (2004)). This result was used in the KMT construction to establish the asymptotic equivalence under the Besov smoothness condition, compact in Besov balls of $B^{1/2}_{2,2}$ and $B^{1/2}_{4,4}$. If one applies the classical Tusnády's inequality, a stronger smoothness condition would be needed to establish the asymptotic equivalence.

Example 2: Asymptotic equivalence of spectral density estimation and Gaussian white noise:

$$\mathbb{E}_n \ : \ y(1), ..., y(n), \text{ a stationary centered Gaussian sequence with spectral density } f$$
$$\mathbb{F}_n \ : \ dy_t = \log f(t)\, dt + 2\pi^{1/2} n^{-1/2} dW_t$$

where f has support on $[-\pi, \pi]$. This asymptotic equivalence between Gaussian spectral density, Gaussian variance regression and Gaussian white noise in Golubev, Nussbaum and Zhou(2005) under a Besov smoothness constraint. In that paper, we used a dyadic KMT-type construction, but different from the classical KMT construction. In the KMT paper, they used a complicate conditional quantile coupling for higher resolutions. It is easy to observe that $L(X|X+Y) = L((X+Y)B_n)$ for two independent and identically distributed random variables X and Y with law χ_n^2, then we can avoid the conditional quantile coupling by considering the coupling for a Beta random variable. The following coupling inequality is then used. Let Z be a standard normal random variable. For every n, there is a mapping $T_n : R \mapsto R$ such that the random variable $B_n = T_n(Z)$ has the Beta $(n/2, n/2)$ law and

$$\left| n\left(1/2 - B_n\right) - \frac{n^{1/2}}{2} Z \right| \leq \min \left\{ \frac{C}{\sqrt{n}} + \frac{C}{n^2} \left| nB_n - n/2 \right|^3, \frac{C}{n} + \frac{C}{n} \left| Z \right|^3 \right\}$$

for $|nB_n - n/2| \leq \varepsilon n$, where C, $\varepsilon > 0$ do not depend on n (cf. Zhou (2004)).

Example 3: Quantile coupling of Median statistics. Let $X_1, X_2, \ldots, X_n$ i.i.d. with density $f(x)$. For simplicity, let $n = 2k + 1$ with some integer $k \geq 1$, and assume that

$$f(0) > 0,\ f'(0) = 0, \text{ and } f \in C^3.$$

Let Z be a standard normal random variable. For every n, there is a mapping $T_n : R \mapsto R$ such that the random variable $X_{med} = T_n(Z)$ has density $f(x)$ and

$$\left| \sqrt{4n} f(0)\, X_{med} - Z \right| \leq C \frac{1}{n} \left(1 + |Z|^3\right), \text{ when } |Z| \leq \varepsilon \sqrt{n}$$

where C, $\varepsilon > 0$ do not depend on n. Details and more general discussions will be presented in Brown, Cai and Zhou (working paper). In this paper, we apply this quantile coupling bound to nonparametric location model with Cauchy noise and consider wavelet regression. Donoho and Yu (2000) considered a similar problem, but minimax property is unclear for their procedure. In wavelet regression setting, Hall and Patil (1996) studied nonparametric location models and achieved the optimal minimax rate, but under an assumption of the existence of finite forth moment. We don't need any moment condition, and the noise can be general and unknown, but achieve optimal minimax rate of convergence. Without the assumption of $f'(0) = 0$ or $f \in C^3$, we may still obtain coupling bounds, but may not as be tight as the bound above. The tightness of the upper bound affects the the underlying smoothness condition we need in deriving asymptotic properties.

References

[1] Bretagnolle, J. and Massart, P. (1989). Hungarian constructions from the nonasymptotic view point, *Ann. Probab.* **17** (1) 239–256.

[2] Brown, L.D., Cai, T. T., Zhou, H. H.. Robust Nonparametric Wavelet Estimation. In preparation.

[3] Brown, L.D., Carter, A.V., Low, M.G. and Zhang, C. (2004). Equivalence theory for density estimation, Poisson processes and Gaussian white noise with drift. *Ann. Statist.* **32**, 2074-2097.

[4] Carter, A. V. and Pollard, D. (2004). Tusnády's inequality revisited. *Ann. Statist.* **32**, 2731-2741

[5] Donoho, D. L. and Johnstone, I. M. (1995). Adapt to unknown smoothness via wavelet shrinkage. *J. Amer. Stat. Assoc.* **90** , 1200-1224.

[6] Donoho, D. L. and Yu , T. P.-Y. (2000). Nonlinear Pyramid Transforms Based on Median-Interpolation. SIAM Journal of Math. Anal., 31(5), 1030-1061.

[7] Dudley, R. M. (2000). Notes on empirical processes. Lecture notes for a course given at Aarhus Univ., August 1999.

[8] Golubev, G. K., Nussbaum, M. and Zhou, H. H., Asymptotic equivalence of spectral density estimation and Gaussian white noise. Available at http://www.stat.yale.edu/~hz68. Submitted.

[9] Grama, I. and Nussbaum, M. (1998). Asymptotic equivalence for nonparametric generalized linear models. *Probab. Theory Relat. Fields* **111** 167-214

[10] Grama, I and Nussbaum, M. (2002). Asymptotic equivalence for nonparametric regression. *Mathematical Methods of Statistics*, **11** (1), 1-36.

[11] P. Hall and P. Patil (1996). On the choice of smoothing parameter, threshold and truncation in nonparametric regression by wavelet methods, *J. Roy. Statist. Soc. Ser. B*, **58**, 361–377.

[12] Komlös, J., Major, P. and Tusnády, G. (1975). An approximation of partial sums of independent rv's and the sample df. I Z. *Wahrsch. verw. Gebiete* **32** 111-131.

[13] Komlös, J., Major, P. and Tusnády, G. (1976). An approximation of partial sums of independent rv's and the sample df. II Z. *Wahrsch. verw. Gebiete* **34** 33-58.

[14] Lawer, G, F. and Trujillo Ferreras, J. A.. Random walk loop soup. Available at http://www.math.cornell.edu/%7Elawler/papers.html

[15] Major, P. (2000). The approximation of the normalized empirical ditribution function by a Brownian bridge. Technical report, Mathematical Institute of the Hungarian Academy of Sciences. Notes available from www.renyi.hu/~major/.

[16] Mason, D. M. (2001). Notes on the KMT Brownian bridge approximation to the uniform empirical process. *In Asymptotic Methods in Probability and Statistics with Applications* (N. Balakrishnan, I. A. Ibragimov and V. B. Nevzorov, eds.) 351–369. Birkhäuser, Boston.

[17] Massart, P. (2002). Tusnády's lemma, 24 years later. *Ann. Inst. H. Poincaré Probab. Statist.* **38** 991–1007.

[18] Nussbaum, M. (1996). Asymptotic equivalence of density estimation and Gaussian white noise. *Ann. Statist.* **24**, 2399-2430

[19] Petrov, V. V. (1975). *Sums of Independent Random Variables.* Springer-Verlag. (Enlish translation from 1972 Russian edition).

[20] Sakhanenko, A. (1984). The rate of convergence in the invariance principle for non-identically distributed variables with exponential moments. Limit theorems for sums of random variables. *Trudy Inst. Matem., Sibirsk. Otdel. AN SSSR.* **3** 3-49 (in Russian).

[21] Sakhanenko, A. I. (1996). Estimates for the Accuracy of Coupling in the Central Limit Theorem. *Siberian Mathematical Journal,* Vol. 37. No 4.

[22] Saulis, L and Statulevicius, V.A. (1991). *Limit theorems for large deviations.* Kluwer Academic Publishers.

[23] Zhou, H. H. (2004). *Minimax Estimation with Thresholding and Asymptotic Equivalence for Gaussian Variance Regression.* Ph.D. Dissertation. Cornell University, Ithaca, NY.