Edgeworth Expansion by Stein’s Method

Xiao Fang and Song-Hao Liu

The Chinese University of Hong Kong and Southern University of Science and Technology

Abstract: Edgeworth expansion provides higher-order corrections to the normal approximation for a probability distribution. The classical proof of Edgeworth expansion is via characteristic functions. As a powerful method for distributional approximations, Stein’s method has also been used to prove Edgeworth expansion results. However, these results assume that either the test function is smooth (which excludes indicator functions of the half line) or that the random variables are continuous (which excludes random variables having only a continuous component). Thus, how to recover the classical Edgeworth expansion result using Stein’s method has remained an open problem. In this paper, we develop Stein’s method for two-term Edgeworth expansions in a general case. Our approach involves repeated use of Stein equations, Stein identities via Stein kernels, and a replacement argument.

AMS 2020 subject classification: 60F05, 62E17

Keywords and phrases: Central limit theorem, Edgeworth expansion, Stein kernel, Stein’s method.

1 Introduction and main results

Let $X_1, X_2, \ldots$ be a sequence of independent and identically distributed (i.i.d.) random variables with mean 0 and variance 1. Let $W_n = (X_1 + \cdots + X_n)/\sqrt{n}$, $n \geq 1$. The central limit theorem states that as $n \to \infty$, the distribution of $W_n$ converges to the standard normal distribution. Under the finite third moment assumption, we have the following non-asymptotic error bound [Berry (1941); Esseen (1942)]:

$$\sup_{x \in \mathbb{R}} |P(W_n \leq x) - P(Z \leq x)| \leq \frac{\mathbb{E}|X_1|^3}{\sqrt{n}},$$

(1.1)

where $Z \sim N(0,1)$ is a standard normal random variable.

In general, the $1/\sqrt{n}$ rate in (1.1) cannot be improved by considering the normal approximation of binomial distributions. However, under additional smoothness conditions, it is possible to devise asymptotic expansions incorporating higher-order correction terms for normal approximation. Such asymptotic expansions are called Edgeworth expansions. For example, if we assume that $X_1$ has a finite fourth moment and its characteristic function $\varphi(t)$ satisfies the Cramér condition:

$$\limsup_{|t| \to \infty} |\varphi(t)| < 1,$$

(1.2)
then we have
\[
\sup_{x \in \mathbb{R}} \left| \mathbb{P}(W_n \leq x) - P(Z \leq x) - \frac{\gamma}{6\sqrt{n}} \mathbb{E}\left[ (Z^3 - 3Z)1_{\{Z \leq x\}} \right] \right| \leq \frac{C}{n}, \tag{1.3}
\]
where \( \gamma = \mathbb{E}[X_1^3] \) and \( C \) is a constant not depending on \( n \). The Cramér condition is satisfied if the distribution of \( X_1 \) has a nonzero and absolutely continuous component with respect to the Lebesgue measure. We refer to Petrov (1975) and Bhattacharya and Rao (1976) for classical Edgeworth expansion results with error bounds in both continuous and discrete cases and in multi-dimensions. As the condition (1.2) indicates, all of the classical proofs of Edgeworth expansion results are obtained via the characteristic function approach.

Stein’s method was introduced in Stein (1972) and has become a main tool for distributional approximations (cf. Chen et al. (2011)). Stein’s method for asymptotic expansions was first studied by Barbour (1986). In the case of two-term Edgeworth expansion, he proved that (dropping the Cramér condition but assuming \( h \) is second-order differentiable, see also Fang (2019, Eq.(3.5)))
\[
\left| \mathbb{E}h(W_n) - \mathbb{E}h(Z) - \frac{\gamma}{6\sqrt{n}} \mathbb{E}\left[ (Z^3 - 3Z)h(Z) \right] \right| \leq \frac{C}{n} \|h''\|_{\infty}, \tag{1.4}
\]
where \( \|h''\|_{\infty} := \sup_{x \in \mathbb{R}} |h''(x)| \). See Rinott and Rotar (2003) for expansions of smooth test functions under local dependence, Fang et al. (2020) for tail probability approximations, and Braverman et al. (2022) for non-normal approximations. Kim and Park (2018) considered asymptotic expansions for Gaussian functionals and Fathi (2021) considered asymptotic expansions using higher-order Stein kernels and obtained Wasserstein bounds for normal approximations.

In summary, to obtain a rate faster than \( n^{-1/2} \) in asymptotic expansions using Stein’s method, studies have thus far assumed that (1) the test function is smooth, or (2) the random variable is continuous. Thus, how to recover the classical Edgeworth expansion result has remained an open problem. In this paper, we take the first step toward solving this problem by proving the following two-term Edgeworth expansion theorems using Stein’s method.

**Theorem 1.1** (Continuous two-term Edgeworth expansion). Let \( n \geq 1 \) and \( X_1, \ldots, X_n \) be i.i.d. with \( \mathbb{E}X_1 = 0, \mathbb{E}X_1^2 = 1, \mathbb{E}X_1^3 = \gamma, \mathbb{E}X_1^4 < \infty \). Suppose \( X_1 \) has a compactly supported continuous component with density bounded away from 0. Let \( W = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i \). Then, for all \( h : \mathbb{R} \to \mathbb{R} \) with \( \|h\|_{\infty} \leq 1 \), we have
\[
\left| \mathbb{E}h(W) - \mathbb{E}h(Z) - \frac{\gamma}{6\sqrt{n}} \mathbb{E}\left[ (Z^3 - 3Z)h(Z) \right] \right| \leq \frac{C}{n}, \tag{1.5}
\]
where \( Z \sim N(0,1) \) and \( C \) is a constant not depending on \( n \). When \( h(w) = 1_{\{w \leq x\}} \), the correction term is
\[
\frac{\gamma}{6\sqrt{2\pi n}} \int_{-\infty}^{x} (y^3 - 3y)e^{-y^2/2}dy,
\]
which is the same as in the classical Edgeworth expansion.
Although our condition is stronger than the Cramér condition, it is a natural condition because we do not use characteristic functions in the proof. It is possible but highly tedious to track the dependence of $C$ on the distribution of $X_1$, as it depends, for example, on the probability $p$ that $X_1$ takes the continuous component and on the density function of the continuous component. Therefore, we do not pursue this in this paper.

The difficulty in proving (1.3) using Stein’s method lies in the fact that using Taylor’s expansion unavoidably results in an error of order $1/\sqrt{n}$. To prove (1.5), we use an approach involving (a) repeated use of Stein equations, (b) Stein identities via Stein kernels, and (c) a replacement argument. Asymptotic expansion using Stein kernels was considered by Fathi (2021). However, Fathi’s approach does not appear to work for the expansion of distribution functions because (1) the existence of a Stein kernel requires the random variable to be continuous, and (2) his definition of higher-order Stein kernels involves higher-order derivatives of Stein equation solutions, which lack sufficient regularity.

Next, we consider the discrete case.

**Theorem 1.2** (Discrete two-term Edgeworth expansion). Let $n \geq 1$ and $X_1, \ldots, X_n$ be i.i.d., integer valued random variables with $EX_1 = \mu$, $\text{Var}(X_1) = \sigma^2 > 0$, $E(X_1 - \mu)^3 = \sigma^3 \gamma_1, \text{EX}_1^4 < \infty$. Suppose the support of $X_1$ is $\{s_0, s_1, s_2, \ldots\}$ such that the greatest common divisor of $\{|s_i - s_0|, i \geq 1\}$ is 1. Let $W = \frac{1}{\sigma \sqrt{n}} \sum_{i=1}^{n} (X_i - \mu)$. Then, for all $h : \mathbb{R} \to \mathbb{R}$ such that $\|h\|_{\infty} \leq 1$ and $h(x)$ equals a constant (depending on $z$) for $x \in (z - \frac{1}{2\sigma \sqrt{n}}, z + \frac{1}{2\sigma \sqrt{n}})$ for any $z$ in the support of $W$, we have

$$\left| Eh(W) - Eh(Z) - \frac{\gamma}{6\sqrt{n}} \mathbb{E}[(Z^3 - 3Z)h(Z)] \right| \leq \frac{C}{n},$$

(1.6)

where $Z \sim N(0,1)$ and $C$ is a constant not depending on $n$.

Note that the restriction on the test functions $h$ in Theorem 1.2 (i.e., the constants $\pm 1/2$) is necessary. As in continuity correction, changing them to other constants will increase the error rate to $1/\sqrt{n}$.

Although we focus on two-term Edgeworth expansion for the i.i.d. case in dimension one, our approach may work for multivariate approximations, for even higher-order expansions, and for some dependent cases. See Remarks 2.1–2.3 for related discussions.

## 2 Proofs

In this section, we prove Theorems 1.1 and 1.2. For the sake of logical flow, we leave some standard computations to Section 3.

### 2.1 Stein kernel

For both proofs of Theorems 1.1 and 1.2, we need the concept of Stein kernel. The term “Stein kernel” first appeared in Ledoux et al. (2015), but the concept goes back to Stein (1986), Chatterjee (2009) and Nourdin and Peccati (2009).
Definition 2.1 (Stein kernel, Saumard (2019)). Let $W$ be a continuous random variable with mean $\mu$, a connected support $(a, b)$, and positive density $p$ in $(a, b)$. The function $\tau : \mathbb{R} \to \mathbb{R}$ is the Stein kernel for $W$ if

$$
\mathbb{E}[(W - \mu)f(W)] = \mathbb{E}[\tau(W)f'(W)]
$$

(2.1)

for any differentiable function $f : \mathbb{R} \to \mathbb{R}$ such that the above expectations exist and $\tau p|_a^b = 0$.

It is known that (cf. Saumard (2019)) $\tau$ can be taken as

$$
\tau(x) = \frac{1}{p(x)} \int_x^\infty (y - \mu)p(y)dy, \quad x \in (a, b).
$$

(2.2)

If, in addition, $\text{Var}(W) = \sigma^2$, then by choosing $f(w) = w$ in (2.1), we find that $\mathbb{E}\tau(W) = \sigma^2$.

Remark 2.1. The existence of a Stein kernel in multi-dimensions is more complicated (cf. Courtade et al. (2019); Fathi (2019, 2021)). Courtade et al. (2019) proved that Stein kernel exists if a probability measure on $\mathbb{R}^d$ satisfies a Poincaré inequality. Therefore, it may be possible to use our approach for multivariate Edgeworth expansions, assuming that the summand $X_i$ has a component uniformly distributed on a Euclidean ball in $\mathbb{R}^d$.

2.2 Proof of Theorem 1.1

In both proofs of Theorems 1.1 and 1.2, we will need the following Edgeworth expansion result for sums of i.i.d. random variables having Stein kernels.

Lemma 2.1 (Stein kernel bound). Let $n \geq 1$ and $X_1, \ldots, X_n$ be i.i.d. with $\mathbb{E}X_1 = 0, \mathbb{E}X_1^2 = 1, \mathbb{E}X_1^4 = \gamma, \mathbb{E}X_1^4 < \infty$. Suppose $X_1$ has a Stein kernel $\tau_1$ such that $\mathbb{E}[\tau_1^2(X_1)] < \infty$. Let $W = \frac{1}{\sqrt{n}} \sum_{i=1}^n X_i$. Then, for all $h : \mathbb{R} \to \mathbb{R}$ with $\|h\|_\infty \leq 1$, we have

$$
\left| \mathbb{E}h(W) - \mathbb{E}h(Z) - \frac{\gamma}{6\sqrt{n}}\mathbb{E}[(Z^3 - 3Z)h(Z)] \right| \leq \frac{C}{n} \left( \mathbb{E}[X_1^4] + \mathbb{E}[\tau_1^2(X_1)] \right),
$$

(2.3)

where $Z \sim N(0, 1)$ and $C$ is a universal constant.

Proof of Lemma 2.1. We use $C$ to denote universal constants in this proof. Its value may differ from line to line. Let $f$ be the bounded solution to the Stein equation for $Z \sim N(0, 1)$:

$$
f'(w) - wf(w) = h(w) - \mathbb{E}h(Z).
$$

(2.4)

It is known that (cf. Chen et al. (2011, Lemma 2.4))

$$
f(w) = e^{w^2/2} \int_{-\infty}^w (h(x) - \mathbb{E}h(Z))e^{-x^2/2}dx
$$

(2.5)

and

$$
\|f\|_\infty \leq \sqrt{\pi/2}\|h - \mathbb{E}h(Z)\|_\infty, \quad \|f'\|_\infty \leq 2\|h - \mathbb{E}h(Z)\|_\infty.
$$

(2.6)
Because $X_1$ has Stein kernel $\tau_1$, we have

$$\mathbb{E}[Wf(W)] = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \mathbb{E}[X_i f(W)] = \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}[\tau_1(X_i) f'(W)].$$

This implies

$$\mathbb{E}h(W) - \mathbb{E}h(Z) = \mathbb{E}f'(W) - \mathbb{E}[Wf(W)] = \mathbb{E}[(1 - \tau)f'(W)],$$

where

$$\tau := \frac{1}{n} \sum_{i=1}^{n} \tau_1(X_i).$$

Therefore, using $\mathbb{E}\tau = \text{Var}(W) = 1$, we have

$$|\mathbb{E}h(W) - \mathbb{E}h(Z)| \leq \frac{2}{\sqrt{n}} \|h - \mathbb{E}h(Z)\|_{\infty} \sqrt{\mathbb{E}[\tau_1^2(X_1)]}. \quad (2.8)$$

This bound is well known in the literature (cf. Chatterjee (2009, Lemma 5.3)). To obtain higher-order expansions, motivated by Kim and Park (2018, Eq.(33)), we make repeated use of the Stein equation as follows. Recall that $f$ is the bounded solution to (2.4). Let $g$ be the bounded solution to the Stein equation

$$g'(w) - wg(w) = f'(w) - \mathbb{E}f'(Z). \quad (2.9)$$

From (2.6) and the condition that $\|h\|_{\infty} \leq 1$, we have

$$\|g\|_{\infty} \leq C, \quad \|g'\|_{\infty} \leq C.$$

Continuing from (2.7) and using $\mathbb{E}\tau = 1$ and (2.9), we have

$$\mathbb{E}h(W) - \mathbb{E}h(Z) = \mathbb{E}[(1 - \tau)(f'(W) - \mathbb{E}f'(Z))] = \mathbb{E}[(1 - \tau)(g'(W) - Wg(W))]. \quad (2.10)$$

Denote $\tau^{(i)} := \tau - \tau_1(X_i)/n$. We have

$$\mathbb{E}[(1 - \tau)Wg(W)]
\begin{align*}
&= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \mathbb{E}[(\frac{n-1}{n} - \tau^{(i)})X_ig(W)] + \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \mathbb{E}[(\frac{1}{n} - \frac{\tau_1(X_i)}{n})X_ig(W)] \\
&= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \mathbb{E}[(\frac{n-1}{n} - \tau^{(i)})\tau_1(X_i)g'(W)] + \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \mathbb{E}[(\frac{1}{n} - \frac{\tau_1(X_i)}{n})X_ig(W)] \\
&= \mathbb{E}[(1 - \tau)g'(W)] - \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \mathbb{E}[(\frac{1}{n} - \frac{\tau_1(X_i)}{n})\tau_1(X_ig)(W)] + \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \mathbb{E}[(\frac{1}{n} - \frac{\tau_1(X_i)}{n})X_ig(W)],
\end{align*}
$$

(2.11)

where we used the definition of Stein kernel in the second equation. From the boundedness of $g'$ and $\mathbb{E}\tau_1(X_i) = \text{Var}(X_i) = 1$, we have

$$\left|\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \mathbb{E}[(\frac{1}{n} - \frac{\tau_1(X_i)}{n})\tau_1(X_ig)(W)]\right| \leq \frac{C}{n} \mathbb{E}[\tau_1^2(X_1)]. \quad (2.12)$$
From the definition of Stein kernel, we have $\tau_1 \geq 0$, $\mathbb{E}[X_i \tau_1(X_i)] = \mathbb{E}[X_i^2]/2 = \gamma/2$ and $\mathbb{E}[X_i^2 \tau_1(X_i)] = \mathbb{E}[X_i^4]/3$. Therefore, using independence, we have

$$\left| \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \mathbb{E}[(\frac{1}{n} - \frac{\tau_1(X_i)}{n})X_i g(W)] + \frac{\gamma}{2\sqrt{n}} \mathbb{E}(W) \right|$$

$$= \left| \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \mathbb{E}[(X_i - X_i \tau_1(X_i) + \gamma/2 (g(W) - g(W^{(i)}))) \right|$$

$$\leq \frac{C}{n} \mathbb{E}[X_i^4],$$

(2.13)

where $W^{(i)} := W - X_i/\sqrt{n}$ and we used the boundedness of $g'$ and the Cauchy-Schwarz inequality in the last inequality. From (2.10)–(2.13), we have

$$\left| \mathbb{E}h(W) - \mathbb{E}h(Z) - \frac{\gamma}{2\sqrt{n}} \mathbb{E}g(W) - \mathbb{E}[(1 - \tau)^2 g'(W)] \right| \leq \frac{C}{n} \left( \mathbb{E}[X_i^4] + \mathbb{E}[\tau_i^2(X_i)] \right).$$

(2.14)

Using the boundedness of $g$, (2.8), $ab \leq (a^2 + b^2)/2$ and the Cauchy-Schwarz inequality, we have

$$\left| \frac{\gamma}{2\sqrt{n}} \mathbb{E}[g(W) - g(Z)] \right| \leq \frac{C}{n} \left( \mathbb{E}[X_i^4] + \mathbb{E}[\tau_i^2(X_i)] \right).$$

(2.15)

Using the boundedness of $g'$ and $\mathbb{E}\tau = 1$, we have

$$\left| \mathbb{E}[(1 - \tau)^2 g'(W)] \right| \leq \frac{C}{n} \mathbb{E}[\tau_i^2(X_i)].$$

(2.16)

Finally, (2.3) follows from (2.14)–(2.16) and

$$\mathbb{E}g(Z) = \frac{1}{3} \mathbb{E}[(Z^3 - Z^2)h(Z)].$$

(2.17)

The last equation is proved in Section 3.1 by a standard computation. □

**Remark 2.2.** As indicated by Kim and Park (2018, Eq.(37)), further repetitions of arguments in the above proof can be made to obtain even higher-order expansions. We do not pursue these in this paper as they unavoidably require tedious notation.

To prove Theorem 1.1, we divide $X_1$ into two components. We use Lemma 2.1 to deal with the continuous component. It then becomes a problem of asymptotic expansion of the expectation of a smooth test function of the remaining components. Finally, we use Lemma 2.1 again to approximate a sum of Gaussian mixtures by a normal distribution.

**Proof of Theorem 1.1.** In this proof, we use $C$ to denote positive constants that depend only on the distribution of $X_1$ and may differ from line to line. We use $O(1)$ to denote a quantity that is bounded by $C$ in absolute value. Without loss of generality, we assume that with probability $p > 0$, $X_1 = U_1$ and

$$\mathbb{E}U_1 = 0, U_1 \text{ is continuous, compactly supported and has density bounded away from 0.}$$

(2.18)
This centering can be achieved by a finite convolution that does not affect the error rate (see details in Section 3.2). We remark that assuming \( \mathbb{E}U_1 = 0 \) is only for convenience, as otherwise the term \( III \) below would result in an additional non-negligible term. Let \( \tilde{U}_1 \) be the other component of \( X_1 \), which may be discrete or even singular. That is,

\[
X_1 = \begin{cases} 
U_1, & \text{with probability } p, \\
\tilde{U}_1, & \text{with probability } 1 - p.
\end{cases}
\]

Suppose

\[
\mathbb{E}U_1 = 0, \quad \mathbb{E}U_1^2 = \sigma_1^2, \quad \mathbb{E}U_1^3 = \gamma_1, \quad \mathbb{E}\tilde{U}_1 = 0, \quad \mathbb{E}\tilde{U}_1^2 = \sigma_2^2, \quad \mathbb{E}\tilde{U}_1^3 = \gamma_2.
\]

Assume \( \sigma_2^2 > 0 \); otherwise, Theorem 1.1 follows directly from Lemma 2.1. (Although we do not do it here, an inspection of Section 3.2 shows that we can further assume \( \mathbb{E}U_1^3 = 0 \). This would slightly simplify the following proof by making \( \gamma_1 = 0 \).) Note that

\[
\mathbb{E}X_1 = 0, \quad \mathbb{E}X_1^2 = 1 = p\sigma_1^2 + (1 - p)\sigma_2^2, \quad \mathbb{E}X_1^3 = \gamma = p\gamma_1 + (1 - p)\gamma_2.
\]

Let

\[
Y_1 = \begin{cases} 
V_1 \sim N(0, \sigma_1^2), & \text{with probability } p, \\
\tilde{V}_1 \sim N(0, \sigma_2^2), & \text{with probability } 1 - p.
\end{cases} \tag{2.19}
\]

Let \( U_1, \ldots, U_n, \tilde{U}_1, \ldots, \tilde{U}_n, V_1, \ldots, V_n, \tilde{V}_1, \ldots, \tilde{V}_n \) be jointly independent, where \( U_1, \ldots, U_n \) are identically distributed (the same applies for \( \tilde{U} \)'s, \( V \)'s and \( \tilde{V} \)'s, respectively). Let \( Z, \tilde{Z} \sim N(0, 1) \) be independent standard normal variables and independent of everything else.

Let \( L \sim Bin(n, p) \) be a binomial random variable independent of everything else. Then, \( W \) has the same distribution as \( (U_1 + \cdots + U_L + \tilde{U}_{L+1} + \cdots + \tilde{U}_n)/\sqrt{n} \). We write \( \mathbb{E}h(W) - \mathbb{E}h(Z) \) into three terms as follows:

\[
\mathbb{E}h(W) - \mathbb{E}h(Z) = \mathbb{E}h\left( \frac{U_1 + \cdots + U_L + \tilde{U}_{L+1} + \cdots + \tilde{U}_n}{\sqrt{n}} \right) - \mathbb{E}h\left( \frac{V_1 + \cdots + V_L + \tilde{V}_{L+1} + \cdots + \tilde{V}_n}{\sqrt{n}} \right)
\]

\[
+ \mathbb{E}h\left( \frac{V_1 + \cdots + V_L + \tilde{V}_{L+1} + \cdots + \tilde{V}_n}{\sqrt{n}} \right) - \mathbb{E}h\left( \frac{V_1 + \cdots + V_L + \tilde{V}_{L+1} + \cdots + \tilde{V}_n}{\sqrt{n}} \right) - \mathbb{E}h(Z)
\]

\[
= I + II + III. \tag{2.20}
\]

Note that \( (V_1 + \cdots + V_L + \tilde{V}_{L+1} + \cdots + \tilde{V}_n)/\sqrt{n} \) has the same distribution as \( (Y_1 + \cdots + Y_n)/\sqrt{n} \), where \( Y_1, \ldots, Y_n \) are i.i.d. with distribution (2.19). From the Stein kernel bound (2.3), \( \mathbb{E}Y_1^3 = 0, \mathbb{E}Y_1^4 < \infty \) and the boundedness of its Stein kernel (this can be easily checked from the expression (2.2), see Lemma 3.4), we have

\[
|III| \leq C \frac{n}{n}. \tag{2.21}
\]
Given $L = l$ such that $|l - np| \leq np/2$, we have, by conditioning on the $\tilde{U}$'s and using the Stein kernel bound (2.3) and the condition (2.18),

$$
\begin{align*}
\mathbb{E} h\left(\frac{U_1 + \cdots + U_i + \tilde{U}_{i+1} + \cdots + \tilde{U}_n}{\sqrt{n}}\right) - \mathbb{E} h\left(\frac{V_1 + \cdots + V_i + \tilde{U}_{i+1} + \cdots + \tilde{U}_n}{\sqrt{n}}\right) \\
= \mathbb{E} h\left(\frac{U_1 + \cdots + U_i}{\sigma_1 \sqrt{l}} \cdot \frac{\sigma_1 \sqrt{l}}{\sqrt{n}} + \frac{\tilde{U}_{i+1} + \cdots + \tilde{U}_n}{\sqrt{n}}\right) - \mathbb{E} h\left(Z \cdot \sigma_1 \sqrt{l} + \frac{\tilde{U}_{i+1} + \cdots + \tilde{U}_n}{\sqrt{n}}\right) \\
= \frac{\gamma_1}{6\sigma_1^2 \sqrt{l}} \mathbb{E}\left[(Z^3 - 3Z)h(Z \cdot \sigma_1 \sqrt{l} + \frac{\tilde{U}_{i+1} + \cdots + \tilde{U}_n}{\sqrt{n}})\right] + O\left(\frac{1}{n}\right). \\
\end{align*}
$$

(2.22)

By regarding the last expectation as an expectation of a smooth test function (because of convolution with normal) of the $\tilde{U}$'s, we have (see Section 3.3)

$$
(2.22) = \frac{\gamma_1}{6\sigma_1^2 \sqrt{l}} \mathbb{E}\left[(Z^3 - 3Z)h(Z \cdot \sigma_1 \sqrt{l} + \tilde{Z} \cdot \frac{\sigma_2 \sqrt{n} - L}{\sqrt{n}})\right] + O\left(\frac{1}{n}\right). \\
$$

(2.23)

Because $L \sim \text{Bin}(n, p)$, the event $\{|L - np| > np/2\}$ has probability $O(1/n)$. Therefore,

$$
I = \frac{1}{6\sqrt{n}} \mathbb{E}\left[\frac{\gamma_1 \sqrt{n}}{\sigma_1 + \sqrt{L}}(Z^3 - 3Z)h(Z \cdot \frac{\sigma_1 \sqrt{l}}{\sqrt{n}} + \tilde{Z} \cdot \frac{\sigma_2 \sqrt{n} - L}{\sqrt{n}})1_{\{L - np \leq np/2\}}\right] + O\left(\frac{1}{n}\right).
$$

By a similar reasoning (conditioning on $L$ and using the smooth function expansion for the $\tilde{U}$'s by applying either (1.4) or a Lindeberg swapping argument), we have

$$
II = \frac{1}{6\sqrt{n}} \mathbb{E}\left[\frac{\gamma_2 \sqrt{n}}{\sigma_1 + \sqrt{L}}(Z^3 - 3Z)h(Z \cdot \frac{\sigma_1 \sqrt{l}}{\sqrt{n}} + \tilde{Z} \cdot \frac{\sigma_2 \sqrt{n} - L}{\sqrt{n}})1_{\{L - np \leq np/2\}}\right] + O\left(\frac{1}{n}\right).
$$

Denote

$$
\sigma_L^2 := \frac{L \sigma_1^2 + (n - L) \sigma_2^2}{n}, \quad \gamma_L := \frac{L \gamma_1 + (n - L) \gamma_2}{n}.
$$

Using Gaussian integration by parts and then combining the two independent Gaussian variables (and by approximating $h$ using arbitrarily close smooth functions in the intermediate step), we have

$$
\begin{align*}
I + II &= -\frac{1}{6\sqrt{n}} \mathbb{E}\left[\gamma_L h'''(\sigma_L Z)1_{\{L - np \leq np/2\}}\right] + O\left(\frac{1}{n}\right) \\
&= -\frac{1}{6\sqrt{n}} \mathbb{E}\left[\frac{\gamma_L}{\sigma_L^3}(Z^3 - 3Z)h(\sigma_L Z)1_{\{L - np \leq np/2\}}\right] + O\left(\frac{1}{n}\right) \\
&= -\frac{1}{6\sqrt{n}} \mathbb{E}\left[\frac{\gamma_L}{\sigma_L^3} h_2(\sigma_L)\right] + O\left(\frac{1}{n}\right) \\
\end{align*}
$$

(2.24)

where $h_2(x) = \mathbb{E}[Z^3 - 3Z)h(xZ)]$ and we removed the indicator in the last equation because the event $\{|L - np| \leq np/2\}$ occurs with overwhelming probability. Using Gaussian
integration by parts, we have $|h_2(x)| \leq C$ and $|h_2'(x)| \leq \frac{C}{x}$ for $x > 0$ (cf. (3.4)). According to the definition of $\sigma_L^2$, we have $\sigma_L^2 \geq \min\{\sigma_1^2, \sigma_2^2\} > 0$. Moreover,

$$\sigma_L^2 - 1 = (\sigma_1^2 - \sigma_2^2) \frac{L - np}{n},$$

$$\gamma_L - \gamma = (\gamma_1 - \gamma_2) \frac{L - np}{n}.$$  

By applying Taylor's expansion to $\gamma_L h_2(\sigma_L)/\sigma_L^3$ with respect to $\sigma_L$ ($\gamma_L$ resp.) at point 1 ($\gamma$ resp.) and then taking expectation with respect to $L$, we obtain

$$I + II = \frac{\gamma}{6\sqrt{n}} \mathbb{E}[(Z^3 - 3Z)h(Z)] + O\left(\frac{1}{n}\right).$$

(2.25)

Combining (2.20), (2.21) and (2.25), we obtain (1.5).

**Remark 2.3.** Our approach may also work for some dependent cases. For example, following the proof of Lemma 2.1, we may obtain an Edgeworth expansion result in normal approximation of multilinear forms of independent random variables having Stein kernels. Then, combining with the replacement argument in the proof of Theorem 1.1, we may deal with multilinear forms of independent random variables having continuous components.

### 2.3 Proof of Theorem 1.2

In this proof, we use $C$ to denote positive constants that depend only on the distribution of $X_1$ and may differ from line to line. We use $O(1)$ to denote a quantity that is bounded by $C$ in absolute value. We divide the proof into four steps following the approach used in the proof of Theorem 1.1. Note that although we deal with a discrete case here, we still need the Stein kernel bound in Lemma 2.1 for approximating a sum of Gaussian mixtures.

**Step 1.** As in the proof of Theorem 1.1, we assume without loss of generality that $X_1$ has a Bernoulli $Ber(1/2)$ component. This is possible for $X_1 + \cdots + X_m - z$ with a positive integer $m$ and an integer $z$ because of the assumption on the support of $X_1$ (see Section 3.6). Such finite grouping does not affect the error rate following similar arguments as (3.11)–(3.16) in Section 3.2. Under this assumption, we have

$$X_1 = \begin{cases} 
U_1 \sim Ber(1/2), & \text{with probability } p, \\
\bar{U}_1 \in \mathbb{Z}, & \text{with probability } 1 - p,
\end{cases}$$

and

$$\mathbb{E}U_1 = \frac{1}{2}, \quad \text{Var}(U_1) = \frac{1}{4}, \quad \mathbb{E}(U_1 - \frac{1}{2})^3 = 0,$$

$$\mathbb{E}\bar{U}_1 =: \mu_2, \quad \text{Var}(\bar{U}_1) =: \sigma_2^2, \quad \mathbb{E}((\bar{U}_1 - \mu_2)^3) =: \gamma_2.$$  

$$\mathbb{E}(X_1 - \mu)^3 = p\left(\frac{3}{4} - \frac{1}{2} - \mu + \frac{1}{2} - \mu)^3\right) + (1 - p)(\gamma_2 + 3\sigma_2^2(\mu_2 - \mu) + (\mu_2 - \mu)^3).$$

(2.26)

Note that if $\sigma_2^2 = 0$, then Theorem 1.2 follows directly from Lemma 2.2 below. Therefore, we assume that $\sigma_2^2 > 0$ in the following. We remark that, unlike (2.18), it appears
impossible to center $U_1$ in the discrete case. This makes the current proof slightly more technical than that of Theorem 1.1. Let

$$Y_1 = \begin{cases} V_1 \sim N\left(\frac{1}{2}, \frac{1}{4}\right), & \text{with probability } p, \\ \tilde{V}_1 \sim N(\mu_2, \sigma_2^2), & \text{with probability } 1 - p. \end{cases}$$  \hspace{1cm} (2.27)$$

Let $U_1, \ldots, U_n, \tilde{U}_1, \ldots, \tilde{U}_n, V_1, \ldots, V_n, \tilde{V}_1, \ldots, \tilde{V}_n$ be jointly independent, where $U_1, \ldots, U_n$ are identically distributed (the same applies for $\tilde{U}$’s, $V$’s and $\tilde{V}$’s, respectively). Let $Z, \tilde{Z} \sim N(0, 1)$ be independent standard normal variables and independent of everything else.

Let $L \sim Bin(n, p)$ be independent of everything else. Then $W$ has the same distribution as $(U_1 + \cdots + U_L + \tilde{U}_{L+1} + \cdots + \tilde{U}_n - n\mu)/(\sigma_1\sqrt{n})$. We write $E_h(W) - E_h(Z)$ into three terms as follows:

$$E_h(W) - E_h(Z) = E_h\left(\frac{U_1 + \cdots + U_L + \tilde{U}_{L+1} + \cdots + \tilde{U}_n - n\mu}{\sigma_1\sqrt{n}}\right) - E_h\left(\frac{V_1 + \cdots + V_L + \tilde{V}_{L+1} + \cdots + \tilde{V}_n - n\mu}{\sigma_1\sqrt{n}}\right)$$

$$+ E_h\left(\frac{V_1 + \cdots + V_L + \tilde{V}_{L+1} + \cdots + \tilde{V}_n - n\mu}{\sigma_1\sqrt{n}}\right) - E_h(Z)$$

$$= I + II + III.$$  \hspace{1cm} (2.28)

Note that $(V_1 + \cdots + V_L + \tilde{V}_{L+1} + \cdots + \tilde{V}_n - n\mu)/(\sigma_1\sqrt{n})$ has the same distribution as $(Y_1 + \cdots + Y_n - n\mu)/(\sigma_1\sqrt{n})$, where $Y_1, \ldots, Y_n$ are i.i.d. with distribution (2.27). From the Stein kernel bound (2.3), $EY_1^4 < \infty$ and the boundedness of its Stein kernel (see Lemma 3.4), we have

$$III = \frac{E(Y_1 - \mu)^3}{6\sqrt{n}\sigma_1^3}E[|Z^3 - 3Z|h(Z)] + O\left(\frac{1}{n}\right).$$  \hspace{1cm} (2.29)

It is straightforward to compute that

$$E(Y_1 - \mu)^3 = p\left(\frac{3}{4}\left(\frac{1}{2} - \mu\right) + \left(\frac{1}{2} - \mu\right)^3\right) + (1 - p)(3\sigma_2^2(\mu_2 - \mu) + (\mu_2 - \mu)^3).$$  \hspace{1cm} (2.30)

**Step 2.** To deal with $I$, we use the following lemma, which is proved in Section 3.4.

**Lemma 2.2.** Let $l \geq 1$, $S \sim Bin(l, 1/2)$ and $Z \sim N(0, 1)$. Let $h : \mathbb{R} \to \mathbb{R}$ be such that $|h(x)| \leq 1$ and $h(x)$ equals a constant in $(z - 1/2, z + 1/2)$ for any $z \in \mathbb{Z}$. Then we have

$$E_h(S) - E_h\left(\frac{l}{2} + Z\sqrt{\frac{l}{4}}\right) = O\left(\frac{1}{l}\right).$$
Given $L = l$ such that $|l - np| \leq np/2$, we have, by conditioning on the $\tilde{U}$’s and using Lemma 2.2 and the condition on $h$ in Theorem 1.2,

$$
\mathbb{E}h\left(\frac{U_1 + \cdots + U_i + \tilde{U}_{i+1} + \cdots + \tilde{U}_n - n\mu}{\sigma\sqrt{n}}\right) - \mathbb{E}h\left(\frac{V_1 + \cdots + V_i + \tilde{U}_{i+1} + \cdots + \tilde{U}_n - n\mu}{\sigma\sqrt{n}}\right)
= O\left(\frac{1}{n}\right).
$$

Because $L \sim \text{Bin}(n, p)$, the event $\{|L - np| > np/2\}$ has probability $O(1/n)$. Therefore,

$$
I = O\left(\frac{1}{n}\right).
$$

(2.31)

**Step 3.** By a similar reasoning as in estimating $II$ in the proof of Theorem 1.1 (conditioning on $L$ and using the smooth function expansion for the $\tilde{U}$’s by either (1.4) or a Lindeberg swapping argument), we have

$$
II = \frac{\gamma_2}{6\sigma_2^2} \mathbb{E}\left[\frac{1}{\sqrt{n} - L} (\tilde{Z}^3 - 3\tilde{Z}) h\left(\frac{\sqrt{L}}{2\sigma\sqrt{n}} Z + \frac{\sigma_2\sqrt{n - L}}{\sigma\sqrt{n}} \tilde{Z} + \mu_L\right) 1_{\{|L-np|\leq np/2\}}\right] + O\left(\frac{1}{n}\right),
$$

(2.32)

where

$$
\mu_L = \frac{L/2 - \mu L}{\sigma\sqrt{n}} + \frac{(n - L)\mu_2 - (n - L)\mu}{\sigma\sqrt{n}}.
$$

(2.33)

Given $L = l$ such that $|l - np| \leq np/2$, let

$$
h_1(x) = \mathbb{E}\left[(\tilde{Z}^3 - 3\tilde{Z}) h\left(x + \frac{\sigma_2\sqrt{n - L}}{\sigma\sqrt{n}} \tilde{Z} + \mu_l\right)\right],
$$

and

$$
h_2(r) = \mathbb{E}h_1(rZ).
$$

Because $\|h_1\|_\infty \leq C$, from (3.4) and Taylor’s expansion, we have

$$
h_2\left(\frac{\sqrt{L}}{2\sigma\sqrt{n}}\right) = h_2\left(\frac{\sqrt{p}}{2\sigma}\right) + O(1) \frac{l - np}{n}.
$$

(2.34)

Plugging (2.34) into (2.32), taking expectation with respect to $L$, and using $\mathbb{E}|L - np| = O(\sqrt{n})$, we obtain

$$
II = \frac{\gamma_2}{6\sigma_2^2} \mathbb{E}\left[\frac{1}{\sqrt{n} - L} (\tilde{Z}^3 - 3\tilde{Z}) h\left(\frac{\sqrt{p}}{2\sigma} Z + \frac{\sigma_2\sqrt{n - L}}{\sigma\sqrt{n}} \tilde{Z} + \mu_L\right) 1_{\{|L-np|\leq np/2\}}\right] + O\left(\frac{1}{n}\right).
$$

(2.35)

Let $h_3(x) = \mathbb{E}h\left(\frac{\sqrt{p}}{2\sigma} Z + x\right)$. By the Gaussian integration by parts formula, we have

$$
II = -\frac{\gamma_2}{6\sqrt{n}\sigma^3} \mathbb{E}\left[\frac{n - L}{n} h_3''\left(\frac{\sigma_2\sqrt{n - L}}{\sigma\sqrt{n}} \tilde{Z} + \mu_L\right) 1_{\{|L-np|\leq np/2\}}\right] + O\left(\frac{1}{n}\right).
$$

(2.36)
Similar to (2.35), we obtain

$$H = -\frac{\gamma_2}{6\sqrt{n}\sigma^3} E\left[\frac{n - L}{n} h''' \left(\frac{\sigma_2\sqrt{n} - np}{\sigma\sqrt{n}} \hat{Z} + \mu_L\right) 1_{\{|L - np| \leq np/2\}}\right] + O\left(\frac{1}{n}\right)$$

$$= -\frac{\gamma_2}{6\sqrt{n}\sigma^3} E\left[\frac{n - L}{n} h''' \left(\frac{\sigma_2\sqrt{n} - np}{\sigma\sqrt{n}} \hat{Z} + \mu_L\right)\right] + O\left(\frac{1}{n}\right)$$

$$= -\frac{\gamma_2(1 - p)}{6\sqrt{n}\sigma^3} E\left[h''' \left(\frac{\sigma_2\sqrt{n} - np}{\sigma\sqrt{n}} \hat{Z} + \mu_L\right)\right] + O\left(\frac{1}{n}\right)$$

(2.37)

where we use the fact that the event \(|L - np| > np/2\) has probability \(O(1/n)\) to drop the indicator in the second equality, and we used the boundedness of \(h'''\) (cf. (3.3)) and the fact that \(E|L - np| = O(\sqrt{n})\) to drop the second term in the last equation.

From (2.33) and \(\mu = p/2 + (1 - p)\mu_2\), we have

$$\mu_L = \left(\frac{1}{2} - \mu_2\right) \frac{L - np}{\sigma\sqrt{n}} = \left(\frac{1}{2} - \mu_2\right) \frac{\sqrt{p(1 - p)}}{\sigma} \frac{L - np}{\sqrt{np(1 - p)}}.$$

Note that \(\mu_L\) converges in distribution to \(N(0, (1/2 - \mu_2)^2 p(1 - p)/\sigma^2)\) as \(n \to \infty\). Since \(h_3\) has bounded derivatives (cf. (3.3)), we have, by Lindeberg’s swapping argument,

$$H = -\frac{\gamma_2(1 - p)}{6\sqrt{n}\sigma^3} E h''' \left(\frac{\sigma_2\sqrt{n} - np}{\sigma\sqrt{n}} \hat{Z} + \left(\frac{1}{2} - \mu_2\right) \frac{\sqrt{p(1 - p)}}{\sigma} \hat{Z}\right) + O\left(\frac{1}{n}\right),$$

(2.38)

where \(\hat{Z} \sim N(0, 1)\) and \(\hat{Z}\) is independent of everything else. Because \(Z, \hat{Z}, \tilde{Z}\) are i.i.d. \(N(0, 1)\), by a straightforward computation, we have

$$\text{Var} \left(\frac{\sqrt{p}}{2\sigma} Z + \frac{\sigma_2\sqrt{n} - np}{\sigma\sqrt{n}} \tilde{Z} + \left(\frac{1}{2} - \mu_2\right) \frac{\sqrt{p(1 - p)}}{\sigma} \tilde{Z}\right) = 1.$$

Combining the three independent Gaussian variables and using Gaussian integration by parts (and approximating \(h\) by arbitrarily close smooth functions in the intermediate step), we have

$$H = -\frac{\gamma_2(1 - p)}{6\sqrt{n}\sigma^3} E h''' \left(\frac{\sqrt{p}}{2\sigma} Z + \frac{\sigma_2\sqrt{n} - np}{\sigma\sqrt{n}} \hat{Z} + \left(\frac{1}{2} - \mu_2\right) \frac{\sqrt{p(1 - p)}}{\sigma} \hat{Z}\right) + O\left(\frac{1}{n}\right)$$

$$= -\frac{\gamma_2(1 - p)}{6\sqrt{n}\sigma^3} E h''' (Z) + O\left(\frac{1}{n}\right)$$

$$= \frac{\gamma_2(1 - p)}{6\sqrt{n}\sigma^3} E[(Z^3 - 3Z) h (Z)] + O\left(\frac{1}{n}\right).$$

(2.39)
Step 4. From (2.26) and (2.30), we have
\[ \mathbb{E}(Y_1 - \mu)^3 + (1 - p)\gamma_2 = \mathbb{E}(X_1 - \mu)^3 = \sigma^3 \gamma. \]  
(2.40)
From (2.39), (2.29) and (2.40), we have
\[ II + III = \frac{\gamma}{6\sqrt{n}} \mathbb{E}[(Z^3 - 3Z)h(Z)] + O\left(\frac{1}{n}\right). \]  
(2.41)
Combining (2.28), (2.31) and (2.41), we obtain (1.6).

3 Appendix

3.1 Proof of (2.17)

From (2.5), the solution to (2.9) is
\[ g(w) = e^{w^2/2} \int_{-\infty}^{w} (f'(x) - \mathbb{E}f'(Z))e^{-x^2/2} dx. \]

By the integration by parts formula,
\[ \mathbb{E}g(Z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \int_{-\infty}^{w} (f'(x) - \mathbb{E}f'(Z))e^{-x^2/2} dx dw \]
\[ = \frac{1}{\sqrt{2\pi}} \left[ w \int_{-\infty}^{w} (f'(x) - \mathbb{E}f'(Z))e^{-x^2/2} dx \right]_{w=\infty}^{w=-\infty} \]
\[ - \frac{1}{\sqrt{2\pi}} \int_{w\in\mathbb{R}} w(f'(w) - \mathbb{E}f'(Z))e^{-w^2/2} dw \]
\[ = - \frac{1}{\sqrt{2\pi}} \int_{w\in\mathbb{R}} w f'(w)e^{-w^2/2} dw. \]
(3.1)
From (2.4) and (2.5), we have
\[ f'(w) = h(w) - \mathbb{E}h(Z) + we^{w^2/2} \int_{-\infty}^{w} (h(x) - \mathbb{E}h(Z))e^{-x^2/2} dx. \]

Plugging this expression into (3.1) and using integration by parts, we have
\[ \mathbb{E}g(Z) = - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} w^3 \int_{-\infty}^{w} (h(x) - \mathbb{E}h(Z))e^{-x^2/2} dx dw \]
\[ - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} w(h(w) - \mathbb{E}h(Z))e^{-w^2/2} dw \]
\[ = - \frac{1}{\sqrt{2\pi}} \left[ w(h(x) - \mathbb{E}h(Z))e^{-x^2/2} dx \right]_{w=\infty}^{w=-\infty} + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{w^3}{3} h(w)e^{-w^2/2} dw \]
\[ - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} wh(w)e^{-w^2/2} dw \]
\[ = \frac{1}{3} \mathbb{E}[(Z^3 - 3Z)h(Z)]. \]
(3.2)
This proves (2.17).
3.2 Proof of (2.18)

We will need the following two lemmas.

**Lemma 3.1.** Let \( h(x) \) be any bounded function and \( P(x) \) be any polynomial. Define \( h_1(x) = \mathbb{E}[h(\sigma Z + x)P(Z)] \) and \( h_2(\sigma) = \mathbb{E}[h(\sigma Z)P(Z)] \), where \( \sigma > 0 \) and \( Z \sim N(0, 1) \). Then for any \( k \geq 0, \) \( h_1(x) \) and \( h_2(\sigma) \) are \( k \)-th order differentiable and

\[
\left| \frac{d^k}{dx^k} h_1(x) \right| \leq \frac{C_{P,k} \|h\|_{\infty}}{\sigma^k}, \quad \forall x \in \mathbb{R}, \tag{3.3}
\]

\[
\left| \frac{d^k}{d\sigma^k} h_2(\sigma) \right| \leq \frac{\tilde{C}_{P,k} \|h\|_{\infty}}{\sigma^k}, \quad \forall x \in \mathbb{R}, \tag{3.4}
\]

where \( C_{P,k} \) and \( \tilde{C}_{P,k} \) are constants depending only on the polynomial \( P \) and \( k \).

**Proof of Lemma 3.1.** First, we rewrite \( h_1(x) \) as the integral form and use change of variable to obtain

\[
h_1(x) = \frac{1}{\sqrt{2\pi}} \int_{y \in \mathbb{R}} P(y) h(\sigma y + x) e^{-\frac{y^2}{2}} dy = \frac{1}{\sigma \sqrt{2\pi}} \int_{y \in \mathbb{R}} P \left( \frac{r - x}{\sigma} \right) h(r) e^{-\frac{(r-x)^2}{2\sigma^2}} dr. \tag{3.5}
\]

The case \( k = 0 \) follows from the boundedness of moments of \( Z \). For \( k = 1 \), taking first order derivative of \( h_1(x) \), we have

\[
\frac{d}{dx} h_1(x) = \frac{1}{\sigma^2 \sqrt{2\pi}} \int_{y \in \mathbb{R}} \left( -P' \left( \frac{r - x}{\sigma} \right) + P \left( \frac{r - x}{\sigma} \right) \frac{r - x}{\sigma} \right) h(r) e^{-\frac{(r-x)^2}{2\sigma^2}} dr
\]

\[
= \frac{1}{\sigma \sqrt{2\pi}} \int_{y \in \mathbb{R}} \left( -P'(y) + yP(y) \right) h(\sigma y + x) e^{-\frac{y^2}{2}} dy \tag{3.6}
\]

\[
= \frac{1}{\sigma} \mathbb{E} \left[ (-P'(Z) + ZP(Z)) h(\sigma Z + x) \right].
\]

From (3.6), we have

\[
\left| \frac{d}{dx} h_1(x) \right| \leq \frac{\|h\|_{\infty}}{\sigma} \mathbb{E} |P'(Z) + ZP(Z)|, \tag{3.7}
\]

and we complete the proof for the case \( k = 1 \). Following similar arguments, we get (3.3) for \( k \geq 2 \) and (3.4) for \( k \geq 0 \). \( \square \)

**Lemma 3.2.** Let \( X_1, \ldots, X_n \) be i.i.d. random variables with \( \mathbb{E} X_1 = 0 \), \( \text{Var}(X_1) = 1 \) and \( \mathbb{E}|X_1|^3 < \infty \). For \( h_1(x) \) defined in Lemma 3.1 and any \( \sigma_1 \geq 0 \), we have

\[
\left| \mathbb{E} h_1 \left( \frac{\sigma_1}{\sqrt{n}} \sum_{i=1}^n X_i \right) - \mathbb{E} h_1(\sigma_1 Z) \right| \leq \frac{C_P \|h\|_{\infty} \sigma_1^3}{\sigma^3 \sqrt{n}} \mathbb{E}(|X_1|^3 + |Z|^3), \tag{3.8}
\]

where \( Z \sim N(0, 1) \) and \( C_P \) is a constant depending only on the polynomial \( P \) in the definition of \( h_1 \).
Proof of Lemma 3.2. We prove this lemma by Lindeberg’s swapping argument. Let \( Z_1, \ldots, Z_n \) be i.i.d. \( N(0, 1) \) random variables. We have

\[
\left| \mathbb{E} h_1 \left( \frac{\sigma_1}{\sqrt{n}} \sum_{i=1}^{n} X_i \right) - \mathbb{E} h_1 (\sigma_1 Z) \right| \\
= \sum_{k=1}^{n} \left| \mathbb{E} h_1 \left( \frac{\sigma_1}{\sqrt{n}} \left( \sum_{i=1}^{k} X_i + \sum_{i=k+1}^{n} Z_i \right) \right) \right| - \mathbb{E} h_1 \left( \frac{\sigma_1}{\sqrt{n}} \left( \sum_{i=1}^{k-1} X_i + \sum_{i=k+1}^{n} Z_i \right) \right) \right| \\
\leq \frac{\sigma_3^3}{6n} \sup_{x \in \mathbb{R}} \left| \frac{d^3}{dx^3} h_1 (x) \right| \mathbb{E} (|X_1|^3 + |Z|^3).
\]

From (3.3), we complete the proof of this lemma. \( \square \)

Proof of (2.18). Suppose that \( X_1 \sim F(x) \) and \( F(x) = p_1 F_1(x) + (1 - p_1) F_2(x) \), where \( F_1(x) \) is compactly supported and has density \( p(x) \) bounded away from 0, but \( \int x dF_1(x) \) is not equal to 0 (otherwise (2.18) follows directly from the condition of Theorem 1.1).

Suppose that the support of \( F_1(x) \) is the interval \([a_1, a_2]\).

Case 1. If \( 0 \in (a_1, a_2) \), then by truncation, \( F_1(x) \) has a component \( U \) which satisfies (2.18). Furthermore, since \( U \) is also a component of \( F(x) \), we complete the proof of (2.18) in this case.

Case 2. If \( 0 \notin (a_1, a_2) \), without loss of generality, we assume that \( a_1 > 0 \). Since \( \mathbb{E} X_1 = 0 \), there must exist a negative constant \( b \) such that \( P(b - \varepsilon \leq X_1 \leq b) > 0 \) for any \( \varepsilon > 0 \). Let

\[
F_3(x) = \frac{F(x) 1_{\{x \leq b\}}}{F(b)},
\]

we then have

\[ F(x) = p_1 F_1(x) + p_3 F_3(x) + p_4 F_4(x), \]

where \( p_3 = F(b) \) and \( p_4 \) and \( F_4(x) \) are determined by this equation. Let \( m_1 \) be the smallest positive integer such that \( m_1 |a_1 - a_2| > 3|b| \). The support of \( F_1^{m_1} (x) \) is \([m_1 a_1, m_1 a_2]\), where \( * \) denotes convolutions. Let \( m_2 \) be the smallest positive integer such that \( m_2 |b| > m_1 a_1 \). Then we have

\[
F_3^{m_2} (m_2 b) - F_3^{m_2} (m_2 b - \varepsilon) > 0
\]

for any \( \varepsilon > 0 \), and

\[
(m_2 - 1)|b| \leq m_1 a_1 < m_2 |b| < m_1 a_2. \quad (3.10)
\]

By (3.10), there exists a component \( \tilde{F}(x) \) of \( F_1^{m_1} * F_3^{m_2} (x) \) such that \( \tilde{F}(x) \) is compactly supported, has density bounded away from 0 and the interior of the support of \( \tilde{F}(x) \) contains 0. Since \( F_1^{m_1} * F_3^{m_2} (x) \) is a component of \( F^{*(m_1+m_2)} (x) \), we have \( \tilde{F}(x) \) is also a component of \( F^{*(m_1+m_2)} (x) \).

By the same argument as that in Case 1 and let \( m = m_1 + m_2 \), we infer that \( F^{*m} (x) \) has a component \( U \) satisfying (2.18).

Let \( n = km + r \) for some integers \( k \geq 1 \) (Theorem 1.1 trivially holds for bounded \( n \)) and \( 0 \leq r < m \). The case \( r = 0 \) follows from the proof of Theorem 1.1. We now
consider the case \( r > 0 \). Let \( Z, \tilde{Z}, Z_1, Z_2, \ldots, Z_n \) be i.i.d. \( N(0, 1) \) random variables and independent of \( \{X_1, \ldots, X_n\} \). We have

\[
\mathbb{E} h(W) - \mathbb{E} h(Z) = \mathbb{E} h\left( \frac{1}{\sqrt{n}} \left( \sum_{i=1}^{km} X_i + \sum_{i=km+1}^{n} X_i \right) \right) - \mathbb{E} h\left( \frac{1}{\sqrt{n}} \left( \sum_{i=1}^{km} Z_i + \sum_{i=km+1}^{n} X_i \right) \right)
+ \mathbb{E} h\left( \frac{1}{\sqrt{n}} \left( \sum_{i=1}^{km} Z_i + \sum_{i=km+1}^{n} X_i \right) \right) - \mathbb{E} h\left( \frac{1}{\sqrt{n}} \left( \sum_{i=1}^{km} Z_i + \sum_{i=km+1}^{n} Z_i \right) \right)
= H_1 + H_2.
\]

For \( H_1 \), conditioning on \( \{X_{km+1}, \ldots, X_n\} \) first and applying Theorem 1.1 (i.e. the case \( r = 0 \)), we have

\[
H_1 = \mathbb{E} h\left( \frac{\sqrt{km}}{\sqrt{n}} \sum_{i=1}^{km} X_i + \frac{1}{\sqrt{n}} \sum_{i=km+1}^{n} X_i \right) - \mathbb{E} h\left( \frac{\sqrt{km}}{\sqrt{n}} Z + \frac{1}{\sqrt{n}} \sum_{i=km+1}^{n} X_i \right)
= \frac{\gamma}{6\sqrt{km}} \mathbb{E} (Z^3 - 3Z)h\left( \frac{\sqrt{km}}{\sqrt{n}} Z + \frac{1}{\sqrt{n}} \sum_{i=km+1}^{n} X_i \right) + O\left( \frac{1}{n} \right)
= \frac{\gamma}{6\sqrt{km}} \mathbb{E} (Z^3 - 3Z)h\left( \frac{\sqrt{km}}{\sqrt{n}} Z + \sqrt{r} \tilde{Z} \right) + O\left( \frac{1}{n} \right),
\]

where we used Lemma 3.2 in the last equality. Using Gaussian integration by parts (and approximating \( h \) by arbitrarily close smooth functions in the intermediate step), we have

\[
\mathbb{E} (Z^3 - 3Z)h\left( \frac{\sqrt{km}}{\sqrt{n}} Z + \sqrt{r} \tilde{Z} \right) = -\frac{\sqrt{k^3 m^3}}{n^3} \mathbb{E} h''\left( \frac{\sqrt{km}}{\sqrt{n}} Z + \sqrt{r} \tilde{Z} \right)
\]

(3.13)

From (3.12) and (3.13), we have

\[
H_1 = \frac{km\gamma}{6\sqrt{n^3}} \mathbb{E} [(Z^3 - 3Z)h(Z)] + O\left( \frac{1}{n} \right) = \frac{\gamma}{6\sqrt{n}} \mathbb{E} [(Z^3 - 3Z)h(Z)] + O\left( \frac{1}{n} \right).
\]

(3.14)

For \( H_2 \), let

\[
h_1(x) = \mathbb{E} h\left( \frac{\sqrt{km}}{\sqrt{n}} Z + x \right).
\]
Then, by Lindeberg’s swapping argument,

\[
H_2 = \mathbb{E}h_1 \left( \sum_{i=km+1}^{n} \frac{X_i}{\sqrt{n}} \right) - \mathbb{E}h_1 \left( \sum_{i=km+1}^{n} \frac{Z_i}{\sqrt{n}} \right)
\]

\[
= \sum_{j=km+1}^{n} \left\{ \mathbb{E}h_1 \left( \sum_{i=km+1}^{j-1} \frac{X_i}{\sqrt{n}} + \sum_{i=j+1}^{n} \frac{Z_i}{\sqrt{n}} \right) \right. \left( \frac{X_j}{\sqrt{n}} - \frac{Z_j}{\sqrt{n}} \right)
\]

\[
+ \mathbb{E}h_1'' \left( \sum_{i=km+1}^{j-1} \frac{X_i}{\sqrt{n}} + \sum_{i=j+1}^{n} \frac{Z_i}{\sqrt{n}} \right) \left( \frac{X_j^2}{2n} - \frac{Z_j^2}{2n} \right)
\]

\[
+ O(1)\|h''_1\|_\infty \mathbb{E} \left( \left| \frac{X_j^3}{n^{3/2}} \right| + \left| \frac{Z_j^3}{n^{3/2}} \right| \right)
\]

\[
= O\left( \frac{r}{n^{3/2}} (\mathbb{E}|X_1|^3 + \mathbb{E}|Z|^3) \right),
\]

where we use Lemma 3.1 in the last equality. Combining (3.14) and (3.15), we have

\[
H_1 + H_2 = \frac{\gamma}{6\sqrt{n}} \mathbb{E}(Z^3 - 3Z)h(Z) + O\left( \frac{1}{n} \right).
\]

Thus, we have proved for the case \( r > 0 \).

\[\square\]

### 3.3 Proof of \((2.23)\)

(2.23) follows immediately from Lemmas 3.1 and 3.2.

### 3.4 Proof of Lemma 2.2

In this subsection, we use \( O(1) \) to denote a quantity which is bounded in absolute value by a universal constant. We will use the following lemma.

**Lemma 3.3.** Let \( n \geq 1 \) and \( X_1, \ldots, X_n \) be i.i.d. \( \text{Ber}(1/2) \) random variables. Then we have, for any constant \( x \) in the set \( A_n := \{ \frac{2z}{\sqrt{n}} - \sqrt{n}, z = 0, \ldots, n \} \cap [-n^{1/4}/8, n^{1/4}/8] \),

\[
\mathbb{P} \left( \frac{2}{\sqrt{n}} \sum_{i=1}^{n} (X_i - \frac{1}{2}) = x \right) = \frac{2}{\sqrt{2\pi n}} e^{-\frac{x^2}{2n}} \left( 1 + O \left( \frac{1 + x^4}{n} \right) \right)
\]

\[
= \frac{1}{\sqrt{2\pi}} \int_{x - \frac{1}{\sqrt{n}}}^{x + \frac{1}{\sqrt{n}}} e^{-\frac{y^2}{2}} dy \left( 1 + O \left( \frac{1 + x^4}{n} \right) \right).
\]
Proof of Lemma 3.3. In this proof, we will use the following inequality:

\[ \alpha - \frac{1}{2} \alpha^2 + \frac{1}{3} \alpha^3 - \alpha^4 \leq \log(1 + \alpha) \leq \alpha - \frac{1}{2} \alpha^2 + \frac{1}{3} \alpha^3, \quad \forall |\alpha| \leq \frac{1}{8}, \quad (3.18) \]

which is proved by Taylor’s expansion. Without loss of generality, we assume that \( x \in A_n, x \geq 0, \) and let \( z \) be the integer such that \( \frac{2z}{\sqrt{n}} - \sqrt{n} = x. \) Assume \( n \) is sufficiently large (otherwise (3.17) is trivial) so that \( z \) and \( n - z \) are sufficiently large. Then

\[ \mathbb{P} \left( \frac{2}{\sqrt{n}} \sum_{i=1}^{n} (X_i - \frac{1}{2}) = x \right) = \mathbb{P} \left( \sum_{i=1}^{n} X_i = z \right) = \frac{n!}{z!(n-z)!} \frac{1}{2^n}. \quad (3.19) \]

Using the Stirling formula

\[ n! = \sqrt{2\pi n} \left( \frac{n}{e} \right)^n \left( 1 + O \left( \frac{1}{n} \right) \right), \]

we have

\[ \frac{n!}{z!(n-z)!} \frac{1}{2^n} = \frac{\sqrt{2\pi n} n^n}{2^n \sqrt{2\pi z} \sqrt{2\pi (n-z) z} (n-z)^{(n-z)}} \frac{(1 + O \left( \frac{1}{n} \right))}{(1 + O \left( \frac{1}{z} \right)) (1 + O \left( \frac{1}{n-z} \right))}. \quad (3.20) \]

Plugging \( z = \frac{1}{2} \sqrt{n} x + \frac{n}{2} \) into (3.20) and using \( x \in [-n^{1/4}/8, n^{1/4}/8], \) we have

\[ \mathbb{P} \left( \frac{2}{\sqrt{n}} \sum_{i=1}^{n} (X_i - \frac{1}{2}) = x \right) = \frac{2}{\sqrt{2\pi n}} \left( 1 - \frac{x^2}{n} \right)^{-\frac{1}{2}} \left( 1 + \frac{x}{\sqrt{n}} \right)^{-\frac{\sqrt{n}}{2} - \frac{x^2}{2n}} \left( 1 - \frac{x}{\sqrt{n}} \right)^{-\frac{\sqrt{n}}{2} + \frac{x^2}{2n}} \times \left( 1 + O \left( \frac{1}{n + \sqrt{n} x} + \frac{1}{n - \sqrt{n} x} \right) \right) \]

\[ = \frac{2}{\sqrt{2\pi n}} \left( 1 + \frac{x}{\sqrt{n}} \right)^{-\frac{\sqrt{n}}{2} - \frac{x^2}{2}} \left( 1 - \frac{x}{\sqrt{n}} \right)^{-\frac{\sqrt{n}}{2} + \frac{x^2}{2}} \times \left( 1 + O \left( \frac{1}{n + \sqrt{n} x} + \frac{1}{n - \sqrt{n} x} + \frac{x^2}{n} \right) \right). \quad (3.21) \]

Because \( |x/\sqrt{n}| \leq |x/n^{1/4}| \leq 1/8, \) applying (3.18), we have

\[ \left( 1 + \frac{x}{\sqrt{n}} \right)^{-\frac{\sqrt{n}}{2} + \frac{x^2}{2}} = e^{-1 - \frac{x^2}{2\sqrt{n}} + O \left( \frac{x^2}{n} \right)}, \]

\[ \left( 1 - \frac{x}{\sqrt{n}} \right)^{-\frac{\sqrt{n}}{2} + \frac{x^2}{2}} = e^{-1 - \frac{x^2}{2\sqrt{n}} + O \left( \frac{x^2}{n} \right)}, \]

\[ \left( 1 - \frac{x^2}{n} \right)^{-\frac{\sqrt{n}}{2} - \frac{x^2}{2}} = e^{-1 + O \left( \frac{x^2}{n} \right)} \quad (3.22) \]
Plugging (3.22) into (3.21), we obtain
\[
\mathbb{P} \left( \frac{2}{\sqrt{n}} \sum_{i=1}^{n} (X_i - \frac{1}{2}) = x \right) = \frac{2}{\sqrt{2\pi n}} e^{-\frac{x^2}{2}} \left( 1 + O \left( \frac{1 + x^4}{n} \right) \right).
\] (3.23)

For the integral on the right hand side of (3.17), we have
\[
\int_{x-\frac{1}{\sqrt{n}}}^{x+\frac{1}{\sqrt{n}}} e^{-\frac{y^2}{2}} dy = e^{-\frac{x^2}{2}} \int_{x-\frac{1}{\sqrt{n}}}^{x+\frac{1}{\sqrt{n}}} e^{-\frac{y^2-x^2}{2}} dy = e^{-\frac{x^2}{2}} \int_{-\frac{1}{\sqrt{n}}}^{\frac{1}{\sqrt{n}}} \left[ 1 - xy - r^2 + O \left( \frac{1 + x^4}{n} \right) \right] dr
\] (3.24)
Combining (3.23) and (3.24), we complete the proof. \(\square\)

We now prove Lemma 2.2.

Proof of Lemma 2.2. Denote
\[
A_l = \{2z/\sqrt{l} - \sqrt{l}, z = 0, \ldots, l\} \cap [-l^{1/4}/8, l^{1/4}/8]
\]
and
\[
\bar{A}_l = \{2z/\sqrt{l} - \sqrt{l}, z = 0, \ldots, l\} \cap [-l^{1/4}/8, l^{1/4}/8].
\]
From the condition that \(h\) equals a constant in \((z - 1/2, z + 1/2)\) for \(z \in \mathbb{Z}\) and Gaussian tail bounds, we have
\[
\mathbb{E} h(S) - \mathbb{E} h \left( \frac{l}{2} + Z \sqrt{\frac{l}{4}} \right)
= \mathbb{E} h \left( \frac{S - \frac{l}{2}}{\sqrt{\frac{l}{2}}} + \frac{l}{2} \right) - \mathbb{E} h \left( \frac{l}{2} + Z \sqrt{\frac{l}{4}} \right)
= \sum_{x \in A_l} h \left( \frac{\sqrt{l}}{2} x + l \right) \left( P \left( \frac{S - l/2}{\sqrt{l}/2} = x \right) - \frac{1}{\sqrt{2\pi}} \int_{x-\sqrt{l}/2}^{x+\sqrt{l}/2} e^{-\frac{y^2}{2}} dy \right)
+ \sum_{x \in \bar{A}_l} h \left( \frac{\sqrt{l}}{2} x + l \right) \left( P \left( \frac{S - l/2}{\sqrt{l}/2} = x \right) - \frac{1}{\sqrt{2\pi}} \int_{x-\sqrt{l}/2}^{x+\sqrt{l}/2} e^{-\frac{y^2}{2}} dy \right) + O \left( \frac{1}{l} \right)
=: R_1 + R_2 + O \left( \frac{1}{l} \right).
\] (3.25)

From Lemma 3.3, we have
\[
\sum_{x \in A_l} \left| P \left( \frac{S - l/2}{\sqrt{l}/2} = x \right) - \frac{1}{\sqrt{2\pi}} \int_{x-\sqrt{l}/2}^{x+\sqrt{l}/2} e^{-\frac{y^2}{2}} dy \right|
= O \left( \frac{1}{l} \right) \sum_{x \in A_l} (1 + x^4) \frac{1}{\sqrt{2\pi}} \int_{x-\sqrt{l}/2}^{x+\sqrt{l}/2} e^{-\frac{y^2}{2}} dy = O \left( \frac{1}{l} \right).
\] (3.26)
From (3.26), we have
\[ R_1 = O\left(\frac{1}{l}\right). \tag{3.27} \]
For \( R_2 \), we have
\[ |R_2| \leq \sum_{x \in \tilde{A}_l} \left( P \left( \frac{S - l/2}{\sqrt{l}/2} = x \right) + \frac{1}{\sqrt{2\pi}} \int_{x-1/\sqrt{l}}^{x+1/\sqrt{l}} e^{-\frac{y^2}{2}} dy \right). \]

For \( x \in \tilde{A}_l \), we have \( |x| > l^{1/4}/8 \). Therefore, from binomial and Gaussian tail bounds, we have
\[ R_2 = O\left(\frac{1}{l}\right). \tag{3.28} \]
Combining (3.25), (3.27) and (3.28), we complete the proof. \( \square \)

### 3.5 Stein kernel for Gaussian mixtures

The following lemma was used in the proofs of Theorems 1.1 and 1.2 to apply the Stein kernel bound (2.3) to Gaussian mixtures.

**Lemma 3.4.** Let
\[ Y = \begin{cases} \begin{align*} Z_1 &\sim N(\mu_1, \sigma_1^2), & \text{with probability } p, \\ Z_2 &\sim N(\mu_2, \sigma_2^2), & \text{with probability } 1 - p. \end{align*} \end{cases} \]

Let \( \tau \) be its Stein kernel. Then we have \( \|\tau\|_\infty \leq C \), where \( C \) is a positive constant depending only on \( p, \mu_1, \sigma_1^2, \mu_2, \sigma_2^2 \).

**Proof of Lemma 3.4.** Assume without loss of generality that \( 0 < p < 1 \). Otherwise, the lemma follows from the fact that the Stein kernel for a Gaussian variable equals its variance. We first consider the case \( x \geq 0 \). Let \( \phi_1(x) \) and \( \phi_2(x) \) be the density of \( N(\mu_1, \sigma_1^2) \) and \( N(\mu_2, \sigma_2^2) \), respectively. Then, \( Y \) has density
\[ p(x) = p\phi_1(x) + q\phi_2(x), \quad q = 1 - p. \]

From the expression of Stein kernel in (2.2), we have
\[ \tau(x) = \frac{1}{p\phi_1(x) + q\phi_2(x)} \int_x^\infty (y - p\mu_1 - q\mu_2)(p\phi_1(y) + q\phi_2(y)) dy. \]

Using
\[ \int_x^\infty (y - \mu_i)\phi_i(y) dy = \sigma_i^2 \phi_i(x), \quad i = 1, 2, \]
we obtain
\[ \tau(x) = \frac{p\sigma_1^2 \phi_1(x) + q\sigma_2^2 \phi_2(x) + q(\mu_1 - \mu_2) \int_x^\infty \phi_1(y) dy + p(\mu_2 - \mu_1) \int_x^\infty \phi_2(y) dy}{p\phi_1(x) + q\phi_2(x)}, \]

20
which is bounded because \( \int_x^\infty \phi_i(y)dy \) decays proportional to \( \phi_i(x)/x \) for large \( x \) and \( i = 1, 2. \)

The case \( x < 0 \) is proved similarly using the alternative expression of the Stein kernel
\[
\tau(x) = -\frac{1}{p(x)} \int_{-\infty}^x (y - \mu)p(y)dy.
\]

\[\square\]

### 3.6 Existence of Bernoulli component

Recall the condition of Theorem 1.2: The support of \( X_1 \) is \( \{s_0, s_1, s_2, \ldots\} \) such that \( \{|s_i - s_0|, i \geq 1\} \) has the greatest common divisor 1. By successively looking for numbers that decrease the common divisor, this condition implies that there exists a finite \( r \) such that \( \{|s_i - s_0|, 1 \leq i \leq r\} \) has the greatest common divisor 1. By shifting invariance, we assume without loss of generality that \( 0 = s_0 < s_1 < \cdots < s_r \). Repeatedly using Bézout’s identity, there exist non-zero integers \( m_1, \ldots, m_r \) such that
\[
m_1s_1 + \cdots + m_rs_r = 1.
\]

Let \( X_{ij}, i, j \geq 1 \) follow the same distribution as \( X_1 \). For \( 1 \leq i \leq r \), let
\[
t_i = \begin{cases} 
0, & \text{if } m_i > 0 \\
-s_i, & \text{if } m_i < 0,
\end{cases}
\]
and
\[
Y_i = \sum_{j=1}^{m_i} (X_{ij} + t_j).
\]

Then, \( \sum_{i=1}^r Y_i \) can take values 0 and 1. This implies that \( \sum_{i=1}^r X_i - \mu \) has a \( Ber(1/2) \) component for \( m = |m_1| + \cdots + |m_r| \) subject to shifting by a fixed integer.

### Acknowledgements

Fang X. was partially supported by Hong Kong RGC GRF 14302418, 14305821 and a CUHK direct grant.

### References

Barbour, A. (1986). Asymptotic expansions based on smooth functions in the central limit theorem. *Probability Theory and Related Fields*, 72(2):289–303.

Berry, A. C. (1941). The accuracy of the gaussian approximation to the sum of independent variates. *Trans. Amer. Math. Soc.*, 49(1):122–136.

Bhattacharya, R. N. and Rao, R. R. (1976). *Normal approximation and asymptotic expansions*. John Wiley & Sons.
Braverman, A., Dai, J., and Fang, X. (2022). High-order steady-state diffusion approximations. *Operations Research*.

Chatterjee, S. (2009). Fluctuations of eigenvalues and second order poincaré inequalities. *Probability Theory and Related Fields*, 143(1):1–40.

Chen, L. H. Y., Goldstein, L., and Shao, Q.-M. (2011). *Normal approximation by Stein's method*. Springer.

Courtade, T. A., Fathi, M., and Pananjady, A. (2019). Existence of Stein kernels under a spectral gap, and discrepancy bounds. *Ann. Inst. Henri Poincaré, Probab. Stat.*, 55(2):777–790.

Esseen, C.-G. (1942). On the liapunov limit error in the theory of probability. *Ark. Mat. Astr. Fys.*, 28:1–19.

Fang, X. (2019). Wasserstein-2 bounds in normal approximation under local dependence. *Electronic Journal of Probability*, 24:1–14.

Fang, X., Luo, L., and Shao, Q. M. (2020). A refined cramér-type moderate deviation for sums of local statistics. *Bernoulli*, 26(3):2319–2352.

Fathi, M. (2019). Stein kernels and moment maps. *Ann. Probab.*, 47(4):2172–2185.

Fathi, M. (2021). Higher-order stein kernels for gaussian approximation. *Studia Math.*, 256(3):241–258.

Kim, Y. T. and Park, H. S. (2018). An edgeworth expansion for functionals of gaussian fields and its applications. *Stochastic Processes and their Applications*, 128(12):3967–3999.

Ledoux, M., Nourdin, I., and Peccati, G. (2015). Stein’s method, logarithmic Sobolev and transport inequalities. *Geometric and Functional Analysis*, 25(1):256–306.

Nourdin, I. and Peccati, G. (2009). Stein’s method on wiener chaos. *Probability Theory and Related Fields*, 145(1):75–118.

Petrov, V. V. (1975). *Sums of independent random variables*. Springer-Verlag.

Rinott, Y. and Rotar, V. (2003). On edgeworth expansions for dependency-neighborhoods chain structures and stein’s method. *Probability Theory and Related Fields*, 126(4):528–570.

Saumard, A. (2019). Weighted poincaré inequalities, concentration inequalities and tail bounds related to stein kernels in dimension one. *Bernoulli*, 25(4B):3978–4006.

Stein, C. (1972). A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. In *Proceedings of the sixth Berkeley symposium on mathematical statistics and probability, volume 2: Probability theory*, volume 6, pages 583–603. University of California Press.

Stein, C. (1986). *Approximate computation of expectations*. IMS.