Probab. Theory Relat. Fields 120, 236–254 (2001)
Digital Object Identifier (DOI) 10.1007/s004400100124

Louis H.Y. Chen · Qi-Man Shao

A non-uniform Berry–Esseen bound via Stein's method

Dedicated to Charles Stein on his eightieth birthday

Received: 2 March 2000 / Revised version: 20 July 2000 /
Published online: 26 April 2001 –

Abstract. This paper is part of our efforts to develop Stein's method beyond uniform bounds in normal approximation. Our main result is a proof for a non-uniform Berry–Esseen bound for independent and not necessarily identically distributed random variables without assuming the existence of third moments. It is proved by combining truncation with Stein's method and by taking the concentration inequality approach, improved and adapted for non-uniform bounds. To illustrate the technique, we give a proof for a uniform Berry–Esseen bound without assuming the existence of third moments.

1. Introduction

Since the publication of his paper on a new method of normal approximation in 1972, Stein's ideas have been applied to many other probability approximations, notably to Poisson, Poisson process, compound Poisson and binomial approximations. Most of these applications have been very successful. In the case of Poisson approximation, Stein's method has also found diverse applications in a wide range of fields. See, for example, Arratia, Goldstein and Gordon (1990), Barbour, Holst and Janson (1992), and Chen (1993). See also Chen (1998) for an account of Stein's method and a brief history of its developments.

Stein's method is of interest not only because it is striking, but also because it works well for dependent random variables, particularly if the dependence is local or of a combinatorial nature, thus lending itself to wide applicability. Despite its immense success, Stein's method is far from being fully exploited. For example, most of the work on Stein's method deals with uniform error bounds. Very little has

L.H.Y. Chen: Department of Mathematics, National University of Singapore, 2 Science Drive 2, Singapore 117543, Republic of Singapore. e-mail: lhychen@math.nus.edu.sg

Q.M. Shao: Department of Mathematics, University of Oregon, Eugene, OR 97403, USA. e-mail: shao@math.uoregon.edu

Research of the first author partially supported by grant RP3892719 at the National University of Singapore, and that of the second author by NSF grant DMS-9802451

Mathematics Subject Classification (2000): Primary 60F05; Secondary 60G50

Key words or phrases: Stein's method – Normal approximation – Non-uniform Berry–Esseen bound – Concentration inequality approach

been done on non-uniform error bounds or large deviations, mainly by Chen and Choi (1992) and Barbour, Chen and Choi (1995) for Poisson approximation with independent summands, and by Chen and Soon (1994) for binomial approximation in the special context of the binary expansion of a random integer. (See Chen (2000) for a brief survey of these works.)

This paper is part of our efforts to develop Stein's method beyond uniform bounds in normal approximation. Our main result is a proof for a non-uniform Berry–Esseen bound for independent and not necessarily identically distributed random variables without assuming the existence of third moments. The bound we have obtained is more general than that of Bikelis (1966), which assumed finiteness of the absolute third moments. It is proved by combining truncation with Stein's method and by taking the concentration inequality approach.

The concentration inequality approach was originally used by Stein for independent and identically distributed random variables (see Ho and Chen (1978)). It was extended by Chen (1986) to dependent and non-identically distributed random variables with arbitrary index set. A proof of the Berry–Esseen Theorem for independent and non-identically distributed random variables using the concentration inequality approach is given in Section 2 of Chen (1998).

In this paper, the concentration inequality approach is improved and extended to non-uniform bounds. The improved approach will be much more effective than that in Chen (1986, 1998) for handling dependent random variables. To illustrate the technique, we give a proof for a uniform Berry–Esseen bound as obtained by Feller (1968) for independent and non-identically distributed random variables without assuming the existence of third moments. Here we obtained 4.1 as the absolute constant in the bound, which is smaller than 6 as obtained by Feller (1968). This is the smallest absolute constant obtained so far by using Stein's method. Other proofs of the Berry–Esseen Theorem which have not been mentioned above are by Barbour and Hall (1984) and Stroock (1993). Their proofs are based on inductive arguments as in Bolthausen (1984), although the latter is concerned with a combinatorial central limit theorem.

2. Statements of theorems

Let $X_1, X_2, \cdots, X_n$ be independent and not necessarily identically distributed random variables with zero means and finite variances. Define $W = \sum_{i=1}^{n} X_i$ and assume that $\mathrm{Var}(W) = 1$. Let F be the distribution function of W and Φ the standard normal distribution function.

Theorem 2.1. *We have*

$$\sup_{-\infty<x<\infty} |F(x) - \Phi(x)| \leq 4.1\Big\{ \sum_{i=1}^{n} EX_i^2 I(|X_i| > 1) + \sum_{i=1}^{n} E|X_i|^3 I(|X_i| \leq 1)\Big\}. \tag{2.1}$$

The bound in (2.1) is more general than the Berry–Esseen bound with absolute third moments. It was obtained by Feller (1968) except that he had 6 as the absolute constant. It was pointed out by Loh (1975) that the truncation at 1 in the bound is optimal in the sense that

$$EX^2 I(|X| > 1) + E|X|^3 I(|X| \le 1) = \inf_A \Big\{ EX^2 I(X \in A) + E|X|^3 I(X \in A^c) \Big\}.$$

He also remarked that this fact seemed to have escaped notice in the literature. The next theorem is our main result.

Theorem 2.2. *There exists an absolute constant C such that for every real number x,*

$$|F(x) - \Phi(x)| \le C \sum_{i=1}^n \Big\{ \frac{EX_i^2 I(|X_i| > 1 + |x|)}{(1+|x|)^2} + \frac{E|X_i|^3 I(|X_i| \le 1 + |x|)}{(1+|x|)^3} \Big\}. \tag{2.2}$$

Non-uniform bounds were first obtained by Esseen (1945) for independent and identically distributed random variables with finite third moments. These were improved to $Cn\gamma/(1+|x|^3)$ by Nagaev (1965), where C is an absolute constant and γ the common absolute third moment of the X_i. Bikelis (1966) generalized Nagaev's result to

$$|F(x) - \Phi(x)| \le \frac{C \sum_{i=1}^n E|X_i|^3}{1 + |x|^3}$$

for independent and not necessarily identically distributed random variables. Paditz (1977) calculated C to be 114.7 and Michel (1981) reduced it to 30.54 for the independent and identically distributed case.

Throughout this paper, C stands for an absolute constant with possibly different values in different places.

3. Concentration inequalities

In this section we prove two concentration inequalities. Let $X_1, \cdots, X_n, W, F$ and Φ be as defined in Section 2. We begin with some useful facts.

(F1) For $a, b > 0$, $\min(a, b) \ge b - b^2/(4a)$;
(F2) For $a < b, y \ge 0, c > 0$,

$$I(a \le w \le b)\, y \ge c\Big(I(a \le w \le b) - (1 - y/c) I(y \le c) \Big).$$

Next we prove a lemma.

Lemma 3.1. *Let $\xi_1, \xi_2, \cdots, \xi_n$ be independent non-negative random variables with $E\xi_j^2 < \infty$ for each $1 \le j \le n$. Put $S = \sum_{j=1}^n \xi_j$, $\mu = \sum_{j=1}^n E\xi_j$ and $\sigma^2 = \sum_{j=1}^n E\xi_j^2$. Then for any $0 < c \le \mu$*

$$E\Big((1 - S/c) I(S \le c) \Big) \le \frac{\sigma}{c} \sqrt{2\pi} \Big(1 - \Phi(\frac{\mu - c}{\sigma}) \Big). \tag{3.1}$$

Proof. For any $t \ge 0$ and $x \le \mu$, we have

$$\begin{aligned} P(S \le x) &\le e^{tx} E e^{-tS} = e^{tx} \prod_{j=1}^{n} E e^{-t\xi_j} \\ &\le e^{tx} \prod_{j=1}^{n} E(1 - t\xi_j + t^2 \xi_j^2/2) \\ &\le \exp\Big(-t(\mu - x) + t^2\sigma^2/2\Big). \end{aligned}$$

Letting $t = (\mu - x)/\sigma^2$ yields

$$P(S \le x) \le \exp\left(-\frac{(\mu - x)^2}{2\sigma^2}\right).$$

Hence

$$\begin{aligned} E\Big((1 - S/c)I(S \le c)\Big) &= \int_0^c (1 - s/c)dP(S \le s) \\ &\le \frac{1}{c}\int_0^c P(S \le s)ds \\ &\le \frac{1}{c}\int_0^c \exp\left(-\frac{(\mu - s)^2}{2\sigma^2}\right)ds \\ &\le \frac{\sigma}{c}\int_{(\mu - c)/\sigma}^{\infty} \exp(-s^2/2)\,ds \\ &= \frac{\sigma}{c}\sqrt{2\pi}\Big(1 - \Phi(\frac{\mu - c}{\sigma})\Big). \end{aligned}$$

□

Let $W^{(i)} = W - X_i$ and let

$$\alpha = \sum_{i=1}^{n} EX_i^2 I(|X_i| > 1), \quad \beta = \sum_{i=1}^{n} E|X_i|^3 I(|X_i| \le 1) \tag{3.2}$$

and

$$\delta = 0.5(0.28\alpha + \beta). \tag{3.3}$$

The following proposition concerns a concentration inequality.

Proposition 3.2. *If* $\alpha + \beta \le 0.14$, *then for* $a < b$

$$P(a \le W^{(i)} \le b) \le 1.5(b - a) + 3.3\delta. \tag{3.4}$$

Proof. Define

$$f(x) = \begin{cases} -\frac{1}{2}(b - a) - \delta & \text{for } x < a - \delta \\ x - \frac{1}{2}(b + a) & \text{for } a - \delta \le x \le b + \delta \\ \frac{1}{2}(b - a) + \delta & \text{for } x > b + \delta \end{cases} \tag{3.5}$$

and let

$$M(w,t) = wI(-w \le t \le 0) - wI(0 < t \le -w).$$

Since X_j and $W^{(i)} - X_j$ are independent for $j \ne i$, $EX_j = 0$ and $M(w,t) \ge 0$, we have

$$\begin{aligned}
E\{W^{(i)} f(W^{(i)})\} &= \sum_{1\le j\le n, j\ne i} E\{X_j(f(W^{(i)}) - f(W^{(i)} - X_j))\} \\
&= \sum_{1\le j\le n, j\ne i} E\left\{X_j \int_{-X_j}^{0} f'(W^{(i)} + t)dt\right\} \\
&= \sum_{1\le j\le n, j\ne i} E\left\{\int_{-\infty}^{\infty} f'(W^{(i)} + t)M(X_j, t)dt\right\} \\
&\ge \sum_{1\le j\le n, j\ne i} E\left\{I(a \le W^{(i)} \le b)\int_{|t|\le\delta} M(X_j,t)\,dt\right\} \\
&= E\left\{I(a \le W^{(i)} \le b)\sum_{1\le j\le n, j\ne i} |X_j|\min(\delta, |X_j|)\right\} \\
&\ge E\left\{I(a \le W^{(i)} \le b)\sum_{j=1}^{n} \xi_j\right\} - P(a \le W^{(i)} \le b)E\xi_i \\
&\ge 0.38\left\{P(a \le W^{(i)} \le b) - E\left[(1 - S/0.38)I(S \le 0.38)\right]\right\} \\
&\quad -P(a \le W^{(i)} \le b)E\xi_i \\
&= (0.38 - E\xi_i)P(a \le W^{(i)} \le b) \\
&\quad -0.38E\left\{(1 - S/0.38)I(S \le 0.38)\right\} \qquad (3.6)
\end{aligned}$$

where $\xi_j = |X_j|I(|X_j| \le 1)\min(\delta, |X_j|I(|X_j| \le 1))$, $S = \sum_{j=1}^n \xi_j$ and we have used (F2) in the last inequality. Noting that $\alpha + \beta \le 0.14$ implies $\beta \le 0.14$ and $\delta \le 0.07$, we obtain from (F1) that

$$\begin{aligned}
ES &\ge \sum_{j=1}^{n} E\Big\{|X_j|I(|X_j| \le 1)\Big[|X_j|I(|X_j| \le 1) - X_j^2 I(|X_j| \le 1)/(4\delta)\Big]\Big\} \\
&= 1 - \alpha - \beta/(4\delta) \ge 1 - (0.28\alpha + \beta)/(4\delta) = 0.5.
\end{aligned}$$

From the definition of ξ_j,

$$E\xi_i \le \delta E|X_i|I(|X_i| \le 1) \le \delta\beta^{1/3} \le 0.5(0.14)^{1+1/3} \le 0.04$$

and

$$\sum_{j=1}^{n} E\xi_j^2 \le \delta^2 \le 0.07^2.$$

Then by Lemma 3.1,

$$0.38E\,[(1-S/0.38)I(S\le 0.38)] \le \delta\sqrt{2\pi}\Big(1-\Phi(\frac{0.5-0.38}{0.07})\Big) \le 0.11\delta.$$

Combining these inequalities with (3.5) and (3.6) and observing that

$$|E\{W^{(i)}f(W^{(i)})| \le \{\frac{1}{2}(b-a)+\delta\}E|W^{(i)}| \le \frac{1}{2}(b-a)+\delta,$$

we have

$$\begin{aligned} P(a\le W^{(i)}\le b) &\le (1/0.34)\Big(0.11\delta + E\{W^{(i)}f(W^{(i)})\}\Big)\\ &\le (1/0.34)\Big(0.11\delta+0.5(b-a)+\delta\Big)\\ &\le 1.5(b-a)+3.3\delta. \end{aligned}$$ □

Remark 3.3. In proving the concentration inequality in Proposition 3.2, we achieved small values of the constants in the bound by using Lemma 3.1, which exploits the independence among $X_1, \cdots, X_n$. If $X_1, \cdots, X_n$ are dependent, we can replace Lemma 3.1 by the following arguments. For $0 < c < ES$,

$$E(1-S/c)I(S\le c) \le P(S\le c) = P(ES-S\ge ES-c) \le \frac{E|S-ES|^p}{(ES-c)^p}$$

for $p \ge 1$.

Let

$$\alpha_y = \sum_{i=1}^n EX_i^2 I(|X_i|>1+y),\quad \beta_y = \sum_{i=1}^n E|X_i|^3 I(|X_i|\le 1+y) \tag{3.7}$$

and

$$\delta_y = \frac{\alpha_y}{(1+y)^2}+\frac{\beta_y}{(1+y)^3} \quad \text{for } y\ge 0. \tag{3.8}$$

The next proposition provides a non-uniform concentration inequality.

Proposition 3.4. *For any* $0\le a<b<\infty$, *we have*

$$P(a\le W^{(i)}\le b) \le C\Big\{\frac{b-a}{(1+a)^3}+\delta_a\Big\} \tag{3.9}$$

where C *is an absolute constant.*

Proof. Let

$$Y_j = X_j I(|X_j| \le 1+a) - EX_j I(|X_j| \le 1+a),$$
$$T^{(i)} = \sum_{1\le j\le n, j\ne i} Y_j \text{ and } r := r_i = \sum_{1\le j\le n, j\ne i} EX_j I(|X_j| > 1+a).$$

We first observe that

$$W^{(i)} = T^{(i)} + \sum_{1\le j\le n, j\ne i} EX_j I(|X_j| \le 1+a) = T^{(i)} - r$$

when $\max_{1\le j\le n} |X_j| \le 1+a$. Hence

$$\begin{aligned} P(a \le W^{(i)} \le b) &\le P(a+r \le T^{(i)} \le b+r) + P(\max_{1\le j\le n} |X_j| > 1+a) \\ &\le P(a+r \le T^{(i)} \le b+r) + \frac{\alpha_a}{(1+a)^2}\,. \end{aligned} \tag{3.10}$$

Next we consider two cases.
Case 1.

$$(1+a)^2\alpha_a + (1+a)\beta_a \ge 1/64. \tag{3.11}$$

By the Rosenthal (1970) inequality (see, for example, Petrov (1995), p. 59), we have

$$E|T^{(i)}|^4 \le C\Big\{(E|T^{(i)}|^2)^2 + \sum_{1\le j\le n, j\ne i} E|X_j|^4 I(|X_j| \le 1+a)\Big\} \le C(1+(1+a)\beta_a) \tag{3.12}$$

and hence

$$\begin{aligned} &P(a+r \le T^{(i)} \le b+r) \\ &\le P(1+a \le 1-r+T^{(i)}) \\ &\le \frac{E|1-r+T^{(i)}|^4}{(1+a)^4} \\ &\le \frac{8(1-r)^4 + 8E|T^{(i)}|^4}{(1+a)^4} \\ &\le \frac{C\Big(1+\sum_{j=1}^n E|X_j|^4 I(|X_j| \le 1+a)\Big)}{(1+a)^4} \\ &\le \frac{C}{(1+a)^4} + \frac{C\beta_a}{(1+a)^3} \\ &\le \frac{C\alpha_a}{(1+a)^2} + \frac{C\beta_a}{(1+a)^3} = C\delta_a \end{aligned} \tag{3.13}$$

where we have used (3.11) in the last inequality and the fact that $|r| \le \sum_{1\le j\le n} EX_j^2 = 1$.
Case 2.

$$(1+a)^2\alpha_a + (1+a)\beta_a < 1/64. \tag{3.14}$$

Let $\kappa = 16\beta_a$ and let

$$f(x) = \begin{cases} 0 & \text{for } x < a + r - \kappa, \\ (1 + x - r + \kappa)^3(x - a - r + \kappa) & \text{for } a + r - \kappa \le x \le b + r + \kappa, \\ (1 + x - r + \kappa)^3(b - a + 2\kappa) & \text{for } x > b + r + \kappa. \end{cases}$$

Note that f is a non-decreasing function satisfying

$$f'(x) \ge (1 + a)^3 \ \text{ for } a + r - \kappa \le x \le b + r + \kappa \ \text{ and } \ \ge 0 \ \text{ otherwise.}$$

Following the proof of Proposition 3.2, we have

$$\begin{aligned} &E\{T^{(i)} f(T^{(i)})\} \\ &\quad = \sum_{1 \le j \le n, j \ne i} E\{Y_j (f(T^{(i)}) - f(T^{(i)} - Y_j))\} \\ &\quad = \sum_{1 \le j \le n, j \ne i} E\left\{\int_{-\infty}^{\infty} f'(T^{(i)} + t) M(Y_j, t) dt\right\} \\ &\quad \ge \sum_{1 \le j \le n, j \ne i} (1 + a)^3 E\left\{I(a + r \le T^{(i)} \le b + r) \int_{|t| \le \kappa} M(Y_j, t)\, dt\right\} \\ &\quad = (1 + a)^3 E\left\{I(a + r \le T^{(i)} \le b + r) \sum_{1 \le j \le n, j \ne i} |Y_j| \min(\kappa, |Y_j|)\right\} \\ &\quad = (1 + a)^3 E\left\{I(a + r \le T^{(i)} \le b + r) \sum_{1 \le j \le n, j \ne i} \eta_j\right\} \\ &\quad \ge 0.5(1 + a)^3 \left\{P(a + r \le T^{(i)} \le b + r) - P(U \le 0.5)\right\} \qquad (3.15) \end{aligned}$$

where $\eta_j = |Y_j| \min(\kappa, |Y_j|)$, $U = \sum_{1 \le j \le n, j \ne i} \eta_j$, and we have used $(F2)$ in the last inequality. By (3.14) and (F1),

$$EY_i^2 \le EX_i^2 I(|X_i| \le 1 + a) \le \beta_a^{2/3} \le 1/16$$

and

$$\begin{aligned} EU &\ge \sum_{1 \le j \le n, j \ne i} E\left\{Y_j^2 - |Y_j|^3/(4\kappa)\right\} \\ &\ge \sum_{1 \le j \le n, j \ne i} \left\{EX_j^2 I(|X_j| \le 1 + a) - (EX_j I(|X_j| > 1 + a))^2\right\} - 2\beta_a/\kappa \\ &\ge \sum_{1 \le j \le n, j \ne i} \left\{EX_j^2 I(|X_j| \le 1 + a) - EX_j^2 I(|X_j| > 1 + a)\right\} - 2\beta_a/\kappa \\ &\ge 1 - EX_i^2 I(|X_i| \le 1 + a) - 2\alpha_a - 2\beta_a/\kappa \\ &\ge 1 - 1/16 - 1/16 - 2\beta_a/\kappa = 0.75\ . \end{aligned}$$

From the definition of η_j and by the Rosenthal inequality,

$$\begin{aligned}
P(U \le 0.5) &\le P(EU - U \ge 0.75 - 0.5)\\
&\le \frac{E|U-EU|^4}{0.25^4}\\
&\le C\Big\{\Big(\sum_{1\le j\le n} E\eta_j^2\Big)^2 + \sum_{1\le j\le n} E\eta_j^4\Big\}\\
&\le C\Big\{\kappa^4 + \kappa^4 \sum_{1\le j\le n} EY_j^4\Big\}\\
&\le C(\kappa^4 + (1+a)\kappa^4\beta_a)\\
&\le \frac{C\beta_a}{(1+a)^3}
\end{aligned}$$

where we have used (3.14) in the last inequality. Combining this with (3.15) and using (3.14), (3.12) and the Rosenthal inequality again, we obtain

$$\begin{aligned}
&P(a + r \le T^{(i)} \le b + r)\\
&\quad\le \frac{C\beta_a}{(1+a)^3} + \frac{2ET^{(i)}f(T^{(i)})}{(1+a)^3}\\
&\quad\le C(1+a)^{-3}\Big\{\beta_a + (b-a+2\kappa)E|T^{(i)}(1+T^{(i)}-r+\kappa)^3|\Big\}\\
&\quad\le C(1+a)^{-3}\Big\{\beta_a + (b-a+\kappa)(1+E|T^{(i)}|^4)\Big\}\\
&\quad\le C(1+a)^{-3}(\beta_a + b - a). \qquad (3.16)
\end{aligned}$$

By combining (3.10), (3.13) and (3.16), the concentration inequality (3.9) is proved. □

4. Proof of Theorem 2.1

We use the same notation as in the previous sections. Let

$$\bar{X}_i = X_i I(|X_i| \le 1) \text{ and } K_i(t) = E\left(\bar{X}_i I(0 \le t \le \bar{X}_i) - \bar{X}_i I(\bar{X}_i \le t < 0)\right).$$

We first derive a Stein identity for W.

Let f be a real-valued, bounded, continuous and piece-wise differentiable function defined on the real line. Then

$$\begin{aligned}
E\{Wf(W)\} &= \sum_{i=1}^n E\{X_i(f(W) - f(W^{(i)}))\}\\
&= \sum_{i=1}^n E\{X_i I(|X_i| \le 1)(f(W) - f(W^{(i)}))\}\\
&\quad + \sum_{i=1}^n E\{X_i I(|X_i| > 1)(f(W) - f(W^{(i)}))\}
\end{aligned}$$

$$\begin{aligned}
&= \sum_{i=1}^n E\Big\{\bar{X}_i \int_0^{\bar{X}_i} f'(W^{(i)}+t)dt\Big\} \\
&\quad + \sum_{i=1}^n E\{X_i I(|X_i|>1)(f(W)-f(W^{(i)}))\} \\
&= \sum_{i=1}^n E\Big\{\bar{X}_i \int_{-\infty}^{\infty} f'(W^{(i)}+t)(I(0\le t\le \bar{X}_i)-I(\bar{X}_i\le t<0))dt\Big\} \\
&\quad + \sum_{i=1}^n E\{X_i I(|X_i|>1)(f(W)-f(W^{(i)}))\} \\
&= \sum_{i=1}^n E\int_{-\infty}^{\infty} f'(W^{(i)}+t)K_i(t)dt \\
&\quad + \sum_{i=1}^n E\{X_i I(|X_i|>1)(f(W)-f(W^{(i)}))\}. \qquad (4.1)
\end{aligned}$$

Let f in (4.1) be the unique bounded solution f_x of the Stein equation

$$f'(w)-wf(w)=I(w\le x)-\Phi(x). \qquad (4.2)$$

By noting that

$$\sum_{i=1}^n \int_{-\infty}^{\infty} K_i(t)dt = \sum_{i=1}^n EX_i^2 I(|X_i|\le 1) = 1-\alpha$$

where α is defined in (3.2), (4.1) yields

$$\begin{aligned}
F(x)-\Phi(x) &= \sum_{i=1}^n E\left\{\int_{-\infty}^{\infty}(f_x'(W^{(i)}+X_i)-f_x'(W^{(i)}+t))K_i(t)dt\right\} \\
&\quad +\alpha Ef_x'(W) - \sum_{i=1}^n E\{X_i I(|X_i|>1)(f_x(W)-f_x(W^{(i)}))\} \\
&= R_1+R_2+R_3+R_4, \qquad (4.3)
\end{aligned}$$

where

$$\begin{aligned}
R_1 &= \sum_{i=1}^n E\left\{I(|X_i|\le 1)\int_{-\infty}^{\infty}(f_x'(W^{(i)}+X_i)-f_x'(W^{(i)}+t))K_i(t)dt\right\}, \\
R_2 &= \sum_{i=1}^n E\left\{I(|X_i|> 1)\int_{-\infty}^{\infty}(f_x'(W^{(i)}+X_i)-f_x'(W^{(i)}+t))K_i(t)dt\right\}, \\
R_3 &= \alpha Ef_x'(W), \\
R_4 &= -\sum_{i=1}^n E\{X_i I(|X_i|>1)(f_x(W)-f_x(W^{(i)}))\}.
\end{aligned}$$

Now the solution f_x of the Stein equation (4.2) is given by (see page 22 in Stein (1986))

$$f_x(w) = \begin{cases} \sqrt{2\pi}e^{w^2/2}(1-\Phi(w))\Phi(x), & w \ge x \\ \sqrt{2\pi}e^{w^2/2}(1-\Phi(x))\Phi(w), & w < x. \end{cases} \tag{4.4}$$

It follows from (59) and (60) on page 28 in Stein (1986) that

$$w \longmapsto wf_x(w) \text{ is increasing in } w. \tag{4.5}$$

It also follows from Lemma 2 and arguments on pages 23 and 24 in Stein (1986) that for all w and v,

$$0 \le f_x(w) \le \min((2\pi)^{1/2}/4,\ 1/|x|), \tag{4.6}$$

$$|wf_x(w)| \le 1, \tag{4.7}$$

$$|f_x'(w)| \le 1 \tag{4.8}$$

and

$$|f_x'(w) - f_x'(v)| \le 1. \tag{4.9}$$

(A proof of (4.9) is given in the Appendix). To bound $F(x) - \Phi(x)$, it suffices to consider $x \ge 0$ as we can simply apply the result to $-W$ when $x < 0$. For $x \ge 0$,

$$\begin{aligned} 1 - F(x) &= \inf_{a\ge 0} P(W + a \ge x + a) \\ &\le \inf_{a \ge 0} \frac{E(W+a)^2}{(x+a)^2} = \frac{1}{1+x^2}. \end{aligned}$$

It follows that for $x \ge 0$,

$$|F(x) - \Phi(x)| \le \sup_{x\ge 0} \Big|\frac{1}{1+x^2} - (1-\Phi(x))\Big| \le 0.55.$$

Hence we only need to consider the case where

$$\alpha + \beta \le 0.14. \tag{4.10}$$

We bound the error terms in (4.3) as follows. By (4.5)–(4.9), we have

$$\begin{aligned} |R_2 + R_3 + R_4| &\le \sum_{i=1}^n E\Big\{I(|X_i| > 1)\int_{-\infty}^{\infty} K_i(t)dt\Big\} \\ &\quad + \alpha + \sum_{i=1}^n E|X_i|I(|X_i|>1)(2\pi)^{1/2}/4 \\ &\le \sum_{i=1}^n P(|X_i|>1)E|X_i|^2 I(|X_i| \le 1) + 1.63\alpha \\ &\le \sum_{i=1}^n EX_i^2 I(|X_i|>1)(E|X_i|^3 I(|X_i|\le 1))^{2/3} + 1.63\alpha \\ &\le \alpha\beta^{2/3} + 1.63\alpha \le 0.14^{2/3}\alpha + 1.63\alpha \\ &\le 1.9\alpha. \end{aligned} \tag{4.11}$$

To bound R_1, we note by (4.2), (4.5)–(4.9) that

$$
\begin{aligned}
& f_x'(w+s) - f_x'(w+t) \\
&= \begin{cases} (w+s)f_x(w+s) - (w+t)f_x(w+t) + 1 & \text{if } w+s \le x, w+t > x \\ (w+s)f_x(w+s) - (w+t)f_x(w+t) - 1 & \text{if } w+s > x, w+t \le x \\ (w+s)f_x(w+s) - (w+t)f_x(w+t) & \text{if } w+s \le x, w+t \le x \\ & \text{or } w+s > x, w+t > x \end{cases} \\
&\le \begin{cases} 1 & \text{if } w+s \le x, w+t > x \\ w(f_x(w+s) - f_x(w+t)) + sf_x(w+s) - tf_x(w+t) & \text{if } s \ge t \\ 0 & \text{otherwise} \end{cases} \\
&\le \begin{cases} 1 & \text{if } w+s \le x, w+t > x \\ (|w| + (2\pi)^{1/2}/4)(|s|+|t|) & \text{if } s \ge t \\ 0 & \text{otherwise.} \end{cases}
\end{aligned}
$$

By Proposition 3.2,

$$
\begin{aligned}
R_1 &\le \sum_{i=1}^n E\left\{ I(|X_i| \le 1) \int_{W^{(i)}+t>x, W^{(i)}+X_i \le x} K_i(t)dt \right\} \\
&\quad + \sum_{i=1}^n E\left\{ I(|X_i| \le 1) \int_{X_i \ge t} (|W^{(i)}| + 0.63)(|X_i|+|t|)K_i(t)dt \right\} \\
&\le \sum_{i=1}^n E\left\{ I(|X_i| \le 1) \int_{t>X_i} P\Big(x - t < W^{(i)} < x - X_i \mid X_i\Big) K_i(t)dt \right\} \\
&\quad + \sum_{i=1}^n E(|W^{(i)}| + 0.63) E\left\{ I(|X_i| \le 1) \int_{X_i \ge t} (|X_i|+|t|)K_i(t)dt \right\} \\
&\le \sum_{i=1}^n E\left\{ I(|X_i| \le 1) \int_{t>X_i} (1.5(|t|+|X_i|) + 3.3\delta) K_i(t)dt \right\} \\
&\quad +1.63 \sum_{i=1}^n E\left\{ I(|X_i| \le 1) \int_{X_i \ge t} (|X_i|+|t|)K_i(t)dt \right\} \\
&\le 3.3\delta + 1.63 \sum_{i=1}^n E\left\{ I(|X_i| \le 1) \int_{-\infty}^{\infty} (|t|+|X_i|)K_i(t)dt \right\} \\
&\le 3.3\delta + 1.63 \sum_{i=1}^n E\left\{ I(|X_i| \le 1)(0.5E|\bar{X}_i|^3 + |X_i|E\bar{X}_i^2) \right\} \\
&\le 3.3\delta + 1.63 \times 1.5\beta \le 0.5\alpha + 4.1\beta. \qquad (4.12)
\end{aligned}
$$

Similarly,

$$
f_x'(w+s) - f_x'(w+t) \ge \begin{cases} -1 & \text{if } w+s > x, w+t \le x \\ -(|w| + (2\pi)^{1/2}/4)(|s|+|t|) & \text{if } s < t \\ 0 & \text{otherwise} \end{cases}
$$

and

$$R_1 \geq -0.5\alpha - 4.1\beta. \tag{4.13}$$

By combining (4.11), (4.12) and (4.13), Theorem 2.1 is proved. □

5. Proof of Theorem 2.2

We use the same notation as in the previous sections. We first prove two lemmas.

Lemma 5.1. *For $x > 0$, we have*

$$E|f_x'(W)| \leq C(1+x)^{-2}.$$

Proof. By (4.8), and (32) and (33) on page 24 in Stein (1986),

$$\begin{aligned} E|f_x'(W)| &= E|f_x'(W)|I(W \leq 0) + E|f_x'(W)|I(W > x/2) \\ &\quad + E|f_x'(W)|I(0 < W \leq x/2) \\ &\leq (1-\Phi(x))P(W \leq 0) + E(1+W)^2/(1+x/2)^2 \\ &\quad + (1-\Phi(x))E\Big(1+\sqrt{2\pi}We^{W^2/2}\Big)I(0 < W \leq x/2) \\ &\leq C(1+x)^{-2} + (1-\Phi(x))(1+2xe^{x^2/8}) \\ &\leq C(1+x)^{-2}. \end{aligned}$$

□

Lemma 5.2. *Let $\tau = \tau_x = x/4$ and let*

$$g(w) = (wf_x(w))'. \tag{5.1}$$

Then for $x \geq 4$ and $|u| \leq 1+\tau$, we have

$$Eg(W^{(i)}+u) \leq C\Big((1+x)^{-3} + x\delta_\tau\Big), \tag{5.2}$$

where δ_τ is as in (3.8).

Proof. From the definition of g and f_x, we get

$$g(w) = \begin{cases} \Big(\sqrt{2\pi}(1+w^2)e^{w^2/2}(1-\Phi(w)) - w\Big)\Phi(x), & w \geq x \\ \Big(\sqrt{2\pi}(1+w^2)e^{w^2/2}\Phi(w) + w\Big)(1-\Phi(x)), & w < x. \end{cases} \tag{5.3}$$

A direct calculation (see Appendix) shows that for $w \geq 0$,

$$0 \leq \sqrt{2\pi}(1+w^2)e^{w^2/2}(1-\Phi(w)) - w \leq \frac{2}{1+w^3}. \tag{5.4}$$

This implies that $g \geq 0$, $g(w) \leq 2(1-\Phi(x))$ for $w \leq 0$ and $\leq \frac{2}{1+w^3}$ for $w \geq x$; furthermore, g is clearly increasing for $0 \leq w < x$. Therefore

$$\begin{aligned}
Eg(W^{(i)}+u) &= Eg(W^{(i)}+u)I(W^{(i)}+u \leq x-1) \\
&\quad + Eg(W^{(i)}+u)I(W^{(i)}+u \geq x) \\
&\quad + Eg(W^{(i)}+u)I(x-1 < W^{(i)}+u < x) \\
&\leq 2(1-\Phi(x)) + g(x-1) + 2(1+x^3)^{-1} \\
&\quad + Eg(W^{(i)}+u)I(x-1 < W^{(i)}+u < x) \\
&\leq C\Big((1+x)^{-3} + x^2 e^{(x-1)^2/2}(1-\Phi(x))\Big) \\
&\quad + Eg(W^{(i)}+u)I(x-1 < W^{(i)}+u < x) \\
&\leq C(1+x)^{-3} + Eg(W^{(i)}+u)I(x-1 < W^{(i)}+u < x). \qquad (5.5)
\end{aligned}$$

Next we observe that δ_x is decreasing in x. Therefore, by Proposition 3.4, (4.5), (4.7) and the fact that $w-u \geq \tau$ for $w \geq x-1 \geq 3$ and $|u| \leq 1+\tau$,

$$\begin{aligned}
&Eg(W^{(i)}+u)I(x-1 < W^{(i)}+u < x) \\
&\quad = \int_{x-1}^{x} -g(w)dP(w < W^{(i)}+u < x) \\
&\quad = g(x-1)P(x-1 < W^{(i)}+u < x) + \int_{x-1}^{x} g'(w)P(w < W^{(i)}+u < x)dw \\
&\quad \leq C(1+x)^{-3} + C\int_{x-1}^{x} g'(w)\Big\{\delta_\tau + (1+x)^{-3}(x-w)\Big\}dw \\
&\quad \leq C\Big\{(1+x)^{-3} + \delta_\tau g(x-) + (1+x)^{-3}\int_{x-1}^{x}(x-w)dg(w)\Big\} \\
&\quad \leq C\Big\{(1+x)^{-3} + \delta_\tau x + (1+x)^{-3}\int_{x-1}^{x} g(w)dw\Big\} \\
&\quad \leq C\Big\{(1+x)^{-3} + \delta_\tau x + (1+x)^{-3} x f_x(x)\Big\} \\
&\quad \leq C\Big\{(1+x)^{-3} + \delta_\tau x\Big\}. \qquad (5.6)
\end{aligned}$$

Combining (5.5) and (5.6) yields (5.2). □

We are now ready to prove Theorem 2.2. In view of Theorem 2.1, we may, without loss of generality, assume $x \geq 4$. If

$$(1+x)^2\alpha_x + (1+x)\beta_x \geq 1/64,$$

then, the arguments in (3.10) and (3.13) yield $P(W \geq x) \leq C\delta_x$, which in turn implies

$$|F(x)-\Phi(x)| \leq P(W \geq x) + 1 - \Phi(x) \leq C\delta_x + C(1+x)^{-4} \leq C\delta_x.$$

Hence (2.2) holds in this case. So we assume

$$(1+x)^2\alpha_x + (1+x)\beta_x < 1/64. \qquad (5.7)$$

Let

$$\tau = \tau_x = x/4.$$

By noting that

$$\delta_\tau \le \frac{(1+x)^3}{(1+\tau)^3}\delta_x \le C\delta_x \ \text{ for } \ x \ge 4$$

where δ_x is defined in (3.8), it suffices to show that

$$|F(x) - \Phi(x)| \le C\delta_\tau. \tag{5.8}$$

Let

$$\begin{aligned} \bar{X}_{i,\tau} &= X_i I(|X_i| \le 1+\tau) \text{ and} \\ K_{i,\tau}(t) &= E\left(\bar{X}_{i,\tau} I(0 \le t \le \bar{X}_{i,\tau}) - \bar{X}_{i,\tau} I(\bar{X}_{i,\tau} \le t < 0)\right). \end{aligned}$$

Using the same arguments as for (4.3), we have

$$F(x) - \Phi(x) = R_{1,\tau} + R_{2,\tau} + R_{3,\tau} + R_{4,\tau} \tag{5.9}$$

where

$$\begin{aligned} R_{1,\tau} &= \sum_{i=1}^n E\left\{ I(|X_i| \le 1+\tau) \int_{|t|\le 1+\tau} (f_x'(W^{(i)}+X_i) - f_x'(W^{(i)}+t))K_{i,\tau}(t)dt \right\}, \\ R_{2,\tau} &= \sum_{i=1}^n E\left\{ I(|X_i| > 1+\tau) \int_{|t|\le 1+\tau} (f_x'(W^{(i)}+X_i) - f_x'(W^{(i)}+t))K_{i,\tau}(t)dt \right\}, \\ R_{3,\tau} &= \alpha_\tau E f_x'(W), \\ R_{4,\tau} &= -\sum_{i=1}^n E\{X_i I(|X_i| > 1+\tau)(f_x(W) - f_x(W^{(i)})\}. \end{aligned}$$

By (4.6), (4.8) and Lemma 5.1,

$$\begin{aligned} &|R_{2,\tau} + R_{3,\tau} + R_{4,\tau}| \\ &\quad \le \sum_{i=1}^n P(|X_i| > 1+\tau) + C\alpha_\tau/(1+x)^2 + \sum_{i=1}^n \frac{E|X_i|I(|X_i| > 1+\tau)}{x} \\ &\quad \le C\alpha_\tau/(1+\tau)^2. \end{aligned} \tag{5.10}$$

Noting that

$$|f_x'(w+s) - f_x'(w+t) - \int_t^s g(w+u)\,du| \le I(x - \max(s,t) < w \le x - \min(s,t))$$

where g is given by (5.3), we have

$$
\begin{aligned}
|R_{1,\tau}| \le\ & \sum_{i=1}^n \left| E\left\{ I(|X_i| \le 1+\tau) \int_{|t|\le 1+\tau} K_{i,\tau}(t) \int_t^{X_i} g(W^{(i)}+u)\,du\,dt \right\}\right| \\
& + \sum_{i=1}^n E\left\{ I(|X_i| \le 1+\tau) \int_{|t|\le 1+\tau} I(x-\max(t,X_i) \le W^{(i)} \le x \right. \\
& \qquad \left. - \min(t,X_i)) K_{i,\tau}(t)\,dt \right\} \\
=\ & \sum_{i=1}^n \left| E\left\{ I(|X_i| \le 1+\tau) \int_{|t|\le 1+\tau} K_{i,\tau}(t) \int_t^{X_i} Eg(W^{(i)}+u)\,du\,dt \right\}\right| \\
& + \sum_{i=1}^n E\left\{ I(|X_i| \le 1+\tau) \int_{|t|\le 1+\tau} P\Big(x-\max(t,X_i) \le W^{(i)} \le x \right. \\
& \qquad \left. - \min(t,X_i) \mid X_i\Big) K_{i,\tau}(t)\,dt \right\} \\
:=\ & R_{5,\tau} + R_{6,\tau}\ . \qquad (5.11)
\end{aligned}
$$

Now by Lemma 5.2 and (5.7),

$$
\begin{aligned}
R_{5,\tau} &\le C \sum_{i=1}^n \left| E\left\{ I(|X_i| \le 1+\tau) \int_{|t|\le 1+\tau} K_{i,\tau}(t) \int_t^{X_i} ((1+x)^{-3} + x\delta_\tau) du\,dt \right\}\right| \\
&\le C((1+x)^{-3} + x\delta_\tau) \sum_{i=1}^n E|\bar{X}_{1,\tau}|^3 \\
&\le C((1+x)^{-3}\beta_\tau + \delta_\tau x \beta_\tau) \\
&\le C\delta_\tau. \qquad (5.12)
\end{aligned}
$$

Next by Proposition 3.4,

$$
\begin{aligned}
R_{6,\tau} &\le C \sum_{i=1}^n E\left\{ I(|X_i| \le 1+\tau) \int_{|t|\le 1+\tau} \Big(\delta_\tau + (1+x)^{-3}(|t|+|X_i|)\Big) K_{i,\tau}(t)\,dt \right\} \\
&\le C\Big(\delta_\tau + (1+x)^{-3} \sum_{i=1}^n E|\bar{X}_{1,\tau}|^3\Big) \\
&\le C\delta_\tau. \qquad (5.13)
\end{aligned}
$$

By combining (5.9)-(5.13), (5.8) and hence Theorem 2.2 is proved. □

Acknowledgements. We thank two anonymous referees for their valuable comments.

6. Appendix

Proof of (5.4). It suffices to show that for $w \geq 0$,

$$\int_w^\infty e^{-x^2/2}dx \geq \frac{w}{1+w^2}e^{-w^2/2} \tag{6.1}$$

and

$$\int_w^\infty e^{-x^2/2}dx \leq \Big(\frac{w}{1+w^2} + \frac{2}{(1+w^2)(1+w^3)}\Big)e^{-w^2/2}. \tag{6.2}$$

Let

$$h(w) = \frac{w}{1+w^2}e^{-w^2/2}.$$

Then

$$h(w) = \int_w^\infty (-h'(x))dx$$

where

$$-h'(x) = \Big(1 - \frac{2}{(1+x^2)^2}\Big)e^{-x^2/2} \leq e^{-x^2/2}.$$

Therefore (6.1) holds.

To prove (6.2), let

$$g(w) = \Big(\frac{w}{1+w^2} + \frac{2}{(1+w^2)(1+w^3)}\Big)e^{-w^2/2}.$$

We have

$$-g'(x)e^{x^2/2} = 1 + 2\Big(\frac{3x-1}{(1+x^2)^2(1+x^3)} + \frac{3x^2}{(1+x^2)(1+x^3)^2}\Big) \geq 1$$

for all $x \geq 1/3$. Therefore (6.2) holds for $w \geq 1/3$. For $0 \leq w \leq 1/3$, the function below is decreasing and so

$$\begin{aligned}&\sqrt{2\pi}(1+w^2)e^{w^2/2}(1-\Phi(w)) - w \\ &\leq \sqrt{2\pi}(1-\Phi(0)) = \sqrt{2\pi}/2 < \frac{2}{1+(1/3)^3} \leq \frac{2}{1+w^3}.\end{aligned}$$

Hence (6.2) remains true for $0 \leq w \leq 1/3$. This proves (5.4). □

Proof of (4.9). By the Stein equation (4.2), we have

$$f_x'(w) = \begin{cases} wf_x(w) + 1 - \Phi(x) & \text{for } w < x \\ wf_x(w) - \Phi(x) & \text{for } w > x. \end{cases}$$

In view of the fact that $wf_x(w)$ is an increasing function of w, by (30) and (31) on page 24 in Stein (1986)

$$0 < f_x'(w) \leq xf_x(x) + 1 - \Phi(x) \;\; \text{for} \;\; w < x$$

and

$$-1 < xf_x(x) - \Phi(x) \leq f_x'(w) < 0 \;\; \text{for} \;\; w > x\ .$$

Hence for any w and v,

$$|f_x'(w) - f_x'(v)| \leq \max(1, xf_x(x) + 1 - \Phi(x) - (xf_x(x) - \Phi(x)) = 1.$$

References

1. Arratia, R., Goldstein, L., Gordon, L.: Poisson approximation and the Chen-Stein method, *Statist. Sci.*, **5**, 403–434 (1990)
2. Barbour, A.D., Chen, L.H.Y., Choi, K.P.: Poisson approximation for unbounded functions, I: independent summands, *Statist. Sinica*, **5**, 749–766 (1995)
3. Barbour, A.D., Hall, P.: Stein's method and the Berry-Esseen theorem, *Austral. J. Statist.*, **26**, 8–15 (1984)
4. Barbour, A.D., Holst, L., Janson, S.: *Poisson Approximation*. Oxford Studies in Probability 2, Clarendon Press, Oxford (1992)
5. Bikelis, A.: Estimates of the remainder in the central limit theorem, *Litovsk. Mat. Sb.*, **6**(3), 323–46 (in Russian) (1966)
6. Bolthausen, E.: An estimate of the remainder in a combinatorial central limit theorem, *Z. Wahrsch. Verw. Gebiete*, **66**, 379–386 (1984)
7. Chen, L.H.Y.: The rate of convergence in a central limit theorem for dependent random variables with arbitrary index set, IMA Preprint Series #243, Univ. Minnesota (1986)
8. Chen, L.H.Y.: Extending the Poisson approximation, *Science*, **262**, 379–380 (1993)
9. Chen, L.H.Y.: Stein's method: some perspectives with applications, *Probability Towards 2000* (L. Accardi and C. C. Heyde, eds.), Lecture Notes in Statistics No. 128, Springer Verlag, 97–122 (1998)
10. Chen, L.H.Y.: Non-uniform bounds in probability approximations using Stein's method, *Probability and Statistical Models with Applications: A Volume in Honor of Theophilos Cacoullos* (Ch. A. Charalambides, M. V. Koutras and N. Balakrishnan, eds.), Chapman & Hall/CRC Press, Boca Raton, Florida, 3–14 (2000)
11. Chen, L.H.Y., Choi, K.P.: Some asymptotic and large deviation results in Poisson approximation, *Ann. Probab.*, **20**, 1867–1876 (1992)
12. Chen, L.H.Y., Soon, S.Y.T.: On the number of ones in the binary expansion of a random integer. Unpublished manuscript (1994)
13. Esseen, C.G.: Fourier analysis of distribution functions: a mathematical study of the Laplace-Gaussian law, *Acta Math.*, **77**, 1–125 (1945)
14. Feller, W.: On the Berry-Esseen theorem, *Z. Wahrsch. Verw. Gebiete*, **10**, 261–268 (1968)
15. Ho, S.T., Chen, L.H.Y.: An L_p bound for the remainder in a combinatorial central limit theorem, *Ann. Probab.*, **6**, 231–249 (1978)
16. Loh, W.Y.: On the normal approximation for sums of mixing random variables. Master Thesis, Department of Mathematics, University of Singapore (1975)
17. Michel, R.: On the constant in the non-uniform version of the Berry-Esseen Theorem, *Z. Wahrsch. Verw. Gebiete*, **55**, 109–117 (1981)
18. Nagaev, S.V.: Some limit theorems for large deviations, *Theory Probab. Appl.*, **10**, 214–235 (1965)
19. Paditz, L.: Über die Annäherung der Verteilungsfunktionen von Summen unabhängiger Zufallsgröben gegen unberrenzt teilbare Verteilungsfunktionen unter besonderer berchtung der Verteilungsfunktion der standarddisierten Normalverteilung. Dissertation, A.TU Dresden (1977)
20. Petrov, V.V.: *Limit Theorems of Probability Theory: Sequences of Independent Random Varaibles*, Oxford Studies in Probability 4, Clarendon Press, Oxford (1995)

21. Rosenthal, H.P.: On the subspaces of $L^p (p > 2)$ spanned by sequences of independent random variables, *Israel J. Math.*, **8**, 273–303 (1970)
22. Stein, C.: A bound for the error in the normal approximation to the distribution of a sum of dependent random variables, *Proc. Sixth Berkeley Symp. Math. Stat. Prob.*, **2**, 583–602, Univ. California Press. Berkeley, Calif (1972)
23. Stein, C.: *Approximation Computation of Expectations*, Lecture Notes 7, Inst. Math. Statist., Hayward, Calif (1986)
24. Stroock, D.W.: *Probability Theory: an analytic view*, Cambridge Univ. Press, Cambridge, U.K. (1993)