Non-uniform Berry-Esseen bound by unbounded exchangeable pairs approach

LIU Da-li[1]   LI Zheng[2]   WANG Han-chao[3,*]   CHEN Zeng-jing[4]

Abstract. In this paper, a new technique is introduced to obtain non-uniform Berry-Esseen bounds for normal and nonnormal approximations by unbounded exchangeable pairs. This technique does not rely on the concentration inequalities developed by Chen and Shao [4,5] and can be applied to the quadratic forms and the general Curie-Weiss model.

§1   Introduction

Since Charles Stein presented his ideas in the seminal paper [15], there have been a lot of research activities around Stein's method. Stein's method is a powerful tool to obtain the approximate error of normal and non-normal approximations. The readers can refer to Chatterjee [2] for recent developments of Stein's method.

While several works on Stein's method pay attention to the uniform error bounds, Stein's method showed to be powerful on the non-uniform error bounds, too. By Stein's method, Chen and Shao [4,5] obtained non-uniform Berry-Esseen bounds for independent or locally dependent random variables. The key point in their works is the concentration inequality, which also has strong connection with another approach called the *exchangeable pairs approach*.

The exchangeable pairs approach turned out to be an important topic within Stein's method. Let W be the random variable under study. The pair (W, W') is called an exchangeable pair if (W, W') and (W', W) share the same distribution. With $\Delta = W - W'$, Rinott and Rotar [12], Shao and Su [11] obtained a Berry-Esseen bound of the normal approximation when Δ is bounded. If Δ is unbounded, Chen and Shao [6] provided a Berry-Esseen bound and got the optimal rate for an independence test. The concentration inequality plays a crucial role

in previous studies, such as Shao and Su [11] , Chen and Shao [6]. Recently, Shao and Zhang [13] made a big step for unbounded Δ and without using the concentration inequality. They obtained a simple bound as seen from the following result.

Theorem 1. *(Shao and Zhang [13]) Let (W, W') be an exchangeable pair, $\Delta = W - W'$, and the relation*

$$E(\Delta|W) = \lambda(W + R), a.s.,$$

holds for some constant $\lambda \in (0, 1)$ and a random variable R . Then,

$$\sup_{z \in \mathbb{R}} \left| P(W \leq z) - \Phi(z) \right|$$

$$\leq E\left|1 - \frac{1}{2\lambda} E(\Delta^2|W)\right| + \frac{1}{\lambda} E\left|E\left(\Delta\Delta^*|W\right)\right| + E|R|, \tag{1}$$

where $\Phi(z)$, $z \in \mathbb{R}$, is the standard normal ditribution function, $\Delta^*(W, W')$ is a random variable satisfying $\Delta^*(W, W') = \Delta^*(W', W)$ and $\Delta^* \geq |\Delta|$, a.s. .

In this paper, inspired by the idea of Shao and Zhang [13], we extend their results and get a non-uniform Berry-Esseen bound for unbounded exchangeable pairs by combining new techniques. In addition, Chatterjee and Shao [3] introduced a new approach for non-normal approximation by Stein's method in the case of bounded Δ. When Δ is unbounded, Shao and Zhang [13] obtained Berry-Esseen bounds for non-normal approximation. In this paper, we extend their result to the non-uniform case.

The main contribution of this paper is threefold. First, we introduce a new technique to obtain non-uniform Berry-Esseen bounds for unbounded exchangeable pairs. Our proof does not rely on the concentration inequality. Second, we present a non-uniform Berry-Esseen bound for non-normal approximation. As far as we know, there are only a few results in this area. Shao, Zhang and Zhang [14] obtained a Cramér-type moderate deviation for non-normal approximation. At last, we apply our results to quadratic forms and the general Curie-Weiss model.

The paper is organized as follows. We present the main result in Section 2. We give some technical lemmas and the proof of the main result in Section 3. The applications of our result are collected in Section 4.

§2   Main result

In this section, we present some notions and notations about Stein's method. Further details can be found in Shao and Zhang [13]. We then state our main result.

Let the function $g(x), x \in \mathbb{R}$, of the class $\mathcal{C}^2$, satisfy the following conditions:

(A1) $g(x)$ is non-decreasing, and $xg(x) \geq 0$ for $x \in \mathbb{R}$;

(A2) $g'(x)$ is abusolutely continuous and $2(g'(x))^2 - g(x)g''(x) \geq 0$ for all $x \in \mathbb{R}$;

(A3) $\lim_{x\downarrow-\infty} g(x)p(x) = 0$ and $\lim_{x\uparrow+\infty} g(x)p(x) = 0$, where

$$p(x) = c_1 e^{-G(x)}, \quad G(x) = \int_0^x g(t)dt, \ x \in \mathbb{R}, \tag{2}$$

and c_1 is the constant such that $\int_{-\infty}^{+\infty} p(x)dx = 1$.

Let us note that if $g(x) = x$, then $p(x), x \in \mathbb{R}$, is the standard normal density function.

Let $F(z), z \in \mathbb{R}$, be the distribution function whose density function is $p(z)$ as defined in (2). For a fixed $z \in \mathbb{R}$, let $f_z(x)$ denote the solution of Stein's equation, here and below $f_z'(x) = \frac{d}{dx} f_z(x)$:

$$f_z'(x) - g(x)f_z(x) = I(x \le z) - F(z), \ x \in \mathbb{R},$$

$I(\cdot)$ is the indicator function.
By Chatterjee and Shao [3],

$$f_z(x) = \begin{cases} \dfrac{F(x)(1 - F(z))}{p(x)}, x \le z, \\ \dfrac{F(z)(1 - F(x))}{p(x)}, x > z. \end{cases}$$

From Shao and Zhang [13], we know that if (A1)$\sim$(A3) hold, then, $f_z(x)$ has the following properties, for any fixed $z \in \mathbb{R}$:

(B1) $0 \le f_z(x) \le \dfrac{1}{c_1}$, $x \in \mathbb{R}$;

(B2) $\|f_z'\| \le 1$ ($\|\cdot\|$ is the sup norm.) ;

(B3) $F(z) - 1 \le g(x)f_z(x) \le F(z)$;

(B4) $g(x)f_z(x)$ is non-decreasing in x.

For a random variable W, applying Stein's equation to it and taking expectation on both sides, we have:

$$P(W \le z) - F(z) = Ef_z'(W) - Eg(W)f_z(W), \quad z \in \mathbb{R}.$$

Before presenting our main result, we introduce another condition we want $g(x)$ to satisfy:

(A4) There is a number $\tau \in (0, 1)$ and a positive constant K_τ such that

$$\frac{g(x)}{g(\tau x)} \le K_\tau, \text{ for all } x \in \mathbb{R}.$$

There is a large class of functions g satisfying condition $(A4)$, besides the conditions (A1)$\sim$(A3). A typical example is $g(x) = sgn(x)|x|^\alpha, \alpha \ge 1$ (α is a real number).

Let $\mathbf{X}$ be a random vector in $\mathbb{R}^n$ and $W = \varphi(\mathbf{X})$ the random variable of interest (φ is a certain measurable function). Denote by $F(z)$, $z \in \mathbb{R}$, the distribution function whose density function is defined by (2). Now we present our main result.

Theorem 2. *Let (W, W') be an exchangeable pair, $\Delta = W - W'$, and let the following relation be satisfied*

$$E(\Delta|\mathbf{X}) = \lambda(g(W) + R) \quad a.s., \tag{3}$$

for some constant $\lambda \in (0, 1)$ and a random variable R. Assume that $g(x)$, $x \in \mathbb{R}$, satisfies (A1) $\sim$(A4) and $Eg^2(W) < \infty$. Then, for any $z \in \mathbb{R}$,

$$\left| P(W \leq z) - F(z) \right| \leq \frac{C}{1 + |g(z)|} \left\{ \sqrt{E \left| (1 - \frac{1}{2\lambda} E(\Delta^2 | \boldsymbol{X})) \right|^2} + \frac{1}{\lambda} \sqrt{E|E(\Delta\Delta^* | \boldsymbol{X})|^2} + E|R| \right\}.$$

$$(4)$$

Here C is a constant depending on τ and $Eg^2(W)$, Δ^* is a random variable such that $\Delta^*(W, W') = \Delta^*(W', W)$ and $\Delta^* \geq |\Delta|$ a.s..

Remark 1. *Shao and Zhang [13] provided the Berry-Esseen bound for non-normal approximation similar to (1). Theorem 2 is a non-uniform refinement of their result.*

Remark 2. *Let $W = \sum\limits_{i=1}^{n} X_i$, where $\{X_i, i = 1, ..., n\}$ are independent random variables with zero mean and $EW^2 = 1$. Our general result (4) cannot directly cover the following classical result in Chen and Shao [?]: there is an absolute constant C such that for any $z \in \mathbb{R}$,*

$$\left| P(W \leq z) - \Phi(z) \right| \leq C \sum_{i=1}^{n} \left\{ \frac{EX_i^2 I(|X_i| > 1 + |z|)}{(1 + |z|)^2} + \frac{E|X_i|^3 I(|X_i| \leq 1 + |z|)}{(1 + |z|)^3} \right\}. \qquad (5)$$

Here, $\Phi(z)$, $z \in \mathbb{R}$ is the standard normal distribution function. However, the technique "leave one out" to deal with sums of independent variables is very similar to the exchangeable pair technique. If we begin with (9) (see in the proof of the main result) and use some results from Chen and Shao [4], it is not difficult to obtain (5). In some applications such as the quadratic forms treated later in this paper, where relation (3) is satisfied with $R = 0$ and $g(x) = x$, the non-uniform part $\frac{C}{1+|z|}$ in (4) can be improved significantly by replacing it with $\frac{C}{(1+|z|)^2}$.

§3   Proof of Theorem 2

Proof. In what follows, C is used to denote a constant whose value may change at each occurrence.

Since (W, W') and (W', W) have the same distribution and $E(\Delta | \mathbf{X}) = \lambda(g(W) + R)$, we obtain

$$0 = E(W - W')(f_z(W) + f_z(W'))$$

$$= E(W - W')(2f_z(W) + f_z(W') - f_z(W))$$

$$= 2\lambda E(g(W)f_z(W)) + 2\lambda Ef_z(W)R - E\Delta \int_{-\Delta}^{0} f_z'(W + t)dt.$$

Thus $Eg(W)f_z(W) = \frac{1}{2\lambda} E\Delta \int_{-\Delta}^{0} f_z'(W + t)dt - Ef_z(W)R$. Then

$$Ef_z'(W) - Eg(W)f_z(W)$$

$$= Ef_z'(W) - \frac{1}{2\lambda} E\Delta \int_{\Delta}^{0} f_z'(W + t)dt + Ef_z(W)R$$

$$= Ef'_z(W)\Big(1 - \frac{1}{2\lambda}\Delta^2\Big) - \frac{1}{2\lambda}E\Delta\int_{-\Delta}^0 f'_z(W+t) - f'_z(W)dt + Ef_z(W)R$$

$$= Ef'_z(W)\Big(1 - \frac{1}{2\lambda}E(\Delta^2|\mathbf{X})\Big) - \frac{1}{2\lambda}E\Delta\int_{-\Delta}^0 f'_z(W+t) - f'_z(W)dt + Ef_z(W)R.$$

With the notation $J = \frac{1}{2\lambda}E\Delta\int_{-\Delta}^0 f'_z(W+t) - f'_z(W)dt$, we find that

$$J = \frac{1}{2\lambda}E\Big(\Delta\int_{-\Delta}^0 g(W+t)f_z(W+t) - g(W)f_z(W)dt\Big)$$

$$+ \frac{1}{2\lambda}E\Big(\Delta\int_{-\Delta}^0 I(W+t \le z) - I(W \le z)dt\Big)$$

$$= J_1 + J_2,$$

where

$$J_1 = \frac{1}{2\lambda}E\Big(\Delta\int_{-\Delta}^0 g(W+t)f_z(W+t) - g(W)f_z(W)dt\Big),$$

$$J_2 = \frac{1}{2\lambda}E\Big(\Delta\int_{-\Delta}^0 I(W+t \le z) - I(W \le z)dt\Big).$$

From Shao and Zhang [13], it is known that

$$|J_1| \le \frac{1}{2\lambda}E\Delta\Delta^* g(W)f_z(W) \tag{6}$$

and

$$|J_2| \le \frac{1}{2\lambda}E\Delta\Delta^* I(W > z). \tag{7}$$

Observe that

$$E\Delta\Delta^* = 0.$$

Then we obtain

$$|J_2| \le \frac{1}{2\lambda}E\Delta\Delta^*\big(I(W > z) - 1\big) = \frac{1}{2\lambda}E\Delta\Delta^* I(W' \le z). \tag{8}$$

Combining (6) and (7), for $z > 0$, we have

$$\big|P(W \le z) - F(z)\big|$$

$$\le E\Big|f'_z(W)\Big(1 - \frac{1}{2\lambda}E(\Delta^2|\mathbf{X})\Big)\Big| + \frac{1}{2\lambda}E\Big|g(W)f_z(W)E(\Delta\Delta^*|\mathbf{X})\Big|$$

$$+ \frac{1}{2\lambda}E\big|E(\Delta\Delta^*|\mathbf{X})I(W > z)\big| + E|f_z(W)R|. \tag{9}$$

For $z \le 0$, using (6) and (8), we have

$$\big|P(W \le z) - F(z)\big|$$

$$\le E\Big|f'_z(W)\Big(1 - \frac{1}{2\lambda}E(\Delta^2|\mathbf{X})\Big)\Big| + \frac{1}{2\lambda}E\Big|g(W)f_z(W)E(\Delta\Delta^*|\mathbf{X})\Big|$$

$$+ \frac{1}{2\lambda}E\big|E(\Delta\Delta^*|\mathbf{X})I(W' \le z)\big| + E|f_z(W)R|. \tag{10}$$

The only difference between (9) and (10) is that the expectation $E\big|E(\Delta\Delta^*|\mathbf{X})I(W > z)\big|$ is replaced by $E\big|E(\Delta\Delta^*|\mathbf{X})I(W' \le z)\big|$.

To prove (4), we first assume that $z > 0$.

Using Cauchy's inequality to the fist term of (9) yields

$$E\Big|f'_z(W)\Big(1 - \frac{1}{2\lambda}E(\Delta^2|\mathbf{X})\Big)\Big| \leq \sqrt{E|f'_z(W)|^2} \cdot \sqrt{E\Big|(1 - \frac{1}{2\lambda}E(\Delta^2 \mid \mathbf{X}))\Big|^2}. \tag{11}$$

We will show now that

$$\sqrt{E|f'_z(W)|^2} \leq \frac{C}{1 + |g(z)|}. \tag{12}$$

Since $g(x), x \in \mathbb{R}$, satisfies (A4), for the τ in (A4) , we have

$$E|f'_z(W)|^2 = E|f'_z(W)|^2I(W \leq 0) + E|f'_z(W)|^2I(0 < W \leq \tau z) + E|f'_z(W)|^2I(W > \tau z).$$

Recall that for $x \leq 0$,

$$f'_z(x) = g(x)f_z(x) + 1 - F(z) = \Big(\frac{F(x)g(x)}{p(x)} + 1\Big) \cdot (1 - F(z)).$$

Because $g(x)f_z(x)$ is increasing in $x \in (-\infty, 0)$ and for fixed z, $F(z) - 1 \leq g(x)f_z(x) \leq F(z)$, we see that

$$-1 \leq \frac{F(x)g(x)}{p(x)} \leq \frac{F(0)g(0)}{p(0)} = 0.$$

Thus we conclude that $\frac{F(x)g(x)}{p(x)} + 1$ is bounded on $(-\infty, 0)$ and it does not depend on z. We notice further that

$$1 - F(z) = \int_z^\infty p(y)dy \leq \int_z^\infty \frac{g(y)}{g(z)}p(y)dy \leq \frac{p(z)}{g(z)}. \tag{13}$$

Hence

$$E|f'_z(W)|^2I(W \leq 0) \leq C(1 - F(z))^2 \leq C\Big(\frac{p(z)}{g(z)}\Big)^2 \leq \frac{C}{g^2(z)}. \tag{14}$$

By (13), we have

$$E|f'_z(W)|^2I(0 < W \leq \tau z) = EI(0 < W \leq \tau z) \cdot (1 + \frac{1}{c_1}F(W) \cdot g(W)e^{G(W)}) \cdot (1 - F(z))^2$$

$$\leq C\Big(1 + g(\tau z)exp\Big(\int_0^{\tau z} g(y)dy\Big)\Big)^2 \cdot exp\Big(-2\int_0^z g(y)dy\Big) \cdot \frac{1}{g^2(z)}.$$

We notice that

$$g(\tau z)exp\Big(\int_0^{\tau z} g(y)dy\Big) \cdot exp\Big(-\int_0^z g(y)dy\Big) = g(\tau z)exp\Big(-\int_{\tau z}^z g(y)dy\Big)$$

$$\leq g(\tau z)e^{-(1-\tau)zg(\tau z)}$$

$$\leq C.$$

Therefore $E|f'_z(W)|^2I(0 < W \leq \tau z) \leq \frac{C}{g^2(z)}$ and C depends on τ.

For the term $E|f'_z(W)|^2I(W > \tau z)$, by (B3) and (A4), we find, by Markov's inequality, that

$$E|f'_z(W)|^2I(W > \tau z) \leq P(W > \tau z)$$

$$\leq \frac{Eg^2(W)}{g^2(\tau z)}$$

$$\leq \frac{C}{g^2(\tau z)} = \frac{g^2(z)}{g^2(\tau z)} \cdot \frac{C}{g^2(z)}$$

$$\leq \frac{K_\tau^2 \cdot C}{g^2(z)}.$$

Thus $E|f'_z(W)|^2 \leq \frac{C}{g^2(z)}$ for $z > 0$ with a constant C depending on τ and $Eg^2(W)$.

The next is to use the fact that $\|f'_z\| \leq 1$ and see that

$$\sqrt{E \mid f'_z(W) \mid^2} \leq \min\left\{1, \frac{C}{|g(z)|}\right\} \leq \frac{C}{1+|g(z)|},$$

which complete the proof of (12) for $z > 0$. By (11) and (12), we have

$$E\left|f'_z(W)\left(1 - \frac{1}{2\lambda}E(\Delta^2|\mathbf{X})\right)\right| \leq \frac{C}{1+|g(z)|} \cdot \sqrt{E\left|(1 - \frac{1}{2\lambda}E(\Delta^2 \mid \mathbf{X}))\right|^2}. \tag{15}$$

Using Cauchy's inequality, for the second term of (9), we find

$$\frac{1}{2\lambda}E\left|g(W)f_z(W)E(\Delta\Delta^*|\mathbf{X})\right| \leq \sqrt{E|g(W)f_z(W)|^2} \cdot \frac{1}{2\lambda}\sqrt{E|E(\Delta\Delta^* \mid \mathbf{X})|^2}. \tag{16}$$

We will show that

$$\sqrt{E|g(W)f_z(W)|^2} \leq \frac{C}{1+|g(z)|}. \tag{17}$$

Since we know that $g(x)f_z(x) = f'_z(x) - \left(I(x \leq z) - F(z)\right)$, $\|gf_z\| \leq 1$ and $E|f'_z(W)|^2 \leq \frac{C}{g^2(z)}$, we only need to show that $E\left(I(W \leq z) - F(z)\right)^2 \leq \frac{C}{g^2(z)}$.

For $z > 0$,

$$E\left(I(W \leq z) - F(z)\right)^2 = E\left(1 - F(z)\right)^2 I(W \leq z) + F^2(z)I(W > z) \leq \frac{C}{g^2(z)}.$$

Thus we have proved (17) for $z > 0$. By (16) and (17), we have

$$\frac{1}{2\lambda}E\left|g(W)f_z(W)E(\Delta\Delta^*|\mathbf{X})\right| \leq \frac{C}{1+|g(z)|} \cdot \frac{1}{2\lambda}\sqrt{E|E(\Delta\Delta^* \mid \mathbf{X})|^2}. \tag{18}$$

For the third term of (9), we obtain

$$\frac{1}{2\lambda}E\left|E(\Delta\Delta^*|\mathbf{X})I(W > z)\right| \leq \sqrt{P(W > z)} \cdot \frac{1}{2\lambda}\sqrt{E\left|E(\Delta\Delta^*|\mathbf{X})\right|^2}. \tag{19}$$

By Markov's inequality,

$$P(W > z) \leq \frac{Eg^2(W)}{g^2(z)}$$

$$\leq \frac{C}{g^2(z)}.$$

Then , (19) becomes

$$\frac{1}{2\lambda}E\left|E(\Delta\Delta^*|\mathbf{X})I(W > z)\right| \leq \frac{C}{1+|g(z)|}\sqrt{\frac{1}{2\lambda}E\left|E(\Delta\Delta^*|\mathbf{X})\right|^2}. \tag{20}$$

From Shao, Zhang and Zhang [14], we know that

$$\|f_z\| \leq \min\left\{\frac{1}{c_1}, \frac{1}{|g(z)|}\right\} \tag{21}$$

for $z \in \mathbb{R}$. For the last term of (9), we have

$$E|f_z(W)R| \leq \frac{C}{1+|g(z)|}E|R|. \tag{22}$$

To show (22) holds, it suffices to consider $z \geq 0$ where $|g(z)| = g(z)$. There is $z_0 \in (0, +\infty]$ such that

$$\min\left\{c_1, \frac{1}{g(z)}\right\} = \begin{cases} c_1, z \leq z_0, \\ \frac{1}{g(z)}, z > z_0. \end{cases}$$

Since $\frac{1+g(z)}{g(z)}$ is bounded for $z > z_0$, $\frac{1}{g(z)} = \frac{g(z)+1}{g(z)} \cdot \frac{1}{1+g(z)} \leq \frac{C}{1+g(z)}$ for $z > z_0$. Also, for

$0 \le z \le z_0$, $c_1 = c_1 \cdot (1 + g(z)) \cdot \frac{1}{1+g(z)} \le \frac{C}{1+g(z)}$. Thus, $\min\left\{c_1, \frac{1}{|g(z)|}\right\} \le \frac{C}{1+|g(z)|}$. Then we have

$$E|f_z(W)R| \le \|f_z\|E|R| \le \min\left\{c_1, \frac{1}{|g(z)|}\right\} \cdot E|R| \le \frac{C}{1+|g(z)|}E|R|.$$

From (9), (15), (18), (20) and (22), it follows that we have proved (4) for $z > 0$.

For $z \le 0$, we take (10) and use Cauchy's inequality. For the third term of (10), it is easy to see that

$$\frac{1}{2\lambda}E\big|E(\Delta\Delta^*|\mathbf{X})I(W' \le z)\big| \le \sqrt{P(W' \le z)}\sqrt{\frac{1}{2\lambda}E\big|E(\Delta\Delta^*|\mathbf{X})\big|^2}$$

$$\le \frac{C}{1+|g(z)|}\sqrt{\frac{1}{2\lambda}E\big|E(\Delta\Delta^*|\mathbf{X})\big|^2}. \tag{23}$$

For the last term of (10), in view of (21),

$$E|f_z(W)R| \le \frac{C}{1+|g(z)|}E|R|. \tag{24}$$

Thus we only need to prove (12) and (17) for $z \le 0$.

For $z \le 0$, we have, for any $\tau \in (0,1)$, that

$$E|f_z'(W)|^2 = E|f_z'(W)|^2 I(W \le \tau z) + E|f_z'(W)|^2 I(\tau z \le W \le 0) + E|f_z'(W)|^2 I(W > 0).$$

By the same arguments as above, we obtain

$$F(z) \le \frac{p(z)}{|g(z)|}, \quad z \le 0. \tag{25}$$

Then following similar steps as in the proof for $z \ge 0$, we establish (12) for $z \le 0$. To prove (17) for $z \le 0$, it suffices to show that $E\big(I(W \le z) - F(z)\big)^2 \le C/|g(z)|^2$ for $z \le 0$. Indeed, by (25) and Markov's inequality,

$$E\big(I(W \le z) - F(z)\big)^2 \le 2P(W \le z) + 2F^2(z) \le \frac{C}{|g(z)|^2}.$$

Let us summarize our findings:(9), (15), (18), (20) and (22) show that the bound (4) is true for $z > 0$, while (12), (17) proved for $z \le 0$, and (24), (25) show that this bound holds for $z \le 0$.

Theorem 2 is proved.

$\square$

§4   Applications

4.1   Quadratic forms

Let $X_1, X_2, ..., X_n$ be i.i.d. random variables with zero mean, unit variance and a finite fourth moment. Let $A = (a_{ij})_{1 \le i,j \le n}$ be a real symmetric matrix with $a_{ii} = 0$ and let

$$W_n = \frac{1}{\sigma_n}\sum_{i \ne j} a_{ij}X_iX_j, \quad \sigma_n^2 = 2\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}^2.$$

This is a classical example which has been widely discussed in the literature. For example, de Jong [7] obtained the asymptotic normality of W_n, Chatterjee [1] gave an L^1 bound and Götze and Tikhomirov [10] studied the Kolmogorov distance between the distribution of W_n and the

distribution of the same quadratic forms with X_{ij} repalced by corresponding Gaussian random variables. Shao and Zhang [13] established the following bound:

$$\sup_{z \in \mathbb{R}} \left| P(W_n \leq z) - \Phi(z) \right| \leq \frac{CEX_1^4}{\sigma_n^2} \left(\sqrt{\sum_i \left(\sum_j a_{ij}^2 \right)^2} + \sqrt{\sum_{i,j} \left(\sum_k (a_{ik}a_{jk})^2 \right)} \right).$$

The next theorem is a non-uniform refinement of this bound.

Theorem 3. *Let* $\{X_1, X_2, ..., X_n\}$ *be i.i.d random variables with zero mean, unit variance and a finite fourth moment. Let* $A = (a_{ij})_{i,j=1}^n$ *be a real symmetric matrix with* $a_{ii} = 0$ *for all* $1 \leq i \leq n$. *Put* $W_n = \frac{1}{\sigma_n} \sum_{i \neq j} a_{ij} X_i X_j$ *and* $\sigma_n^2 = 2 \sum_{i=1}^n \sum_{j=1}^n a_{ij}^2$. *Then,*

$$\left| P(W_n \leq z) - \Phi(z) \right| \leq \frac{CEX_1^4}{(1+|z|)^2 \sigma_n^2} \left(\sqrt{\sum_i \left(\sum_j a_{ij}^2 \right)^2} + \sqrt{\sum_{i,j} \left(\sum_k (a_{ik}a_{jk})^2 \right)} \right), \quad (26)$$

where C *is an absolute constant depending on* EX_1^4.

Proof. Let $(X_1', X_2', ..., X_n')$ be an independent copy of $(X_1, X_2, ..., X_n)$ and θ a disrete uniformly distributed random variable over the set $\{1, 2, ..., n\}$ and independent of all oher random variables. Define

$$W_n' = W_n - \frac{2}{\sigma_n} \sum_{j=1}^n a_{\theta j} X_\theta X_j + \frac{2}{\sigma_n} \sum_{j=1}^n a_{\theta j} X_\theta' X_j.$$

Then (W, W') is an exhcangeable pair. It is easy to see that

$$\Delta = W_n - W_n' = \frac{2}{\sigma_n} \sum_{i=1}^n I\{\theta = i\} \sum_{j=i}^n a_{ij} X_j (X_i - X_i')$$

and

$$E(\Delta | \mathbf{X}) = \frac{2}{n} W_n.$$

These relations imply that condition (3) is satisfied with $g(x) = x$, $\lambda = \frac{2}{n}$ and $R = 0$. By Shao and Zhang [13],

$$E \left| 1 - \frac{1}{2\lambda} E(\Delta^2 | \mathbf{X}) \right|^2 \leq C \sigma_n^{-4} \left(E(X_1^4) \right)^2 \left(\sum_{i=1}^n \left(\sum_{j=1}^n a_{ij}^2 \right)^2 + \sum_{i,j=1}^n \left(\sum_{k=1}^n (a_{ik}a_{jk})^2 \right) \right) \quad (27)$$

and

$$Var \left(\frac{1}{\lambda} E(\Delta |\Delta|| \mathbf{X}) \right) \leq C \sigma_n^{-4} E^2 (X_1^4) \sum_i^n \left(\sum_j^n a_{ij}^2 \right)^2. \quad (28)$$

Note that $EX_1^4 < \infty$ and $EW_n^4 < C$ for any $n = 1, 2 \cdots$. Then,

$$P(|W_n| > z) \leq \frac{EW_n^4}{z^4} \wedge 1 = \min\{1, \ C/z^4\} \leq \frac{C}{(1+|z|)^4}.$$

For τ involved in (A4), we can take, for example $\tau = \frac{1}{2}$ and derive that

$$
\begin{aligned}
E|f_z'(W_n)|^2 I(0 < W \le \frac{1}{2}z) &\le \left[C\left(1 + ze^{\int_0^{z/2} y\, dy}\right)^2 \cdot e^{-2\int_0^z y\, dy} \cdot \frac{1}{z^2} \right] \wedge 1 \\
&\le \left[C\left(e^{-z^2}/z^2 + e^{-3z^2/4}\right) \right] \wedge 1 \\
&= \left[C\left(z^2 e^{-z^2}/z^4 + z^4 e^{-3z^2/4}/z^4\right) \right] \wedge 1 \\
&\le \min\{1,\ C/z^4\} \\
&\le \frac{C}{(1+|z|)^4}.
\end{aligned}
$$

By(14), we have

$$
E|f_z'(W_n)|^2 I(W \le 0) \le C\left(\frac{p(z)}{g(z)}\right)^2 = Ce^{-z^2/2}/z^2 = Cz^2 e^{-z^2/2}/z^4 \le C/z^4.
$$

Then,

$$
E|f_z'(W_n)|^2 \le \frac{C}{(1+|z|)^4}.
$$

Using the same arguments as those in the proof of the main result, we find that

$$
\sqrt{E|f_z'(W_n)|^2} \le \frac{C}{(1+|z|)^2},
$$

$$
\sqrt{E|W_n f_z(W_n)|^2} \le \frac{C}{(1+|z|)^2}.
$$

Hence the bound $\frac{C}{1+|z|}$ in (4) can be improved replacing it by $\frac{C}{(1+|z|)^2}$. Thus, referring to Theorem 2, in view of (27) and (28), we complete the proof of this theorem. $\qquad\square$

4.2   General Curie-Weiss model

The Curie-Weiss model is important in statistical physics and has been extensively discussed in the literature. For some history and the first asymptotic results, the reader is referred to Ellis and Newman [8], [9]. Using the technique of exchangeable pair approach, Chatterjee and Shao [3] studied a kind of Curie-Weiss model. Shao and Zhang [13] studied a general Curie-Weiss model and got the optimal convergence rate. In this subsection, we refine the bound in Shao and Zhang [13] to the non-uniform case.

Let $L(x), x \in \mathbb{R}$, be a distribution function satisfying the conditions:

$$
\int_{-\infty}^{+\infty} x\, dL(x) = 0 \quad \text{and} \quad \int_{-\infty}^{+\infty} x^2\, dL(x) = 1 \tag{29}
$$

For a positive integer k and a real number λ, say that L be of type k with strength λ, if

$$
\int_{-\infty}^{+\infty} x^j\, d\Phi(x) - \int_{-\infty}^{+\infty} x^j\, dL(x) = \begin{cases} 0, & \text{for } j = 1, \cdots, 2k-1, \\ \lambda, & \text{for } j = 2k, \end{cases}
$$

where, to recall that $\Phi(x), x \in \mathbb{R}$, is the standard normal distribution function.

Let $(X_1, ..., X_n)$ be a random vector with joint distribution function $P_{n,\beta}(\mathbf{x})$, $\mathbf{x} = (x_1, ..., x_n) \in$

$\mathbb{R}^n$, such that

$$dP_{n,\beta}(\mathbf{x}) = \frac{1}{K_n} \exp\left(\frac{\beta(x_1 + \dots + x_n)^2}{2n}\right) \prod_{i=1}^{n} dL(x_i) \tag{30}$$

where K_n is the normalizing constant. Let ξ be a random variable with distribution function L. Moreover, assume that:

(1) for $0 < \beta < 1$, there exists a constant $b > \beta$ such that

$$Ee^{t\xi} \le e^{t^2/2b}, \quad t \in \mathbb{R}. \tag{31}$$

(2) for $\beta = 1$, there exist constants $b_0 > 0$, $b_1 > 0$ and $b_2 > 1$ such that:

$$Ee^{t\xi} \le \begin{cases} \exp(t^2/2 - b_1 t^{2k}), & |t| \le b_0, \\ \exp(t^2/2b_2), & |t| > b_0. \end{cases} \tag{32}$$

We have the following results:

Theorem 4. *Suppose that the distribution function of the random vector $(X_1, X_2, ..., X_n)$ is given by (30), where L satisfies (29) and let $S_n = X_1 + \cdots + X_n$.*
(i) If $0 < \beta < 1$ and (31) is satisfied, $W_n = S_n/\sqrt{n}$. Then

$$\left|P(W_n \le z) - F_1(z)\right| \le \frac{C}{1 + (1-\beta)|z|} \cdot \frac{1}{\sqrt{n}}, \tag{33}$$

where $F_1(z)$, $z \in \mathbb{R}$, is the distribution function of a random variable $Z_1 \sim \mathcal{N}(0, \frac{1}{1-\beta})$ and C is a constant depending on b and β.
(ii) If $\beta = 1, L$ is of type k, (32) holds and $W_n = S_n/n^{1-1/2k}$, then

$$\left|P(W_n \le z) - F_k(z)\right| \le \frac{C}{1 + 2kc_2|z|^{2k-1}} \cdot \frac{1}{n^{1/2k}}, \tag{34}$$

where C is a constant depending on b_0, b_1, b_2 and k. $F_k(z)$, $z \in \mathbb{R}$, is the distribution function whose density function is $p_k(z) = c_1 e^{-c_2 y^{2k}}$, $c_2 = \dfrac{H^{(2k)}(0)}{(2k)!}$, c_1 is the normalizing constant, and

$$H(s) = \frac{s^2}{2} - \ln\left(\int_{-\infty}^{+\infty} exp(sx)dL(x)\right), \ s \in \mathbb{R}.$$

.

Proof. Recall that $S_n = \sum\limits_{i=1}^{n} X_i$. We first construct an exchangeable pair as follows. For a fixed i, $1 \le i \le n$, given $\{X_j, j \neq i\}$, let X_i' be a random variable which is conditionally independent of X_i and has the same conditional distribution as X_i. Let θ be a random index uniformly distributed over $\{1, \cdots, n\}$ and independent of all other random variables. Let $S_n' = S_n - X_\theta + X_\theta'$. Then (S_n, S_n') is an exhangeable pair.

When $0 < \beta < 1$, let $W_n = S_n/\sqrt{n}$ and $W_n' = S_n'/\sqrt{n}$. Then (W_n, W_n') is an exchangeable

pair. By Shao and Zhang [13], the following relations are satisfied:

$$E(W_n - W_n'|\mathbf{X}) = \frac{1}{n}\Big((1-\beta)W_n + \sqrt{n}R_2\Big); \tag{35}$$

$$E|R_2| \le Cn^{-1/2}; \tag{36}$$

$$E\Big|\frac{1}{2\lambda}E((W_n - W_n')^2|\mathbf{X}) - 1\Big|^2 \le Cn^{-1}; \tag{37}$$

$$E\Big|\frac{1}{2\lambda}E((S_n - S_n')|S_n - S_n'||\mathbf{X}) \Big|^2 \le n^{-1}. \tag{38}$$

Here C depends on β and b. Thus (3) is satisfied with $g(x) = (1-\beta)x$, and $\lambda = \frac{1}{n}$. Using (36), (37), (38) and Theorem 2 , we obtained (33).

When $\beta = 1$, recall that $W_n = S_n/n^{1-1/2k}$ and define $W_n' = S_n'/n^{1-1/2k}$, so (W_n, W_n') is an exchangeable pair. By Shao and Zhang [13], we obtain the following:

$$E(W_n - W_n'|\mathbf{X}) = n^{-2+1/k}\Big(\frac{H^{(2k)}(0)}{(2k-1)!}W_n^{2k-1} + n^{-1+1/2k}R_1\Big);$$

$$E|R_1| \le Cn^{-1/2k}; \tag{39}$$

$$E\Big|\frac{1}{2\lambda}E((W_n - W_n')^2|\mathbf{X}) - 1\Big|^2 \le Cn^{-1/k}; \tag{40}$$

$$E\Big|\frac{1}{2\lambda}E((W_n - W_n')^2|\mathbf{X}) - 1\Big|^2 \le Cn^{-1}. \tag{41}$$

Here C depends on β and b. Thus $g(x) = \frac{H^{(2k)}(0)}{(2k-1)!}x^{2k-1} = 2kc_2x^{2k-1}$ and $\lambda = n^{-2+\frac{1}{2k}}$. By (39),(40), (41) and Theorem 2, we obtain (34).   □

Acknowledgement

References

[1] S Chatterjee. *A new method of normal approximation*, Ann Probab, 2008, 36: 1584-1610.

[2] S Chatterjee. *A short survey of Stein's method*, Proceedings of the International Congress of Mathematicians-Seoul, 2014, 2014, IV: 1-24.

[3] S Chatterjee, Q Shao. *Nonnormal approximation by Stein's method of exchangeable pairs with application to the Curie-Weiss model*, Ann Appl Probab, 2014, 21: 464-483.

[4] L Chen, Q Shao. *A non-uniform Berry-Esseen bound via Stein's method*, Probab Theory Related Fields, 2001, 32: 1985-2028.

[5] L Chen, Q Shao. *Normal approximation under local dependence*, Ann Probab, 2004, 20: 236-254.

[6] Y Chen, Q Shao. *Berry-Esseen inequality for unbounded exchangeable pairs*, Probability approximations and beyond, Lect Notes Stat, Springer, New York, 2012, 205: 13-30.

[7] P de Jong. *A central limit theorem for generalized quadratic forms*, Probab Theory Related Fields, 1987, 75: 261-277.

[8] R Ellis, C Newman. *Limit theorems for sums of dependent random variables occurring in statistical mechanics*, Z Wahrsch Verw Gebiete, 1978, 44: 117-139.

[9] R Ellis, C Newman. *The statistics of Curie-Weiss models*, J Statist Phys, 1978, 19: 149-161.

[10] F Götze, A Tikhomirov. *Asymptotic distribution of quadratic forms*, Ann Probab, 1999, 27: 1072-1098.

[11] Q Shao, Z Su. *The Berry-Esseen bound for character ratios*, Proc Amer Math Soc, 2006, 134: 2153-2159.

[12] Y Rinott, V Rotar. *On coupling constructions and rates in the CLT for dependent summands with applications to the antivoter model and weighted U-statistics*, Ann Appl Probab, 1997, 7: 1080-1105.

[13] Q Shao, Z Zhang. *Berry-Esseen bounds of normal and nonnormal approximation for unbounded exchangeable pairs*, Ann Probab, 2019, 47: 61-108.

[14] Q Shao, M Zhang, Z Zhang. *Cramér-type moderate deviation theorems for non-normal approximation*, 2019, arXiv:1809.07966.

[15] C Stein. *A bound for the error in the normal approximation to the distribution of a sum of dependent random variables*, Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, Probability theory, 1972, II: 583-602.

[1]Institute for Financial Studies, Shandong University, Jinan 250100, China.

[2]School of Mathematics, Shandong University, Jinan 250100, China.

[3]Institute for Financial Studies, Shandong University, Jinan 250100, China.

[4]School of Mathematics, Shandong University, Jinan 250100, China.

Email: wanghanchao@sdu.edu.cn