Submitted to the Annals of Probability

CRAMÉR-TYPE MODERATE DEVIATION THEOREMS FOR NONNORMAL APPROXIMATION

By Qi-Man Shao*, Mengchen Zhang and Zhuo-Song Zhang

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

A Cramér-type moderate deviation theorem quantifies the relative error of the tail probability approximation. It provides theoretical justification when the limiting tail probability can be used to estimate the tail probability under study. Chen, Fang and Shao [11] obtained a general Cramér-type moderate result using Stein's method when the limiting was a normal distribution. In this paper, Cramér-type moderate deviation theorems are established for nonnormal approximation under a general Stein identity, which is satisfied via the exchangeable pair approach and Stein's coupling. In particular, a Cramér-type moderate deviation theorem is obtained for the Curie–Weiss model and the imitative monomer-dimer mean-field model.

1. Introduction. Let W_n be a sequence of random variables that converge to Y in the distribution. The Cramér-type moderate deviation quantifies the relative error of the distribution approximation, that is,

$$\frac{\mathrm{P}(W_n \geq x)}{\mathrm{P}(Y \geq x)} = 1 + \text{error} \to 1 \tag{1.1}$$

for $0 \leq x \leq a_n$, where $a_n \to \infty$ as $n \to \infty$. When Y is the normal random variable and W_n is the standardized sum of the independent random variables, the Cramér-type moderate deviation is well understood. In particular, for independent and identically distributed random variables $X_1, \cdots, X_n$

*Research partially supported by Hong Kong RGC GRF 14302515 and 14304917

MSC 2010 subject classifications: Primary 60F10; secondary 60F05

Keywords and phrases: Moderate deviation, Nonnormal approximation, Stein's method, Curie–Weiss model, Imitative monomer-dimer mean-field model

with $\mathrm{E}X_i = 0, \mathrm{E}X_i^2 = 1$ and $\mathrm{E}e^{t_0\sqrt{|X_1|}} < \infty$, it holds that

$$\frac{\mathrm{P}(W_n \geq x)}{1 - \Phi(x)} = 1 + O(1)(1 + x^3)/\sqrt{n} \tag{1.2}$$

for $0 \leq x \leq n^{1/6}$, where $W_n = (X_1 + \cdots + X_n)/\sqrt{n}$. The finite-moment-generating function of $|X_1|^{1/2}$ is necessary, and both the range $0 \leq x \leq n^{1/6}$ and the order of the error term $(1 + x^3)/\sqrt{n}$ are optimal. We refer to Linnik [18] and Petrov [21] for details.

Considering general dependent random variables whose dependence is defined in terms of a Stein identity, Chen, Fang and Shao [11] obtained a general Cramér-type moderate deviation result for normal approximation using Stein's method. Stein's method, introduced by Stein [26], is a completely different approach to distribution approximation than the classical Fourier transform. It works not only for independent random variables but also for dependent random variables. It can also make the distribution approximation accurate. Extensive applications of Stein's method to obtain Berry–Esseen-type bounds can be found in, for example, Diaconis [15], Stein [27], Barbour [3], Goldstein and Reinert [17], Chen and Shao [9, 10], Chatterjee [5], Nourdin and Peccati [19] and Shao and Zhang [24]. We refer to Chen, Goldstein and Shao [12], Nourdin and Peccati [20] and Chatterjee [6] for comprehensive coverage of the method's fundamentals and applications. In addition to the normal approximation, Chatterjee and Shao [8] obtained a general nonnormal approximation via the exchangeable pair approach and the corresponding Berry–Esseen-type bounds. We also refer to Shao and Zhang [23] for a more general result.

The main purpose of this paper is to obtain a Cramér-type moderate deviation theorem for nonnormal approximation. Our main tool is based on Stein's method, combined with some techniques in Chatterjee and Shao [8] and Chen, Fang and Shao [11]. The paper is organized as follows. Section 2 presents a Cramér-type moderate deviation theorem under a general Stein identity setting, which recovers the result of Chen, Fang and Shao [11] as a special case. In Section 3, the result is applied to two examples: the Curie–

Weiss and imitative monomer-dimer models. The proofs of the main results are given in Sections 4 and 5.

2. Main Results. Let $W := W_n$ be the random variable of interest. Following the setting in Chatterjee and Shao [8] and Chen, Fang and Shao [11], we assume that there exists a constant δ, a nonnegative random function $\hat{K}(t)$, a function g and a random variable R such that

$$\mathrm{E}(f(W)g(W)) = \mathrm{E}\left(\int_{|t|\le\delta} f'(W+t)\hat{K}(t)dt\right) + \mathrm{E}(f(W)R(W)) \tag{2.1}$$

for all absolutely continuous functions f for which the expectation of either side exists. Let

$$\hat{K}_1 = \int_{|t|\le\delta} \hat{K}(t)dt \tag{2.2}$$

and

$$G(y) = \int_0^y g(t)dt. \tag{2.3}$$

Let Y be a random variable with the probability density function

$$p(y) = c_1\, e^{-G(y)}, \quad y \in \mathbb{R}, \tag{2.4}$$

where c_1 is a normalizing constant.

In this section, we present a Cramér-type moderate theorem for general distribution approximation under Stein's identity in general and under an exchangeable pair and Stein's couplings in particular.

Before presenting the main theorem, we first give some of the conditions of g.

Assume that

(A1) The function g is nondecreasing and $g(0) = 0$.

(A2) For $y \neq 0,\ yg(y) > 0$.

(A3) There exists a positive constant c_2 such that for $x, y \in \mathbb{R}$,

$$|g(x+y)| \le c_2\left(|g(x)| + |g(y)| + 1\right). \tag{2.5}$$

(A4) There exists $c_3 \geq 1$ such that for $y \in \mathbb{R}$,

$$|g'(y)| \leq c_3 \left(\frac{1 + |g(y)|}{1 + |y|} \right). \tag{2.6}$$

A large class of functions satisfy conditions (A1)–(A4). A typical example is $g(y) = \mathrm{sgn}(y)|y|^p,\ p \geq 1$.

We are now ready to present our main theorem.

THEOREM 2.1. *Let W be a random variable of interest satisfying (2.1). Assume that conditions (A1)–(A4) are satisfied. Additionally, assume that there exist $\tau_1 > 0, \tau_2 > 0, \delta_1 > 0$ and $\delta_2 \geq 0$ such that*

$$|\,\mathrm{E}(\hat{K}_1|W) - 1| \leq \delta_1\big(|g(W)|^{\tau_1} + 1\big), \tag{2.7}$$

$$|R(W)| \leq \delta_2\big(|g(W)|^{\tau_2} + 1\big). \tag{2.8}$$

In addition, there exist constants $d_0 \geq 1, d_1 > 0$ and $0 \leq \alpha < 1$ such that

$$\mathrm{E}(\hat{K}_1|W) \leq d_0, \tag{2.9}$$

$$\delta|g(W)| \leq d_1, \tag{2.10}$$

$$|R(W)| \leq \alpha(|g(W)| + 1). \tag{2.11}$$

Then, we have

$$\begin{aligned} \frac{\mathrm{P}(W > z)}{\mathrm{P}(Y > z)} = 1 + O(1)\Big(&\delta\big(1 + zg^2(z)\big) \\ &+ \delta_1\big(1 + zg^{\tau_1+1}(z)\big) + \delta_2\big(1 + zg^{\tau_2}(z)\big)\Big) \end{aligned} \tag{2.12}$$

for $z \geq 0$ satisfying $\delta zg^2(z) + \delta_1 zg^{\tau_1+1}(z) + \delta_2 zg^{\tau_2}(z) \leq 1$, where $O(1)$ is bounded by a finite constant depending only on $d_0, d_1, c_1, c_2, c_3, \tau_1, \tau_2, \alpha$ and $\max(g(1), |g(-1)|)$.

The condition (2.1) is called a general Stein identity, see Chen, Goldstein and Shao [12, Chapter 2]. We use the exchangeable pair approach and Stein's coupling to construct $\hat{K}(t)$ and $R(W)$ as follows.

Let (W, W') be an exchangeable pair, that is, (W, W') has the same joint distribution as (W', W). Let $\Delta = W - W'$. Assume that

$$\mathrm{E}(\Delta|W) = \lambda(g(W) - R(W)), \tag{2.13}$$

where $0 < \lambda < 1$. Assume that $|\Delta| \le \delta$ for some constant $\delta > 0$. It is known (see, e.g., Chatterjee and Shao [8]) that (2.1) is satisfied with

$$\hat{K}(t) = \frac{1}{2\lambda}\Delta(I(-\Delta \le t \le 0) - I(0 < t \le \Delta)).$$

Clearly, we have

$$\hat{K}_1 = \frac{1}{2\lambda}\Delta^2.$$

For exchangeable pairs, we have the following corollary.

COROLLARY 2.1. *Assume that $g(W)$, $\hat{K}_1$ and $R(W)$ satisfy the conditions (A1)–(A4) and (2.7)–(2.11) stated in Theorem 2.1; then,* (2.12) *holds.*

Stein's coupling introduced by Chen and Röllin [13] is another way to construct the general Stein identity.

A triple (W, W', G) is called a g-Stein's coupling if there is a function g such that

$$\mathrm{E}(Gf(W') - Gf(W)) = \mathrm{E}(f(W)g(W)) \tag{2.14}$$

for all absolutely continuous function f, such that the expectations on both sides exist. Assume that $|W' - W| \le \delta$. Then, by Chen and Röllin [13], we have

$$\mathrm{E}(f(W)g(W)) = \mathrm{E}\Big(\int_{|t|\le\delta} f'(W+t)\hat{K}(t)dt\Big),$$

where

$$\hat{K}(t) = G(I(0 \le t \le W' - W) - I(W' - W \le t < 0)).$$

It is easy to see that $\hat{K}_1 = G(W - W')$.

The following corollary presents a moderate deviation result for Stein's coupling.

COROLLARY 2.2. *If $\hat{K}_1 \geq 0$, and $g(W)$ and $\hat{K}_1$ satisfy the conditions (A1)–(A4) and (2.7), (2.9) and (2.10) stated in Theorem 2.1, then (2.12) holds with $\delta_2 = 0$.*

REMARK 2.1. Condition (2.11) may not be satisfied when $|W|$ is large in some applications. Following the proof of Theorem 2.1, when (2.11) is replaced by the following condition, there exist $0 \leq \alpha < 1$, $d_2 \geq 0, d_3 > 0$ and $\kappa > 0$ such that

$$|R(W)| \leq \alpha\,(|g(W)| + 1) + d_2\, I(|W| > \kappa), \tag{2.15}$$

and

$$d_2 P(|W| > \kappa) \leq d_3 e^{-2s_0 d_1^{-1}\delta^{-1}}, \tag{2.16}$$

where d_1 is given in (2.10) and $s_0 = \max\{s : \delta s g^2(s) \leq 1\}$, Theorem 2.1 and Corollaries 2.1 and 2.2 remain valid with $O(1)$ bounded by a finite constant depending only on $d_0, d_1, d_3, c_1, c_2, c_3, \tau_1, \tau_2, \alpha$ and $\max(g(1), |g(-1)|)$.

3. Applications. In this section, we apply the main results to the Curie–Weiss model at the critical temperature and the imitative monomer-dimer model.

3.1. *The Curie–Weiss Model at the critical temperature.* The Curie–Weiss model is a simple statistical mechanical model of ferromagnetic interaction. For $\sigma = (\sigma_1, ..., \sigma_n)$, representing "spins" in $\{-1, 1\}^n$, σ is assigned probability

$$p_\sigma(\sigma) = C_\beta \exp\left(\frac{\beta}{n} \sum_{1\leq i<j\leq n} \sigma_i\sigma_j\right),\ \beta > 0 \tag{3.1}$$

where C_β is a normalizing constant and in this context $1/\beta$ is referred to as the temperature. It is interesting to understand the behavior of the total magnetization. For $0 < \beta < 1$, Chen, Fang and Shao [11] obtained a Cramér-type moderate deviation result for this model, in which case $n^{-1/2}\sum_i \sigma_i$

converges to a normal distribution. When $\beta = 1$, Simon and Griffiths [25] proved that the law of

$$W = n^{-3/4} \sum_{i=1}^{n} \sigma_i \tag{3.2}$$

converges to $\mathcal{W}(4, 12)$ as $n \to \infty$, with the probability density function

$$f_Y(y) = Ce^{-\frac{y^4}{12}}. \tag{3.3}$$

where $C = \sqrt{2}/(3^{1/4}\Gamma(1/4))$. Chatterjee and Dey [7] used the exchangeable pair technique to derive a tail probability

$$\mathrm{P}(|W| \geq t) \leq 2e^{-ct^4},$$

where $c > 0$ is an absolute constant. Here, we present the moderate deviation for this model.

THEOREM 3.1. *Let W be the scaled total spin (3.2) in the Curie–Weiss model, where the vector σ of spins has p.d.f. (3.1) with $\beta = 1$. Let Y be a random variable with p.d.f. given by (3.3). Then, we have*

$$\frac{\mathrm{P}(W > z)}{\mathrm{P}(Y > z)} = 1 + O(1)\big((1 + z^6)n^{-1/2}\big), \tag{3.4}$$

for $0 \leq z \leq n^{1/12}$, where $O(1)$ is bounded by an absolute constant.

3.2. *The imitative monomer-dimer mean-field model.* In this subsection, we consider the imitative monomer-dimer model and give the moderate deviation result. A pure monomer-dimer model can be used to study the properties of diatomic oxygen molecules deposited on tungsten or liquid mixtures with molecules of unequal size [see 16, 22, for example]. Chang [4] studied the attractive component of the van der Waals potential, while Alberici, Contucci, Fedele and Mingione [1] and Alberici, Contucci and Mingione [2] considered the asymptotic properties. Chen [14] recently obtained the Berry–Esseen bound by using Stein's method. In this subsection, we apply our main theorem to obtain the moderate deviation result.

For $n \geq 1$, let $G = (V, E)$ be a complete graph with vertex set $V = \{1, \cdots, n\}$ and edge set $E = \big\{uv = \{u, v\} : u, v \in V, u < v\big\}$. A dimer configuration on the graph G is a set D of pairwise nonincident edges satisfying the following rule: if $uv \in D$, then for all $w \neq v$, $uw \notin D$. Given a dimer configuration D, the set of monomers $\mathcal{M}(D)$ is the collection of dimer-free vertices. Let $\mathbf{D}$ denote the set of all dimer configurations. Denote the element number of a set by $\#(\cdot)$. Then, we have

$$2\#(D) + \#(\mathcal{M}(D)) = n.$$

We now introduce the imitative monomer-dimer model. The Hamiltonian of the model with an imitation coefficient $J \geq 0$ and an external field $h \in \mathbb{R}$ is given by

$$-T(D) = n(Jm(D)^2 + bm(D))$$

for all $D \in \mathbf{D}$, where $m(D) = \#(\mathcal{M}(D))/n$ is called the monomer density and the parameter b is given by

$$b = \frac{\log n}{2} + h - J.$$

The associated Gibbs measure is defined as

$$p(D) = \frac{e^{-T(D)}}{\sum_{D \in \mathbf{D}} e^{-T(D)}}.$$

Let

$$H(x) = -Jx^2 - \frac{1}{2}\left(1 - g\big(\tau(x)\big) + \log\big(1 - g(\tau(x))\big)\right), \tag{3.5}$$

where

$$g(x) = \frac{1}{2}\left(\sqrt{e^{4x} + 4e^{2x}} - e^{2x}\right), \quad \tau(x) = (2x - 1)J + h.$$

Alberici, Contucci and Mingione [2] showed that the imitative monomer-dimer model exhibits the following three phases. Let

$$J_c = \frac{1}{4(3 - 2\sqrt{2})}, \quad h_c = \frac{1}{2}\log(2\sqrt{2} - 2) - \frac{1}{4}.$$

There exists a function $\gamma : (J_c, \infty) \to \mathbb{R}$ with $\gamma(J_c) = h_c$ such that if $(J, h) \notin \Gamma$, where $\Gamma := \{(J, \gamma(J)) : J > J_c\}$, then the function $H(x)$ has a unique maximizer m_0 that satisfies $m_0 = g(\tau(m_0))$. Moreover, if $(J, h) \notin \Gamma \cup \{(J_c, h_c)\}$, then $H''(m_0) < 0$. If $(J, h) = (J_c, h_c)$, then $m_0 = m_c := 2 - \sqrt{2}$ and

$$H'(m_c) = H''(m_c) = H^{(3)}(m_c) = 0,$$

but

$$H^{(4)}(m_c) < 0.$$

If $(J, h) \in \Gamma$, then $H(s)$ has two distinct maximizers; therefore, in this case, $m(D)$ may not converge. Hence, we consider only the cases when $(J, h) \notin \Gamma$.

Alberici, Contucci and Mingione [2] showed that when $(J, h) \notin \Gamma \cup \{(J_c, h_c)\}$, $n^{1/2}(m(D) - m_0)$ converges to a normal distribution with zero mean and variance $\lambda_0 = -(H''(m_0))^{-1} - (2J)^{-1}$. However, when $(J, h) = (J_c, h_c)$, then $n^{1/4}(m(D) - m_0)$ converges to Y, whose p.d.f. is given by

$$p(y) = c_1 e^{-\lambda_c y^4/24} \tag{3.6}$$

with $\lambda_c = -H^{(4)}(m_c) > 0$ and c_1 is a normalizing constant. Chen [14] obtained the Berry–Esseen bound using Stein's method.

Similar to the Curie–Weiss model, we use the following notations. Let $\Sigma = \{0, 1\}^n$. For each $\sigma = (\sigma_1, ..., \sigma_n) \in \Sigma$, define a Hamiltonian

$$-T(\sigma) = n(Jm(\sigma)^2 + bm(\sigma)),$$

where $m(\sigma) = n^{-1}(\sigma_1 + ... + \sigma_n)$ is the magnetization of the configuration σ. Denote by $\boldsymbol{A}(\sigma)$ the set of all sites $i \in V$ such that $\sigma_i = 1$. Also, let $D(\sigma)$ denote the total number of dimer configurations $D \in \mathbf{D}$ with $\mathcal{M}(D) = \boldsymbol{A}(\sigma)$. Therefore, the Gibbs measure can be written as

$$p(\sigma) = \frac{D(\sigma)\exp(-T(\sigma))}{\sum_{\tau \in \Sigma} D(\tau)\exp(-T(\tau))}.$$

The following result gives a Cramér-type moderate deviation for the magnetization.

THEOREM 3.2. *If* $(J,h) \notin \Gamma \cup \{J_c, h_c\}$*, then, for* $0 \le z \le n^{1/6}$,

$$\frac{\mathrm{P}(n^{1/2}(m(\sigma)-m_0) \ge z)}{\mathrm{P}(Z_0 > z)} = 1 + O(1)n^{-1/2}(1+z^3), \tag{3.7}$$

where Z_0 *follows normal distribution with zero mean and variance* $\lambda_0 = -(H''(m_0))^{-1} - (2J)^{-1}$. *If* $(J,h) = (J_c, h_c)$*, then for* $0 \le z \le n^{1/20}$,

$$\frac{\mathrm{P}(n^{1/4}(m(\sigma)-m_c) \ge z)}{\mathrm{P}(Y \ge z)} = 1 + O(1)(n^{-1/4}(1+z^5)), \tag{3.8}$$

where Y *is a random variable with the probability density function given in* (3.6).

4. Proofs of main results. In this section, we give the proofs of the main theorems. In what follows, we use C or $C_1, C_2, \cdots$ to denote a finite constant depending only on $c_1, c_2, c_3, d_0, d_1, \tau_1, \tau_2, \mu_1$ and α, where $\mu_1 = \max(g(1), |g(-1)|) + 1$.

4.1. *Proof of Theorem 2.1.* Let Y be a random variable with a probability density function given in (3.6) and $F(z)$ be the distribution function of Y. We start with a preliminary lemma on the properties of $(1-F(w))/p(w)$ and $F(w)/p(w)$, whose proof is postponed to Section 4.2.

LEMMA 4.1. *Assume that conditions (A1)–(A4) are satisfied. Then, we have*

$$\frac{1}{\max(1,c_3)(1+g(w))} \le \frac{1-F(w)}{p(w)} \le \min\left\{\frac{1}{g(w)}, 1/c_1\right\} \quad \text{for} \quad w > 0 \tag{4.1}$$

and

$$\frac{F(w)}{p(w)} \le \min\left\{\frac{1}{|g(w)|}, 1/c_1\right\} \quad \text{for} \quad w < 0. \tag{4.2}$$

PROOF OF THEOREM 2.1. Let f_z be the solution to Stein's equation

$$f'(w) - f(w)g(w) = I(w \le z) - F(z). \tag{4.3}$$

As shown in Chatterjee and Shao [8], the solution f_z can be written as

(4.4)
$$f_z(w) = \begin{cases} \dfrac{1}{p(w)} \int_{-\infty}^{w} (I(y \le z) - F(z))p(y)dy = \dfrac{F(w)(1-F(z))}{p(w)}, & w \le z; \\ -\dfrac{1}{p(w)} \int_{w}^{\infty} (I(y \le z) - F(z))p(y)dy = \dfrac{F(z)(1-F(w))}{p(w)}, & w > z. \end{cases}$$

From (2.1), we have

(4.5)
$$\begin{aligned} & \mathrm{E}(f_z(W)g(W) - f_z(W)R(W)) \\ & \quad = \mathrm{E}\left(\int_{|t|\le\delta} f_z'(W+t)\hat{K}(t)dt\right) \\ & \quad = \mathrm{E}\Big(\int_{|t|\le\delta} \big(f_z(W+t)g(W+t) \\ & \qquad\qquad + P(Y>z) - I(W+t>z)\big)\hat{K}(t)dt\Big) \\ & \quad \le \mathrm{E}\Big(\int_{|t|\le\delta} \big(f_z(W+t)g(W+t) - f_z(W)g(W)\big)\hat{K}(t)dt\Big) \\ & \qquad + \mathrm{E}(\hat{K}_1 f_z(W)g(W)) \\ & \qquad + \mathrm{E}\left(\hat{K}_1\big(P(Y>z) - I(W>z+\delta)\big)\right) \\ & \quad \le \mathrm{E}\left(\int_{|t|\le\delta} \big|f_z(W+t)g(W+t) - f_z(W)g(W)\big|\hat{K}(t)dt\right) \\ & \qquad + \mathrm{E}(\hat{K}_1 f_z(W)g(W)) \\ & \qquad + \mathrm{E}\left(|\,\mathrm{E}(\hat{K}_1|W) - 1|\big|P(Y>z) - I(W>z+\delta)\big|\right) \\ & \qquad + P(Y>z) - P(W>z+\delta). \end{aligned}$$

Rearranging (4.5) leads to

$$
\begin{aligned}
&\mathrm{P}(W > z+\delta) - \mathrm{P}(Y > z) \\
&\quad \le \mathrm{E}\Big(\int_{|t|\le\delta} \big|f_z(W+t)g(W+t) - f_z(W)g(W)\big|\hat{K}(t)dt\Big) \\
&\qquad + \mathrm{E}(|\,\mathrm{E}(\hat{K}_1|W) - 1||f_z(W)g(W)|) \\
&\qquad + \mathrm{E}\left(|(\mathrm{E}(\hat{K}_1|W) - 1)(P(Y>z) - I(W>z+\delta))|\right) \\
&\qquad + \mathrm{E}(f_z(W)|R(W)|) \\
&\quad := I_1 + I_2 + I_3 + I_4,
\end{aligned} \tag{4.6}
$$

where

$$
\begin{aligned}
I_1 &= \mathrm{E}\Big(\int_{|t|\le\delta} \big|f_z(W+t)g(W+t) - f_z(W)g(W)\big|\hat{K}(t)dt\Big), \\
I_2 &= \mathrm{E}(|(\mathrm{E}(\hat{K}_1|W) - 1)f_z(W)g(W)|), \\
I_3 &= \mathrm{E}\left(|(\mathrm{E}(\hat{K}_1|W) - 1)(P(Y>z) - I(W>z+\delta))|\right), \\
I_4 &= \mathrm{E}(f_z(W)|R(W)|).
\end{aligned}
$$

The following propositions provide estimates of I_1, I_2, I_3 and I_4, whose proofs are given in Section 4.3.

PROPOSITION 4.1. *If $\delta \le 1$, then*

$$I_1 \le C\delta. \tag{4.7}$$

Assume that $z \ge 0, \max(\delta,\delta_1,\delta_2) \le 1$ and $\delta z g^2(z) + \delta_1 z g^{\tau_1+1}(z) + \delta_2 z g^{\tau_2}(z) \le 1$. Then, we have

$$I_1 \le C\delta(1 + zg^2(z))(1 - F(z)). \tag{4.8}$$

PROPOSITION 4.2. *We have*

$$I_2 + I_3 \le C\delta_1, \quad I_4 \le C\delta_2. \tag{4.9}$$

For $z > 0, \max(\delta, \delta_1, \delta_2) \le 1$ *and* $\delta z g^2(z) + \delta_1 z g^{\tau_1+1}(z) + \delta_2 z g^{\tau_2}(z) \le 1$, *we have*

$$I_2 + I_3 \le C\delta_1\big(1 + zg^{\tau_1+1}(z)\big)(1 - F(z)), \tag{4.10}$$

$$I_4 \le C\delta_2\big(1 + zg^{\tau_2}(z)\big)(1 - F(z)). \tag{4.11}$$

We now go back to the proof of Theorem 2.1. First, we use (4.6) and Propositions 4.1 and 4.2 to prove the Berry–Esseen bound

$$|\mathrm{P}(W > z) - \mathrm{P}(Y > z)| \le C(\delta + \delta_1 + \delta_2), \tag{4.12}$$

where $C \ge 1$. By (4.6), (4.7) and (4.9), for $\delta \le 1$, we have

$$\mathrm{P}(W > z + \delta) - P(Y > z) \le C(\delta + \delta_1 + \delta_2). \tag{4.13}$$

Together with

$$\mathrm{P}(Y > z) - \mathrm{P}(Y > z + \delta) \le c_1 \int_z^{z+\delta} e^{-G(w)} dw \le c_1 \delta,$$

we have

$$\mathrm{P}(W > z) - \mathrm{P}(Y > z) \le C(\delta + \delta_1 + \delta_2).$$

Similarly, we have

$$\mathrm{P}(W > z) - \mathrm{P}(Y > z) \ge -C(\delta + \delta_1 + \delta_2).$$

This proves the inequality (4.12) for $\delta \le 1$. For $\delta > 1$, (4.12) is trivial because $C \ge 1$.

Next, we move to prove (2.12). Let $z_0 > 1$ be a constant such that

$$\min\big\{z_0 g^2(z_0), z_0 g^{\tau_1+1}(z_0), z_0 g^{\tau_2}(z_0), z_0\big\} \ge 1.$$

For $0 \le z \le z_0$, (2.12) follows from (4.12). In fact,

$$\frac{\mathrm{P}(W > z) - \mathrm{P}(Y > z)}{\mathrm{P}(Y > z)} \le \frac{C(\delta + \delta_1 + \delta_2)}{1 - F(z_0)}, \tag{4.14}$$

where C is a constant. For $z > z_0$, we can assume $\max\{\delta, \delta_1, \delta_2\} \le 1$; otherwise, it would contradict the condition $\delta z g^2(z) + \delta_1 z g^{\tau_1+1}(z) + \delta_2 z g^{\tau_2}(z) \le 1$.

Recall by (4.6),

$$\mathrm{P}(W > z + \delta) - \mathrm{P}(Y > z) \le I_1 + I_2 + I_3 + I_4.$$

It follows from Propositions 4.1 and 4.2 that

$$\begin{aligned}
&\mathrm{P}(W > z + \delta) - \big(1 - F(z)\big) \\
&\quad \le C(1 - F(z))\Big(\delta(1 + zg^2(z)) \\
&\qquad\qquad + \delta_1(1 + zg^{\tau_1+1}(z)) + \delta_2(1 + zg^{\tau_2}(z))\Big).
\end{aligned} \tag{4.15}$$

By replacing z with $z-\delta$, and noting that g is nondecreasing, we can rewrite (4.15) as

$$\begin{aligned}
&\mathrm{P}(W > z) - (1 - F(z - \delta)) \\
&\quad \le C(1 - F(z - \delta))\Big(\delta(1 + zg^2(z)) + \delta_1(1 + zg^{\tau_1+1}(z)) \\
&\qquad\qquad + \delta_2(1 + zg^{\tau_2}(z))\Big).
\end{aligned} \tag{4.16}$$

As $p(y)$ is decreasing in $[z - \delta, z]$, we have

$$\begin{aligned}
F(z) - F(z - \delta) &= \int_{z-\delta}^{z} p(t)dt \\
&\le \delta p(z - \delta) \le e^{\delta g(z)} \delta p(z).
\end{aligned}$$

By our assumptions, $\delta(1 + g^2(z)) \le 2$ and hence $\delta g(z) \le 1$. By (4.1), we also have

$$p(z) \le \max(1, c_3)(1 + g(z))(1 - F(z));$$

then,

$$F(z) - F(z - \delta) \le C\delta(1 + g(z))(1 - F(z))$$

for some constant C. Recall that $\delta(1 + g(z)) \le 2$; then,

$$1 - F(z - \delta) \le C(1 - F(z)).$$

Together with (4.16), we get

$$\begin{aligned}
&\mathrm{P}(W > z) - (1 - F(z)) \\
&\quad \le \mathrm{P}(W > z) - (1 - F(z - \delta)) + F(z) - F(z - \delta) \\
&\quad \le C(1 - F(z - \delta))\Big(\delta(1 + zg^2(z)) + \delta_1(1 + zg^{\tau_1+1}(z)) + \delta_2(1 + zg^{\tau_2}(z))\Big) \\
&\qquad + C\delta(1 + g(z))(1 - F(z)) \\
&\quad \le C(1 - F(z))\Big(\delta(1 + zg^2(z)) + \delta_1(1 + zg^{\tau_1+1}(z)) + \delta_2(1 + zg^{\tau_2}(z))\Big).
\end{aligned}$$

Similarly, we can prove the lower bound as follows:

$$\begin{aligned}
&\mathrm{P}(W > z) - (1 - F(z)) \\
&\quad \ge \; - C(1 - F(z))\Big(\delta(1 + zg^2(z)) + \delta_1(1 + zg^{\tau_1+1}(z)) + \delta_2(1 + zg^{\tau_2}(z))\Big).
\end{aligned}$$

This completes the proof of Theorem 2.1. □

4.2. *Proof of Lemma 4.1.* For $w \ge 0$, by the monotonicity of $g(\cdot)$, we have

$$\begin{aligned}
1 - F(w) &= \int_w^\infty p(t)dt \\
&= c_1 \int_w^\infty e^{-G(t)} dt \\
&= c_1 \int_w^\infty \frac{1}{g(t)} e^{-G(t)} dG(t) \\
&\le \frac{c_1}{g(w)} e^{-G(w)} \\
&= \frac{p(w)}{g(w)}.
\end{aligned}$$

Let $H(w) = 1 - F(w) - p(w)/c_1$; then,

$$H'(w) = p(w)(g(w)/c_1 - 1).$$

Note that $g(w)/c_1 = 1$ has at most one solution in $(0, +\infty)$ and that $g(0) = 0$; then, $H(w)$ takes the maximum at either 0 or $+\infty$. We have

$$H(w) \le \max\big\{H(0), \lim_{w\to\infty} H(w)\big\} \le 0.$$

This proves (4.1). The inequality (4.2) can be obtained similarly.

To finish the proof, we need to prove that for $w \geq 0$,

$$\frac{p(w)}{1+g(w)} \leq c_3(1-F(w)). \tag{4.17}$$

Consider

$$\zeta(w) = \frac{1}{1+g(w)} e^{-G(w)}. \tag{4.18}$$

As $g'(w) \leq c_3(1+g(w))$, we have

$$-\zeta'(w) = \frac{g(w)}{1+g(w)} e^{-G(w)} + \frac{g'(w)}{(1+g(w))^2} e^{-G(w)} \leq \max(1, c_3) e^{-G(w)}.$$

As $g(w)$ is nondecreasing and $g(w) > 0$ for $w > 0$, then $G(w) = \int_0^w g(t)dt \to \infty$ as $w \to \infty$. Therefore, $\lim_{w\to\infty} p(w) = 0$. Taking the integration on both sides yields

$$\zeta(w) = -\int_w^\infty \zeta'(t)dt \leq \max(1, c_3) \int_w^\infty e^{-G(t)} dt,$$

which leads to (4.17). This completes the proof.

4.3. *Proofs of Propositions 4.1 and 4.2.* Throughout this subsection, we assume that conditions (A1)–(A4) are satisfied, and $g(t) = G'(t)$. To prove Propositions 4.1 and 4.2, we first present some preliminary lemmas, whose proofs are postponed in the following several subsections. Lemmas 4.2 and 4.3 give the properties of g and g'.

LEMMA 4.2. *Assume that $0 < \delta \leq 1$. Then, we have*

$$\sup_{|t|\leq\delta} |g(w+t)| \leq c_2(|g(w)| + \mu_1), \tag{4.19}$$

where $\mu_1 = \max(g(1), |g(-1)|) + 1$.

Also, for $w > s > 0$ and any positive number $a > 1$, there exists $b(a)$ depending on a, c_2 and c_3, such that

$$g(w) - g(w-s) \leq \frac{1}{a} g(w) + b(a)(g(s)+1), \tag{4.20}$$

where one can choose

$$b(a) = \big((2c_2) + \cdots + (2c_2)^{m(a)}\big) + 1/a,$$

and $m(a) = [\log(ac_3 + 1)] + 1$.

LEMMA 4.3. *For $w \geq 0$ and any $a > 0$, we have*

$$g'(w) \leq \frac{1}{a} g(w) + c_3(g(ac_3) + 1) + 1/a. \tag{4.21}$$

Let W be the random variable defined as in 2.1. For $0 \leq \tau \leq \max(2, \tau_1 + 1, \tau_2)$ and $s > 0$, Lemmas 4.4 and 4.5 give the properties of $\mathrm{E}|g(W)|^\tau$, $\mathrm{E}|g(W)|^\tau e^{G(W)} I(0 \leq W \leq s)$ and $\mathrm{E}|g(W)|^\tau e^{G(W) - G(W-s)} I(W > s)$, which play a key role in the proof of Propositions 4.1 and 4.2.

For $s > 0$, define

$$f(w, s) = \begin{cases} e^{G(w) - G(w-s)} - 1, & w > s, \\ e^{G(w)} - 1, & 0 \leq w \leq s, \\ 0, & w \leq 0. \end{cases} \tag{4.22}$$

LEMMA 4.4. *For $0 \leq \tau \leq \max(2, \tau_1 + 1, \tau_2)$, we have*

$$\mathrm{E}|g(W)|^\tau \leq C. \tag{4.23}$$

Moreover, for $s > 0$ and $\delta \leq 1$, we have

$$\mathrm{E}\left(e^{G(W) - G(W-s)} g^\tau(W) I(W > s)\right) \leq C(1 + g^\tau(s))(\mathrm{E}(f(W, s)) + 1), \tag{4.24}$$

and

$$\mathrm{E}\left(e^{G(W)} g^\tau(W) I(0 \leq W \leq s)\right) \leq C(1 + g^\tau(s))(\mathrm{E}(f(W, s)) + 1). \tag{4.25}$$

LEMMA 4.5. *Let $0 < \delta \leq 1$ and $s > 0$. Then, we have*

$$\begin{aligned} &\mathrm{E}(f(W, s) + 1) \\ &\leq C(1 + s) \exp\Big\{C\Big(\delta\big(1 + sg^2(s)\big) + \delta_1\big(1 + sg^{\tau_1 + 1}(s)\big) \\ &\qquad + \delta_2\big(1 + sg^{\tau_2}(s)\big)\Big)\Big\}. \end{aligned} \tag{4.26}$$

The next Lemma gives the properties of the Stein solution.

LEMMA 4.6. *Let f_z be the solution to Stein's equation (4.3). Then, for $z \geq 0$,*

$$|f_z(w)g(w)| \leq \begin{cases} 1 - F(z), & w \leq 0, \\ F(z), & w > 0, \end{cases} \tag{4.27}$$

$$f_z(w) \leq \begin{cases} (1 - F(z))/c_1, & w \leq 0, \\ F(z)/c_1, & w > 0, \end{cases} \tag{4.28}$$

and

$$|f_z'(w)| \leq \begin{cases} 2(1 - F(z)), & w \leq 0, \\ 1, & 0 < w \leq z, \\ 2F(z), & w > z. \end{cases} \tag{4.29}$$

LEMMA 4.7. *For $z > 0$ and $0 \leq \tau \leq \max(2, \tau_1 + 1, \tau_2)$,*

$$\mathrm{E}(f_z(W)|g(W)|^{\tau}) \leq C(1 + zg^{\tau}(z))(1 - F(z)), \tag{4.30}$$

provided that $\max(\delta, \delta_1, \delta_2) \leq 1$ *and* $\delta z g^2(z) + \delta_1 z g^{\tau_1+1}(z) + \delta_2 z g^{\tau_2}(z) \leq 1$.

We are now ready to give the proofs of Propositions 4.1 and 4.2.

PROOF OF PROPOSITION 4.1. Recalling (2.9), we have

$$\begin{aligned} I_1 &\leq d_0 \,\mathrm{E}\Big(\sup_{|t|\leq\delta}\big|f_z(W+t)g(W+t) - f_z(W)g(W)\big|\Big) \\ &\leq \delta d_0 \,\mathrm{E} \sup_{|t|\leq\delta} \big|(f_z(W+t)g(W+t))'\big| . \end{aligned} \tag{4.31}$$

We first prove (4.7). By Lemma 4.6, $\|f_z\| \leq 1/c_1$ and $\|f_z'\| \leq 2$. Thus, for

$\delta \le 1$,

$$\begin{aligned}
&\mathrm{E}\Big(\sup_{|t|\le\delta} |\big(f_z(w+t)g(w+t)\big)'|\Big)\\
&\le \mathrm{E}\Big(\sup_{|t|\le\delta}\big(\big|f_z(W+t)g'(W+t)\big| + \big|f_z'(W+t)g(W+t)\big|\big)\Big)\\
&\le (2+1/c_1)\,\mathrm{E}\Big(\sup_{|t|\le\delta}\big(|g'(W+t)|+|g(W+t)|\big)\Big)\\
&\le 4c_3(1+1/c_1)(1+c_2)\big(\mathrm{E}|g(W)|+\mu_1\big),
\end{aligned}\tag{4.32}$$

where in the last inequality we use (2.6) and Lemma 4.2. This proves (4.7) by (4.23), (4.31) and (4.32).

Next, we prove (4.8). Similarly, we first calculate the following term:

$$\mathrm{E}\Big(\sup_{|t|\le\delta} |(f_z(W+t)g(W+t))'|\Big).$$

Note that

$$(f_z(w)g(w))' = \begin{cases} \frac{p(w)g(w)+F(w)g'(w)+F(w)g^2(w)}{p(w)}(1-F(z)), & w \le z,\\ \frac{-p(w)g(w)+(1-F(w))g'(w)+(1-F(w))g^2(w)}{p(w)}F(z), & w > z. \end{cases}\tag{4.33}$$

For $w+t \le 0$, we have

$$\begin{aligned}
&|(f_z(w+t)g(w+t))'|\\
&\le (1-F(z))\Big(2|g(w+t)| + \frac{g'(w+t)}{\max\{c_1, |g(w+t)|\}}\Big)\\
&\le (1-F(z))\,(2\,|g(w+t)| + c_3(1+1/c_1))\\
&\le C(1-F(z))(|g(w)|+1).
\end{aligned}$$

Thus, by (4.19) and (4.23),

$$\mathrm{E}\Big(\sup_{|t|\le\delta} |(f_z(W+t)g(W+t))'| I(W+t\le 0)\Big) \le C(1-F(z)).\tag{4.34}$$

For $w+t>z$, and $|t|\le\delta$, again by (4.19), we have

$$\begin{aligned}
&|(f_z(w+t)g(w+t))'|\\
&\le F(z)\Big(|g(w+t)|+\frac{1-F(z)}{p(z)}(|g'(w+t)|+|g(w+t)|^2)\Big)\\
&\le C(1+|g(w+t)|^2)\\
&\le C(|g(w)|^2+1).
\end{aligned} \tag{4.35}$$

Hence, by Lemmas 4.4 and 4.5, we have

$$\begin{aligned}
&\mathrm{E}\sup_{|t|\le\delta}|(f_z(W+t)g(W+t))'|I(W+t\ge z)\\
&\le\ \mathrm{E}\sup_{|t|\le\delta}\big(|g(W+t)|^2I(W+t>z)\big)\\
&\le C\,\mathrm{E}\Big(\big(|g(W)|^2+1\big)I(W>z-\delta)\Big)\\
&\le Cp(z-\delta)\,\mathrm{E}\left(e^{G(W)-G(W-z+\delta)}|g(W)|^2I(W>z-\delta)\right)\\
&\le Ce^{\delta g(z)}p(z)(1+g^2(z))\,\mathrm{E}\left(e^{G(W)-G(W-z+\delta)}I(W>z-\delta)\right)\\
&\le Ce^{\delta g(z)}(1+zg^2(z)).
\end{aligned} \tag{4.36}$$

Also note that $\delta g(z)\le\delta+\delta zg^2(z)\le 2$ for $z\ge 1$ and $\delta g(z)\le\mu_1$ for $0\le z\le 1$. Thus, (4.36) yields

$$\mathrm{E}\bigg(\sup_{|t|\le\delta}|(f_z(W+t)g(W+t))'|I(W+t>z)\bigg)\le C(1+zg^2(z))(1-F(z)). \tag{4.37}$$

For $w+t\in(0,z)$ and $|t|\le\delta$, noting that $\delta g(z)\le\max(2,\mu_1)$, by (4.19), we have

$$\begin{aligned}
&|(f_z(w+t)g(w+t))'|\\
&\qquad\le C(1-F(z))e^{G(w+t)}\big(1+g(w+t)^2\big)\\
&\qquad\le C(1-F(z))e^{G(w)+\delta g(z)}\big(1+|g(w)|^2\big)\\
&\qquad\le C(1-F(z))e^{G(w)}\big(1+|g(w)|^2\big).
\end{aligned} \tag{4.38}$$

By Lemmas 4.4 and 4.5 and (4.19), we have

$$
\begin{aligned}
&\mathrm{E}\Big(\sup_{|t|\le\delta}|(f_z(W+t)g(W+t))'|I(0\le W+t\le z)\Big)\\
&\quad\le C(1-F(z))\,\mathrm{E}e^{G(W)}(1+|g(W)|^2)I(-\delta\le W\le z+\delta)\\
&\quad\le Ce^{\mu_1}(1+\mu_1^2)(1-F(z))\\
&\qquad\quad+C(1-F(z))\,\mathrm{E}e^{G(W)}(1+|g(W)|^2)I(0\le W\le z+\delta)\\
&\quad\le C(1-F(z))\big(1+(z+\delta)g^2(z+\delta)\big)\\
&\quad\le C(1-F(z))(1+zg^2(z)).
\end{aligned}
\tag{4.39}
$$

Putting together (4.34), (4.37) and (4.39) gives

$$
\mathrm{E}\Big(\sup_{|t|\le\delta}|(f_z(W+t)g(W+t))'|\Big)\le C(1+zg^2(z))(1-F(z)). \tag{4.40}
$$

This completes the proof of (4.8). ☐

Proof of Proposition 4.2. By Lemma 4.6, we have $\|f_zg\|\le 1$; thus, by (4.23),

$$
I_2+I_3\le C\,\mathrm{E}|\,\mathrm{E}(\hat{K}_1|W)-1|\le C\delta_1\big(\mathrm{E}(|g(W)|^{\tau_1})+1\big)\le C\delta_1.
$$

To bound I_4, by (2.8), (4.23) and (4.28), we have

$$
I_4\le C\delta_2.
$$

This proves (4.9).

We now move to prove (4.10) and (4.11). As to I_2, By (2.5) and Lemma 4.6, for $z\ge 0,\max(\delta,\delta_1,\delta_2)\le 1$ and $\delta zg^2(z)+\delta_1 zg^{\tau_1+1}(z)+\delta_2 zg^{\tau_2}(z)\le 1$, we have

$$
\begin{aligned}
I_2&\le\delta_1\,\mathrm{E}\,(f_z(W)|g(W)|(|g(W)|^{\tau_1}+1)\\
&\le C\delta_1\,\mathrm{E}\,\big(f_z(W)(1+|g(W)|^{\tau_1+1})\big)\\
&\le C\delta_1(1+zg^{\tau_1+1}(z))(1-F(z)).
\end{aligned}
\tag{4.41}
$$

As to I_3, note that

$$
I(W>z)\le\frac{e^{G(W)-G(W-z)}}{e^{G(z)}}I(W>z).
$$

By Lemmas 4.4 and 4.5,

$$
\begin{aligned}
&\mathrm{E}\left((1+|g(W)|^{\tau_1})I(W>z)\right)\\
&\quad\le Cp(z)\,\mathrm{E}\left(e^{G(W)-G(W-z)}(1+|g(W)|^{\tau_1})I(W>z)\right)\\
&\quad\le C(1+zg^{\tau_1}(z))p(z)\\
&\quad\le C(1+zg^{\tau_1+1}(z))(1-F(z)),
\end{aligned}
\tag{4.42}
$$

where we use (4.1) in the last inequality. Thus, by Lemma 4.6,

$$
\begin{aligned}
I_3 &\le \delta_1\big(1-F(z)\big)E\big(|g(W)|^{\tau_1}+1\big)\\
&\qquad + \delta_1\,\mathrm{E}\Big(\big(|g(W)|^{\tau_1}+1\big)I(W>z+\delta)\Big)\\
&\le \delta_1\big(1-F(z)\big)E\big(|g(W)|^{\tau_1}+1\big)\\
&\qquad + \delta_1\,\mathrm{E}\Big(\big(|g(W)|^{\tau_1}+1\big)I(W>z)\Big)\\
&\le C\delta_1(1+zg^{\tau_1+1}(z))(1-F(z)).
\end{aligned}
\tag{4.43}
$$

(4.10) now follows by (4.41) and (4.43).

As to I_4, because $|R(W)|\le\delta_2(1+|g(W)|^{\tau_2})$, by (4.30), we have

$$
I_4 \le C\delta_2(1+zg^{\tau_2}(z))(1-F(z)).
\tag{4.44}
$$

This completes the proof of Proposition 4.2. □

4.4. *Proof of Lemmas 4.2 and 4.3.*

PROOF OF LEMMA 4.2. The inequality (4.19) can be derived immediately from (2.5). Meanwhile, (4.20) remains to be shown. For $a>1$, consider two cases.

Case 1. If $s<w\le(ac_3+1)s$, denote $m:=m(a)=[\log_2(ac_3+1)]+1$. As g is nondecreasing and by (2.5), we have

$$
g(w)\le g\big(2^m s\big)\le 2c_2 g\big(2^{m-1}s\big)+c_2.
$$

By induction, we have

$$
\begin{aligned}
g(w) &\le (2c_2)^m g(s)+c_2(1+(2c_2)+\cdots+(2c_2)^{m-1})\\
&\le b(a)(g(s)+1),
\end{aligned}
\tag{4.45}
$$

where $b(a) = 2c_2(1 + (2c_2) + \cdots + (2c_2)^{m(a)-1}) + 1/a$.
Case 2. If $w > (ac_3 + 1)s$, by (2.6), we have

$$g(w) - g(w-s) = \int_0^s g'(w-t)dt$$
$$\le c_3 \int_0^s \frac{1 + g(w-t)}{1 + (w-t)} dt \tag{4.46}$$
$$\le \frac{1}{a}(g(w) + 1).$$

By (4.45) and (4.46), this completes the proof. □

PROOF OF LEMMA 4.3. Recall that (2.6) states that for $w \ge 0$,

$$g'(w) \le c_3 \left(\frac{1 + g(w)}{1 + w} \right).$$

Fix $a > 0$. When $w > ac_3$, we have

$$g'(w) \le \frac{1}{a}(g(w) + 1).$$

When $w \le ac_3$, by the monotonicity property of g, we have

$$g'(w) \le c_3(g(ac_3) + 1).$$

This completes the proof. □

4.5. *Proofs of Lemmas 4.4 and 4.5.* Before giving the proofs of Lemmas 4.4 and 4.5, we first consider a ratio property of $f(w, s)$. It is easy to see that $f(w, s)$ is absolutely continuous with respect to both w and s, and the partial derivatives are

$$\frac{\partial}{\partial w} f(w,s) = e^{G(w)-G(w-s)}(g(w) - g(w-s))I(w > s) \tag{4.47}$$
$$+ e^{G(w)} g(w) I(0 \le w \le s)$$

and

$$\frac{\partial}{\partial s} f(w,s) = e^{G(w)-G(w-s)} g(w-s) I(0 < s \le w). \tag{4.48}$$

LEMMA 4.8. *Let $f(w) := f(w,s)$ be defined as in (4.22). For $\delta|g(w)| \le d_1$ and $\delta \le 1$, we have*

$$\sup_{|u|\le\delta}\Big|\frac{f(w+u)+1}{f(w)+1}\Big| I(w+u\ge 0) \le \mu_2, \tag{4.49}$$

where $\mu_2 = \exp(c_2(d_1+\mu_1)+\mu_1)$. Moreover, we have

$$\sup_{|u|\le\delta} |f''(w+u)| \le \mu_3 (g^2(w)+1)(f(w)+1). \tag{4.50}$$

where $\mu_3 = 2c_2^2(c_3+1)(\mu_1^2+1)\mu_2$.

PROOF. Recall that $\mu_1 = \max(g(1), |g(-1)|)+1$. When $w+u\ge 0$ and $w \ge 0$, as g is nondecreasing, we have

$$\begin{aligned}\sup_{|u|\le\delta}\Big|\frac{f(w+u)+1}{f(w)+1}\Big| &\le e^{G(w+\delta)-G(w)}\\ &\le e^{\delta|g(w+\delta)|} \le e^{c_2(d_1+\mu_1)},\end{aligned}$$

where in the last inequality we use (4.19). When $w+u\ge 0$, $w<0$ and $|u|\le\delta$, we have $0\le w+u<\delta\le 1$; hence, by the nondecreasing property of g,

$$\sup_{|u|\le\delta}\Big|\frac{f(w+u)+1}{f(w)+1}\Big| \le \sup_{|u|\le\delta} e^{G(w+u)} \le e^{G(\delta)} \le e^{\mu_1}.$$

This proves (4.49).

For $f''(w)$, by (4.47),

$$\begin{aligned}f''(w) = {}& e^{G(w)-G(w-s)}\big(g(w)-g(w-s)\big)^2 I(w>s)\\ &+ e^{G(w)-G(w-s)}(g'(w)-g'(w-s))I(w>s)\\ &+ e^{G(w)}g^2(w)I(0\le w\le s)\\ &+ e^{G(w)}g'(w)I(0\le w\le s).\end{aligned}$$

As g is nondecreasing, we have $g'(w-s)\ge 0$; thus, $g'(w)-g'(w-s)\le g'(w)$. For $w>s$, $0\le g(w)-g(w-s)\le g(w)$. Therefore,

$$f''(w) \le \big(g'(w)+g^2(w)\big)\big(f(w)+1\big)I(w\ge 0).$$

By (2.6), for $c_3 > 1$, we have

$$g^2(w) + g'(w) \le g^2(w) + c_3\big(1 + g(w)\big) \le 2(c_3 + 1)\big(g^2(w) + 1\big).$$

Hence,

$$f''(w) \le 2(c_3 + 1)(g^2(w) + 1)(f(w) + 1).$$

By (4.19) and (4.49), we have

$$\sup_{|u|\le\delta} |f''(w + u)| \le \mu_3(g^2(w) + 1)(f(w) + 1),$$

where $\mu_3 = 2c_2^2(c_3+1)(\mu_1^2+1)\mu_2$. This completes the proof of Lemma 4.8. $\square$

We now give the proofs of Lemmas 4.4 and 4.5.

Proof of Lemma 4.4. We first prove (4.23). Without loss of generality, we consider only the case where $\tau \ge 2$. As $\delta|g(W)| \le d_1$, we have $\mathrm{E}|g(W)|^\tau < \infty$. To bound $\mathrm{E}|g(W)|^\tau$, without loss of generality, we consider only $\mathrm{E}g^\tau(W)I(W \ge 0)$. Let $g_+(w) := g(w)I(w \ge 0)$. As $g(0) = 0$ and g is differentiable, we find that $g_+(w)$ is absolutely continuous. By (2.1), we have

$$\begin{aligned}
\mathrm{E}g^\tau(W)I(W \ge 0) &= \mathrm{E}g(W)\cdot g_+^{\tau-1}(W)\\
&= (\tau - 1)\,\mathrm{E}\int_{|u|\le\delta} g_+^{\tau-2}(W + u)g'(W + u)I(W + u \ge 0)\hat{K}(u)du\\
&\quad + \mathrm{E}R(W)g_+^{\tau-1}(W)\\
&:= Q_1 + Q_2, \qquad (4.51)
\end{aligned}$$

where

$$\begin{aligned}
Q_1 &= (\tau - 1)\,\mathrm{E}\int_{|u|\le\delta} g_+^{\tau-2}(W + u)g'(W + u)I(W + u \ge 0)\hat{K}(u)du,\\
Q_2 &= \mathrm{E}R(W)g_+^{\tau-1}(W).
\end{aligned}$$

The following inequality is well known: for any $a > 0, x, y \ge 0$ and $\tau > 1$

$$x^{\tau-1}y \le \frac{\tau - 1}{a\tau}x^\tau + \frac{a^{\tau-1}}{\tau}y^\tau. \qquad (4.52)$$

For the first term Q_1, by (2.6), we have

$$g'(w+u) \le c_3\big(1+|g(w+u)|\big).$$

Thus, for $w+u\ge 0$,

$$\begin{aligned}
&\sup_{|u|\le\delta} g_+^{\tau-2}(w+u)g'(w+u)\\
&\quad\le c_3 \sup_{|u|\le\delta}\big(g_+^{\tau-1}(w+u)+g_+^{\tau-2}(w+u)\big)\\
&\quad\le 2c_3 \sup_{|u|\le\delta}\big(g_+^{\tau-1}(w+u)+1\big)\\
&\quad\le \frac{1-\alpha}{8\times(2c_2)^\tau d_0(\tau-1)}\sup_{|u|\le\delta}|g(w+u)|^\tau + D_{1,0},
\end{aligned}$$

where we use (4.52) in the last inequality. Here and in the sequel, $D_{1,0}$, $D_{2,0}$, etc. denote constants depending on $c_2, c_3, d_0, d_1, \mu_1, \alpha$ and τ. By (4.19), we have

$$\sup_{|u|\le\delta}|g(w+u)|^\tau \le (2c_2)^\tau(|g(w)|^\tau+\mu_1^\tau).$$

Then, by (2.9), we have

$$Q_1 \le \frac{1-\alpha}{8}\,\mathrm{E}|g(W)|^\tau + D_{2,0}. \tag{4.53}$$

For Q_2, by (2.11) and using (4.52) again, we have

$$Q_2 \le \alpha\,\mathrm{E}g_+^\tau(W) + \frac{1-\alpha}{4}\,\mathrm{E}g_+^\tau(W) + \Big(\frac{4}{1-\alpha}\Big)^{\tau-1}. \tag{4.54}$$

Hence, by (4.51), (4.53) and (4.54), we have

$$\mathrm{E}g_+^\tau(W) \le \frac{1}{6}\,\mathrm{E}|g(W)|^\tau + D_{3,0}.$$

Similarly, we have

$$\mathrm{E}g_-^\tau(W) \le \frac{1}{6}\,\mathrm{E}|g(W)|^\tau + D_{4,0}.$$

Combining the two foregoing inequalities yields (4.23).

As to (4.24) and (4.25), we first consider the case $\tau \geq 2$. Write $f(w) := f(w,s)$. By (2.1) and (4.47), we have

$$
\begin{aligned}
&\mathrm{E}\big(g(W)^\tau f(W)\big)\\
&= \mathrm{E}\int_{|u|\le\delta} g^\tau(W+u)e^{G(W+u)}I(0\le W+u\le s)\hat{K}(u)du\\
&\quad + \mathrm{E}\int_{|u|\le\delta} g^{\tau-1}(W+u)\big(g(W+u)-g(W+u-s)\big)\\
&\qquad\qquad \times e^{G(W+u)-G(W+u-s)}I(W+u>s)\hat{K}(u)du\\
&\quad + (\tau-1)\,\mathrm{E}\int_{|u|\le\delta} g^{\tau-2}(W+u)g'(W+u)f(W+u)\hat{K}(u)du\\
&\quad + \mathrm{E}R(W)g^{\tau-1}(W)f(W)\\
&:= M_1+M_2+M_3+M_4,
\end{aligned}
\tag{4.55}
$$

where

$$
\begin{aligned}
M_1 &= \mathrm{E}\int_{|u|\le\delta} g^\tau(W+u)e^{G(W+u)}I(0\le W+u\le s)\hat{K}(u)du,\\
M_2 &= \mathrm{E}\int_{|u|\le\delta} g^{\tau-1}(W+u)\big(g(W+u)-g(W+u-s)\big)\\
&\qquad\qquad \times e^{G(W+u)-G(W+u-s)}I(W+u>s)\hat{K}(u)du,\\
M_3 &= (\tau-1)\,\mathrm{E}\int_{|u|\le\delta} g^{\tau-2}(W+u)g'(W+u)f(W+u)\hat{K}(u)du,\\
M_4 &= \mathrm{E}R(W)g^{\tau-1}(W)f(W).
\end{aligned}
\tag{4.56}
$$

We next give the bounds of M_1, M_2, M_3 and M_4. For M_1, by (2.9) and (4.49) and noting that g is nondecreasing, we have

$$
\begin{aligned}
M_1 &\le d_0 g^\tau(s)\,\mathrm{E}\sup_{|u|\le\delta}(f(W+u)+1)I(0\le W+u\le s)\\
&\le d_0\mu_2 g^\tau(s)\,\mathrm{E}(f(W)+1).
\end{aligned}
\tag{4.57}
$$

To bound M_2, we first give the bound of $g(W+u)$ and $g(W+u)-g(W+u-s)$ for $|u|\le\delta$. By (4.19), we have

$$
\sup_{|u|\le\delta}|g(W+u)| \le c_2(|g(W)|+\mu_1). \tag{4.58}
$$

Furthermore, by (4.20), for $w+u>s$, there exists a constant D_1 depending on $c_2, c_3, d_0, d_1, \mu_1, \alpha$ and τ such that

$$\begin{aligned}\sup_{|u|\le\delta}&|g(w+u)-g(w+u-s)| \\ &\le \frac{1-\alpha}{2^{\tau+3}d_0\mu_2c_2^\tau}\sup_{|u|\le\delta}|g(w+u)|+D_1(g(s)+1).\end{aligned}\tag{4.59}$$

By (4.52), (4.58) and (4.59), we have

$$\begin{aligned}\sup_{|u|\le\delta}&\big|g(W+u)^{\tau-1}(g(W+u)-g(W+u-s))\big| \\ &\le \Big(\frac{1-\alpha}{2^{\tau+3}d_0\mu_2c_2^\tau}\sup_{|u|\le\delta}|g(W+u)|+D_1(g(s)+1)\Big)\sup_{|u|\le\delta}|g(W+u)|^{\tau-1} \\ &\le \frac{1-\alpha}{2^{\tau+2}d_0\mu_2c_2^\tau}\sup_{|u|\le\delta}|g(W+u)|^\tau+\frac{2^{\tau+3}d_0\mu_2c_2^\tau}{\tau(1-\alpha)}\times D_1^\tau(1+g(s))^\tau \\ &\le \frac{1-\alpha}{4d_0\mu_2}\Big(|g(W)|^\tau+\mu_1^\tau\Big)+\frac{2^{\tau+3}d_0\mu_2c_2^\tau}{\tau(1-\alpha)}\times D_1^\tau(1+g(s))^\tau \\ &\le \frac{1-\alpha}{4d_0\mu_2}|g(W)|^\tau+D_2(1+g^\tau(s)),\end{aligned}$$

where

$$D_2=\frac{2^{2\tau+3}d_0\mu_2c_2^\tau}{\tau(1-\alpha)}\times D_1^\tau+\frac{(1-\alpha)\mu_1^\tau}{4d_0\mu_2}.$$

By (2.9) and (4.49), we have

$$\begin{aligned}M_2\le{}&\frac{1-\alpha}{4}\,\mathrm{E}|g(W)|^\tau(f(W)+1) \\ &+d_0\mu_2D_2(1+g^\tau(s))\,\mathrm{E}(f(W)+1).\end{aligned}\tag{4.60}$$

For M_3, by Lemma 4.3 and similar to M_2, we have

$$\begin{aligned}M_3\le{}&\frac{1-\alpha}{4}\,\mathrm{E}|g(W)|^\tau(f(W)+1) \\ &+D_3(1+g^\tau(s))\,\mathrm{E}(f(W)+1),\end{aligned}\tag{4.61}$$

where D_3 is a finite constant depending on $c_2, c_3, d_0, d_1, \mu_1, \alpha$ and τ.

For M_4, by (2.11) and (4.52), we have

$$\begin{aligned}M_4&\le\alpha\,\mathrm{E}|g(W)|^\tau f(W)+\alpha\,\mathrm{E}|g(W)|^{\tau-1}f(W) \\ &\le\Big(\alpha+\frac{1-\alpha}{4}\Big)\,\mathrm{E}|g(W)|^\tau f(W)+\Big(\frac{4\alpha}{1-\alpha}\Big)^{\tau-1}\mathrm{E}f(W).\end{aligned}\tag{4.62}$$

By (4.55), (4.57) and (4.60)–(4.62), we have

$$\begin{aligned}\mathrm{E}|g(W)|^\tau f(W) \le{}& \left(\alpha + \frac{3(1-\alpha)}{4}\right)\mathrm{E}|g(W)|^\tau f(W)\\ &+ (D_4 + \mathrm{E}|g(W)|^\tau)(1+g^\tau(s))\,\mathrm{E}(f(W)+1),\end{aligned}$$

where D_4 is a constant depending on $c_2, c_3, d_0, d_1, \mu_1, \alpha$ and τ. Rearranging the inequality gives

$$\mathrm{E}|g(W)|^\tau f(W) \le \frac{4(D_4 + \mathrm{E}|g(W)|^\tau)}{1-\alpha}(1+g^\tau(s))\,\mathrm{E}(f(W)+1). \tag{4.63}$$

Combining (4.23) and (4.63), we have

$$\mathrm{E}|g(W)|^\tau (f(W)+1) \le D_5(1+g^\tau(s))\,\mathrm{E}(f(W)+1), \tag{4.64}$$

where D_5 is a constant depending on $c_2, c_3, d_0, d_1, \mu_1, \alpha$ and τ. This proves (4.24) and (4.25) for $\tau \ge 2$.

For $0 \le \tau < 2$ with $\mathrm{E}|g(W)|^2 < \infty$. By the Cauchy inequality, we have

$$(1+g^{2-\tau}(s))|g(w)|^\tau \le 1 + g^2(s) + 2g^2(w).$$

Recalling that for $s > 0$ and $g(s) > 0$,

$$|g(w)|^\tau \le g^\tau(s) + \frac{1+2g^2(w)}{1+g^{2-\tau}(s)}. \tag{4.65}$$

By (4.64) with $\tau = 2$, we have

$$\mathrm{E}|g(W)|^2 (f(W)+1) \le D_6(1+g^2(s))\,\mathrm{E}(f(W)+1), \tag{4.66}$$

where D_6 is a constant depending on $c_2, c_3, d_0, d_1, \mu_1, \alpha$ and τ.

Thus, for $0 \le \tau < 2$, by (4.65) and (4.66), we have

$$\begin{aligned}\mathrm{E}|g(W)|^\tau (f(W)+1) &\le g^\tau(s)\,\mathrm{E}(f(W)+1)\\ &\qquad + \frac{\mathrm{E}(f(W)+1) + 2\,\mathrm{E}g^2(W)(f(W)+1)}{1+g^{2-\tau}(s)}\\ &\le D_7(1+g^\tau(s))\,\mathrm{E}(f(W)+1),\end{aligned}$$

where D_7 is a constant depending on $c_2, c_3, d_0, d_1, \mu_1, \alpha$ and τ. This completes the proof together with (4.64). $\square$

PROOF OF LEMMA 4.5. Let $h(s) = \mathrm{E} f(W, s)$ and let $f(W) := f(W, s)$. By (4.47) and (4.48), for $s > 0$, we have

$$\begin{aligned} h'(s) &= \mathrm{E}\left(e^{G(W)-G(W-s)} g(W-s) I(W > s)\right) \\ &= \mathrm{E}(f(W)g(W)) + E(g(W)I(W > 0)) - E(f'(W)). \end{aligned}$$

We first show that $h'(s)$ can be bounded by a function of $h(s)$. We then solve the differential inequality to obtain the bound of $h(s)$, using an idea similar to that in the proof of Lemma 4.4.

By (2.1), we have

$$\begin{aligned} &\mathrm{E}(f(W)g(W)) - \mathrm{E}(f'(W)) \\ &\quad = \mathrm{E}\Big(\int_{|u|\le\delta} \big(f'(W+u) - f'(W)\big)\hat{K}(u)du\Big) \\ &\qquad + \mathrm{E}\big(f'(W)(1 - \mathrm{E}(\hat{K}_1|W)\big) + \mathrm{E}(f(W)R(W)) \\ &\quad := T_1 + T_2 + T_3, \end{aligned} \tag{4.67}$$

where

$$\begin{aligned} T_1 &= \mathrm{E}\Big(\int_{|u|\le\delta} \big(f'(W+u) - f'(W)\big)\hat{K}(u)du\Big), \\ T_2 &= \mathrm{E} f'(W)(1 - \mathrm{E}(\hat{K}_1|W)), \\ T_3 &= \mathrm{E}(f(W)R(W)). \end{aligned}$$

We next give the bounds of T_1, T_2 and T_3.

i). The bound of T_1. By (4.50), we have

$$\begin{aligned} &\sup_{|u|\le\delta} |f'(w+u) - f'(w)| \\ &\le \delta \sup_{|u|\le\delta} |f''(w+u)| \\ &\le \delta\mu_3(g^2(w)+1)(f(w)+1). \end{aligned}$$

By (2.9), we have

$$\begin{aligned} |T_1| &\le \delta d_0 \mu_3 \, \mathrm{E}(g^2(W)+1)(f(W)+1) \\ &\le D_8 \delta(1+g^2(s)) \, \mathrm{E}(f(W)+1), \end{aligned} \tag{4.68}$$

where D_8 is a constant depending on $c_2, c_3, d_0, d_1, \mu_1$ and α.

ii). The bound of T_2. By (2.7) and Lemma 4.4, we have

$$
\begin{aligned}
|T_2| &\le \delta_1 \operatorname{E}\Big(|g(W)|\big(|g(W)|^{\tau_1}+1\big)(f(W)+1)\Big) \\
&\le 2\delta_1 \operatorname{E}\big(|g(W)|^{\tau_1+1}+1\big)\big(f(W)+1\big) \\
&\le D_9\delta_1(1+g^{\tau_1+1}(s)) \operatorname{E}\big(f(W)+1\big),
\end{aligned} \tag{4.69}
$$

where D_9 is a constant depending on $c_2, c_3, d_0, d_1, \mu_1, \tau_1$ and α.

iii). The bound of T_3. For T_3, by (2.8) and Lemma 4.4, we have

$$
\begin{aligned}
T_3 &\le \delta_2 \operatorname{E}(|g(W)|^{\tau_2}+1)f(W) \\
&\le D_{10}\delta_2(1+g^{\tau_2}(s)) \operatorname{E}\big(f(W)+1\big),
\end{aligned} \tag{4.70}
$$

where D_{10} is a constant depending on $c_2, c_3, d_0, d_1, \mu_1, \tau_2$ and α.

By (4.23), we have

$$
\operatorname{E}g(W)I(W>0) \le D_{11}, \tag{4.71}
$$

where D_{11} is a constant depending on $c_2, c_3, d_0, d_1, \mu_1$ and α. By (4.67)–(4.71), we have

$$
\begin{aligned}
h'(s) \le D_{11} + D_{12}\big(&\delta\big(1+g^2(s)\big)+\delta_1\big(1+g^{\tau_1+1}(s)\big)+\delta_2\big(1+g^{\tau_2}(s)\big)\big) \\
&\times \operatorname{E}\big(f(W)+1\big),
\end{aligned}
$$

where $D_{12} = \max\big(D_8, D_9, D_{10}\big)$. Therefore,

$$
\begin{aligned}
h'(s) \le\ & D_{12}\big(\delta\big(1+g^2(s)\big)+\delta_1\big(1+g^{\tau_1+1}(s)\big)+\delta_2\big(1+g^{\tau_2}(s)\big)\big)h(s) \\
&+ D_{11} + D_{12}\big(\delta\big(1+g^2(s)\big)+\delta_1\big(1+g^{\tau_1+1}(s)\big)+\delta_2\big(1+g^{\tau_2}(s)\big)\big),
\end{aligned}
$$

By solving the differential inequality and given that $s + sg^{\tau}(s) \le 1 + (1 + g^{-\tau}(1))sg^{\tau}(s)$ for $\tau > 0$ and $s \ge 0$, we have

$$
\begin{aligned}
\operatorname{E}(f(W)+1) \le C_1(1+s)\exp\Big\{C_2\Big(&\delta\big(1+sg^2(s)\big)+\delta_1\big(1+sg^{\tau_1+1}(s)\big) \\
&+\delta_2\big(1+sg^{\tau_2}(s)\big)\Big)\Big\},
\end{aligned}
$$

where C_1 and C_2 are constants depending on $c_2, c_3, d_0, d_1, \mu_1, \tau_1, \tau_2$ and α. This completes the proof. $\square$

4.6. *Proof of Lemmas 4.6 and 4.7 .*

PROOF OF LEMMA 4.6. Our first step is to prove (4.27). By (4.4), we have

$$f_z(w)g(w) = \begin{cases} \frac{F(w)g(w)(1-F(z))}{p(w)}, & w \le z, \\ \\ \frac{F(z)g(w)(1-F(w))}{p(w)}, & w > z. \end{cases} \tag{4.72}$$

Without loss of generality, we must consider only three case when $z > 0$:

1. $w \le 0$: By (4.2),
$$|f_z(w)g(w)| \le 1 - F(z).$$
2. $0 \le w \le z$: Since $w \le z$, $1 - F(z) \le 1 - F(w)$, thus by (4.1),
$$|f_z(w)g(w)| \le \frac{F(w)|g(w)|(1 - F(w))}{p(w)} \le F(w) \le F(z).$$
3. $w > z$: By (4.1),
$$|f_z(w)g(w)| \le F(z).$$

We can have a similar argument when $z \le 0$, which completes the proof of (4.27). Additionally, (4.28) can be shown similarly. (4.29) follows directly from (4.3) and (4.27). □

PROOF OF LEMMA 4.7. By (4.4),

$$\begin{aligned}
&\mathrm{E}(f_z(W)|g(W)|^\tau) \\
&= F(z)\,\mathrm{E}\left(\frac{1 - F(W)}{p(W)}|g(W)|^\tau I(W > z)\right) \\
&\quad + (1 - F(z))\,\mathrm{E}\left(\frac{F(W)}{p(W)}|g(W)|^\tau I(W < 0)\right) \\
&\quad + (1 - F(z))\,\mathrm{E}\left(\frac{F(W)}{p(W)}|g(W)|^\tau I(0 \le W \le z)\right) \\
&:= T_4 + T_5 + T_6.
\end{aligned}$$

i). For T_4, we first consider the case when $\tau \geq 1$. As $g(w)$ is increasing, $e^{G(w)-G(w-z)}$ is also increasing with respect to w; thus,

$$I(W > z) \leq \frac{e^{G(W)-G(W-z)} I(W > z)}{e^{G(z)}}.$$

By Lemma 4.5, we have $\max(\delta, \delta_1, \delta_2) \leq 1$ and z, satisfying that $\delta z g^2(z) + \delta_1 z g^{\tau_1+1}(z) + \delta_2 z g^{\tau_2}(z) \leq 1$,

$$\mathrm{E}(f(W, z) + 1) \leq C(1 + z).$$

Hence, by (4.1) and Lemma 4.4, we have

$$\begin{aligned}
T_4 &\leq C e^{-G(z)} \mathrm{E}|g(W)|^{\tau-1} e^{G(W)-G(W-z)} I(W > z) \\
&\leq C e^{-G(z)} (1 + g^{\tau-1}(z)) \mathrm{E}(f(W, z) + 1) \\
&\leq C e^{-G(z)} (1 + z g^{\tau-1}(z)) \\
&\leq C(1 + z g^{\tau}(z))(1 - F(z)),
\end{aligned} \tag{4.73}$$

for $\max(\delta, \delta_1, \delta_2) \leq 1$ and z, satisfying that $\delta z g^2(z) + \delta_1 z g^{\tau_1+1}(z) + \delta_2 z g^{\tau_2}(z) \leq 1$. If $0 \leq \tau < 1$, then $g^{\tau}(w) \leq 2\big(1 + g(w)\big)/\big(1 + g^{1-\tau}(z)\big)$ for $w > z$. Therefore, (4.73) also holds for $0 \leq \tau < 1$.

ii). As to T_5, because $F(w)/p(w) \leq 1/c_1$ for $w \leq 0$,

$$T_5 \leq (1 - F(z)) \mathrm{E}|g(W)|^{\tau} I(W < 0).$$

By (4.23), we have

$$T_5 \leq C(1 - F(z)) \tag{4.74}$$

for some constant C.

iii). We now bound T_6. By Lemmas 4.4 and 4.5,

$$\begin{aligned}
T_6 &\leq C(1 - F(z)) \mathrm{E} e^{G(W)} |g(W)|^{\tau} I(0 \leq W \leq z) \\
&\leq C(1 - F(z))(1 + g^{\tau}(z)) \mathrm{E} e^{G(W)} I(0 \leq W \leq z) \\
&\leq C(1 - F(z))(1 + z g^{\tau}(z)).
\end{aligned} \tag{4.75}$$

By (4.73)–(4.75), we have

$$\mathrm{E}(f_z(W)|g(W)|^{\tau}) \leq C(1 + z g^{\tau}(z))(1 - F(z)),$$

which completes the proof. ☐

4.7. *Proof of Remark 2.1.* In this subsection, we assume that the condition (2.11) is replaced by (2.15) and (2.16). The conclusion of Remark 2.1 follows from the proof of Theorem 2.1 and the following lemma.

LEMMA 4.9. *Let the conditions in Remark 2.1 be satisfied. Furthermore, $0 < \delta \leq 1$, and $s > 0$ such that $\delta s g^2(s) \leq 1$. For $0 \leq \tau \leq \max\{2, \tau_1 + 1, \tau_2\}$, inequalities (4.23)–(4.25) hold.*

PROOF. Recall that $s_0 = \max\big\{s : \delta s g^2(s) \leq 1\big\}$ and $\delta \leq 1$. We have

$$s_0 \geq s_1 \quad \text{and} \quad \delta s_1 g^2(s_1) = 1.$$

Following the proof of Lemma 4.4, it suffices to prove the following two inequalities.

For Q_2 defined in (4.51),

$$Q_2 \leq \Big(\alpha + \frac{1-\alpha}{4}\Big) \mathrm{E} g_+^\tau(W) + C, \tag{4.76}$$

and for M_4 defined in (4.56),

$$M_4 \leq \Big(\alpha + \frac{1-\alpha}{4}\Big) \mathrm{E}|g(W)|^\tau f(W) + \Big(\frac{4\alpha}{1-\alpha}\Big)^{\tau-1} \mathrm{E} f(W) + C. \tag{4.77}$$

For Q_2, by (2.15) and similar to (4.54), we have

$$Q_2 \leq \Big(\alpha + \frac{1-\alpha}{4}\Big) \mathrm{E} g_+^\tau(W) + \Big(\frac{4}{1-\alpha}\Big)^{\tau-1} + d_2\, \mathrm{E} g_+^\tau(W) I(W > \kappa).$$

For the last term, by (2.10) and (2.16) and noting that $0 \leq \tau \leq \max\{2, \tau_1 + 1, \tau_2\}$, we obtain

$$\begin{aligned} d_2\, \mathrm{E} g_+^\tau(W) I(W > \kappa) &\leq d_1^{-\tau} d_2 \delta^{-\tau} \mathrm{P}(W > \kappa) \\ &\leq d_1^{-\tau} d_3 \delta^{-\tau} \exp(-2 s_0 d_1^{-1} \delta^{-1}) \\ &\leq d_1^{-\tau} d_3 \sup_{\delta > 0} \delta^{-\tau} \exp(-2 s_1 d^{-1} \delta^{-1}) \\ &= d_3 \Big(\frac{\tau}{2 s_1}\Big)^\tau e^{-\tau}, \end{aligned} \tag{4.78}$$

where the equality holds when $\delta = 2s_1/(d_1\tau)$. The inequality (4.76) follows from (4.54) and (4.78).

As to M_4, by (2.15), we have

$$M_4 \le \left(\alpha + \frac{1-\alpha}{4}\right) \mathrm{E}|g(W)|^\tau f(W) + \left(\frac{4\alpha}{1-\alpha}\right)^{\tau-1} \mathrm{E} f(W) + d_2\,\mathrm{E}\big|g^\tau(W)\big|e^{G(W)-G(W-s)} I(W > \kappa).$$

For the last term, by (2.10) and (2.16) and noting that $g(\cdot)$ is nondecreasing and $s \le s_0$, similar to (4.78), we have

$$\begin{aligned} d_2\,\mathrm{E}\big|g^\tau(W)\big|e^{G(W)-G(W-s)} I(W > \kappa) &\\ \le d_1^{-\tau} d_2 \delta^{-\tau} e^{s d_1^{-1}\delta^{-1}} \mathrm{P}(W > \kappa) &\\ \le d_1^{-\tau} d_3 \delta^{-\tau} e^{-s_0 d_1^{-1}\delta^{-1}} &\\ \le d_1^{-\tau} d_3 \sup_{\delta>0} \delta^{-\tau} e^{-s_1 d_1^{-1}\delta^{-1}} &\\ = d_3 \left(\frac{\tau}{s_1}\right)^\tau e^{-\tau}, & \end{aligned} \tag{4.79}$$

where the equality holds when $\delta = s_1/(d_1\tau)$. Combining (4.62) and (4.79), inequality (4.77) holds. Following the proof of Lemma 4.4 and replacing (4.54) and (4.62) with (4.76) and (4.77), respectively, we complete the proof of Lemma 4.7. □

5. Proofs of Theorems 3.1 and 3.2.

5.1. *Proof of Theorem 3.1.* We apply Corollary 2.1. Following Chatterjee and Shao [8], let σ have p.d.f. (3.1) and I be a uniformly distributed random variable among $\{1, 2, ..., n\}$, and independent with σ. Let σ_i' be generated from the conditional distribution of σ_i given $\{\sigma_j, j \ne i\}$ and let $W' = W - \sigma_I + \sigma_I'$. Then, (W, W') is an exchangeable pair. Let $\Delta = W - W'$ and $g(w) = w^3/3$. Clearly, we have

$$|W| \le n^{1/4}, \quad |\Delta| = |W - W'| \le 2n^{-3/4} := \delta.$$

Thus, we have

$$\delta|g(W)| \le 2/3.$$

Let $\lambda = n^{-3/2}$ and $\hat{K}_1 = \Delta^2/(2\lambda)$. As shown in Chatterjee and Shao [8] (see Section 5), we have

$$\mathrm{E}(\Delta|W) = \lambda(g(W) + R(W)) \tag{5.1}$$

and

$$|\,\mathrm{E}(\hat{K}_1|W) - 1| \le n^{-1/2}(1 + W^2) \le \delta_1\big(|g(W)|^{2/3} + 1\big),$$

where

$$|R(W)| \le n^{-1/2}\left(\frac{2}{15}|W|^5 + |W| + \frac{2}{3}n^{-3/4}\right)$$

and $\delta_1 = 6n^{-1/2}$.

Moreover, $|\hat{K}_1| \le 2$ and

$$|R(W)| \le n^{-1/2}\left(\frac{1}{6}|W|^5 + 5\right) \le \delta_2|g(W)|^{5/3} + \delta_3,$$

where $\delta_2 = n^{-1/2}/6$ and $\delta_3 = 5n^{-1/2}$. Thus, $\delta_2|g(W)|^{2/3} \le 1/6$. Then,

$$|R(W)| \le \frac{1}{2}(|g(W)| + 1).$$

Therefore, all of the conditions in Theorem 2.1 are satisfied. Hence, Theorem 3.1 holds.

5.2. *Proof of Theorem 3.2.* In this subsection, we use Remark 2.1 to prove the result.

PROOF OF THEOREM 3.2. The proof is organized as follows. We first construct an exchangeable pair as in Chen [14].

For any $\sigma \in \Sigma, uv \in D$ and $s, t = 0, 1$, let σ_{uv}^{st} denote the configuration $\tau \in \Sigma$, such that $\tau_i = \sigma_i$ for $i \ne u, v$ and $\tau_u = s$, $\tau_v = t$. Let (σ'_u, σ'_v) be independent of (σ_u, σ_v) and follow the conditional distribution

$$\mathrm{P}(\sigma'_u = s, \sigma'_v = t|\sigma) = \frac{p(\sigma_{uv}^{st})}{\sum_{s,t\in\{0,1\}} p(\sigma_{uv}^{st})}.$$

Let $M = \sum_{i=1}^n \sigma_i$ and $M' = M - \sigma_u - \sigma_v + \sigma'_u + \sigma'_v$. Then, by Chen [14], (M, M') is exchangeable. Also, by Chen [14, Proposition 2], we have

$$\mathrm{E}[M - M'|\sigma] = L_1(m(\sigma)) + R_1(m(\sigma)), \tag{5.2}$$

$$\mathrm{E}[(M - M')^2|\sigma] = L_2(m(\sigma)) + R_2(m(\sigma)), \tag{5.3}$$

where $m(\sigma) = M/n$ and

$$\begin{gathered} L_1(x) = \tfrac{2(1-x)(x^2-(1-x)e^{2\tau(x)})}{(1-x)+e^{2\tau(x)}}, \text{ for } 0 < x < 1, \\ L_2(x) = \tfrac{4(1-x)(x^2+(1-x)e^{2\tau(x)})}{(1-x)+e^{2\tau(x)}}, \text{ for } 0 < x < 1, \\ |R_1(x)| + |R_2(x)| \le \tfrac{C}{n} \end{gathered}$$

for some constant C. Next, we consider two cases. In the first case, $(J, h) \notin \Gamma \cup \{(J_c, h_c)\}$, and in the second case, $(J, h) = (J_c, h_c)$.

Case 1. When $(J, h) \notin \Gamma \cup \{(J_c, h_c)\}$. Define $W = n^{-1/2}(M - nm_0)$ and $W' = n^{-1/2}(M' - nm_0)$; then, (W, W') is also an exchangeable pair. Moreover,

$$|W - W'| \le 2n^{-1/2} := \delta.$$

Note that $L_1(m_0) = 0$ by observing $m_0^2 = (1 - m_0)e^{2\tau(m_0)}$. Moreover, we have

$$L'_1(m_0) = \frac{1}{2\lambda_0} L_2(m_0) > 0,$$

where $\lambda_0 = (-1/H''(m_0)) - (1/2J) > 0$. By the Taylor expansion, we have

$$L_1(m(\sigma)) = L'_1(m_0)(m(\sigma) - m_0) + \int_{m_0}^{m(\sigma)} L''_1(s)(m(\sigma) - s)ds.$$

Let $\lambda = L_2(m_0)/(2n)$, and we have

$$n^{-1/2} L_1(m(\sigma)) = \lambda \left(\lambda_0^{-1} W + r(W)\right),$$

where

$$r(W) = 2n^{1/2} L_2^{-1}(m_0) \int_{m_0}^{m(\sigma)} L''_1(s)(m(\sigma) - s)ds.$$

Therefore, together with the definition of (W, W') and (5.2), we have

$$\mathrm{E}(W - W'|W) = n^{-1/2}(L_1(m(\sigma)) + R_1(m(\sigma))) = \lambda(g(W) + R(W)),$$

where

$$g(W) = W/\lambda_0 \quad \text{ and } \quad R(W) = r(W) + \frac{2n^{1/2}}{L_2(m_0)} R_1(m(\sigma)).$$

Thus, conditions (A1)–(A4) holds for $g(w) = w/\lambda_0$. Furthermore, $\delta|g(W)| \leq \frac{2}{\lambda_0}$, as $n^{-1/2}|W| \leq 1$.

By Chen [14, Lemma 1], there exist constants $C_0, C_1 > 0$ such that

$$|R(W)| \leq C_0 n^{-1/2}(W^2 + 1), \tag{5.4}$$

and

$$\left| \frac{1}{2\lambda} \mathrm{E}((W - W')^2|W) - 1 \right| \leq C_1 n^{-1/2}(|W| + 1)$$

and $|\hat{K}_1| = \frac{\Delta^2}{2\lambda} \leq 4/L_2(m_0)$. Therefore, (2.7)–(2.10) are satisfied with $\tau_1 = 1, \tau_2 = 2, \delta_1 = \delta_2 = O(1)n^{-1/2}$ and $d_0 = 4/L_2(m_0)$ and $d_1 = 2/\lambda_0$.

It suffices to prove (2.15) and (2.16). By (5.4), we have for $|W| \leq \frac{\sqrt{n}}{2\lambda_0 C_0}$,

$$|R(W)| \leq \frac{1}{2}(|g(W)| + 1), \tag{5.5}$$

and for $|W| > \frac{\sqrt{n}}{2\lambda_0 C_0}$, recalling that $|W| \leq 1$, we have $|R(W)| \leq C_0(\sqrt{n}+1)$. Then, (2.15) holds with $\alpha = 1/2$, $d_2 = C_0(\sqrt{n} + 1)$ and $\kappa = \sqrt{n}/(2\lambda_0 C_0)$. By Chen [14, Lemma 2], when $(J, h) \notin \Gamma \cup \{(J_c, h_c)\}$, for any $u > 0$, there exists a constant $\eta > 0$ such that

$$\mathrm{P}(|m(\sigma) - m_0| \geq u) \leq Ce^{-n\eta}$$

for some constant C. Hence,

$$d_2 \mathrm{P}(|W| > \kappa) \leq C(\sqrt{n} + 1)e^{-n\eta}.$$

Note that $s_0 = \max\{s : \delta s g^2(s) \leq 1\}$, $g(w) = w/\lambda_0$, $d_1 = \frac{2}{\lambda_0}$ and $\delta = 2n^{-1/2}$, then $s_0 = (\lambda_0/2)^{1/3} n^{1/6}$. Therefore, (2.16) is satisfied. By Remark 2.1, we have

$$\frac{\mathrm{P}(W \geq z)}{\mathrm{P}(Z_0 \geq z)} = 1 + O(1)n^{-1/2}(1 + z^3)$$

for $0 \leq z \leq n^{1/6}$.

Case 2. When $(J,h) = (J_c, h_c)$. Define $W = n^{-3/4}(M - nm_c)$ and $W' = n^{-3/4}(M' - nm_c)$; then, (W, W') is an exchangeable pair. By (5.2), we have

$$\mathrm{E}(W - W'|W) = n^{-3/4}(L_1(m(\sigma)) + R_1(m(\sigma))).$$

By Chen [14, p. 14], we have

$$L_1(m_c) = L_1'(m_c) = L_1''(m_c) = 0, \quad L_1^{(3)} = \frac{\lambda_c}{2} L_2(m_c),$$

where λ_c is given in (3.6). Then, by the Taylor expansion, we have

$$L_1(m(\sigma)) = \frac{L_1^{(3)}(m_c)}{6}(m(\sigma) - m_c)^3 + \frac{1}{6}\int_{m_c}^{m(\sigma)} L_1^{(4)}(s)(m(\sigma) - s)^3 ds.$$

Then, by taking $\lambda = L_2(m_c)/(2n^{3/2})$, by Chen [14, Lemma 1], we have

$$\mathrm{E}(W - W'|W) = \lambda(g(W) + R(W)),$$

where $g(W) = (\lambda_c/6)W^3$ and

$$R(W) = \frac{n^{3/4}}{2L_2(m_c)} \int_{m_0}^{m(\sigma)} L_1^{(4)}(s)(m(\sigma) - m_c)^3 ds + \frac{2n^{3/4}}{L_2(m_c)} R_1(W).$$

Hence, $G(w) = \frac{\lambda_c}{24} w^4$. Based again on Chen [14, Lemma 1], for some constant C, we have

$$|R(W)| \le Cn^{-1/4}(|W|^4 + 1) \le Cn^{-1/4}(|g(W)|^{4/3} + 1), \tag{5.6}$$

and

$$\left| \frac{1}{2\lambda} \mathrm{E}((W - W')^2|W) - 1 \right| \le Cn^{-1/4}(|g(W)|^{1/3} + 1).$$

As $|W - W'| \le 2n^{-3/4}$ and $|W| \le Cn^{1/4}$, $n^{-3/4}|g(W)|$ and $\hat{K}_1 = (W - W')^2/(2\lambda)$ are bounded by constants d_1 and d_0, respectively. Thus, (2.9) and (2.10) are satisfied. Furthermore, (2.7) and (2.8) hold with $\delta = 2n^{-3/4}, \delta_1 = \delta_2 = O(1)n^{-1/4}$ and $\tau_1 = 1/3, \tau_2 = 4/3$. It suffices to show that (2.15) and (2.16) are satisfied. By (5.6), there exists a constant $c > 0$ such that for $|W| \le cn^{1/4}$,

$$|R(W)| \le \frac{1}{2}(|g(W)| + 1).$$

For $|W| \geq cn^{1/4}$, noting that $|W| \leq Cn^{1/4}$, we have $|R(W)| \leq Cn^{3/4}$. Thus, (2.15) is satisfied with $\alpha = 1/2$, $d_2 = Cn^{3/4}$ and $\kappa = cn^{1/4}$. Furthermore, as $\delta = 2n^{-3/4}$ and $g(w) = (\lambda_c/6)w^3$, we have $s_0 = (18/\lambda_c)^{1/7} n^{3/28}$. In addition, by Chen [14, Lemma 2], when $(J, h) = (J_c, h_c)$, for any $u > 0$, there exists a constant $\eta > 0$ such that

$$\mathrm{P}(|m(\sigma) - m_c| \geq u) \leq Ce^{-n\eta}.$$

Thus,

$$d_2 \mathrm{P}(|W| \geq \kappa) \leq Cn^{3/4} e^{-n\eta} \leq Ce^{-2s_0 d_1^{-1} \delta^{-1}}.$$

Then, (2.16) holds. By Remark 2.1, we complete the proof of Theorem 3.2. $\square$

References.

[1] Alberici, D., Contucci, P., Fedele, M. and Mingione, E. (2016). *Limit theorems for monomer-dimer mean-field models with attractive potential. Commun. Math. Phys.*, **346**(3), pp. 781–799.

[2] Alberici, D., Contucci, P. and Mingione, E. (2014). *A mean-field monomer-dimer model with attractive interaction: Exact solution and rigorous results. J. Math. Phys.*, **55**(6), 063301.

[3] Barbour, A. (1990). *Stein's method for diffusion approximations. Probab. Theory Relat. Fields*, **84**(3), pp. 297–322.

[4] Chang, T. S. (1939). *Statistical theory of the adsorption of double molecules. Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, **169**(939), pp. 512–531.

[5] Chatterjee, S. (2008). *A new method of normal approximation. Ann. Probab.*, **36**(4), pp. 1584–1610.

[6] Chatterjee, S. (2014). *A short survey of Stein's method. Proceedings of ICM 2014.*, **4**, pp. 1–24.

[7] Chatterjee, S. and Dey, P. S. (2010). *Applications of Stein's method for concentration inequalities. Ann. Probab.*, **38**(6), pp. 2443–2485.

[8] Chatterjee, S. and Shao, Q.-M. (2011). *Nonnormal approximation by Stein's method of exchangeable pairs with application to the Curie–Weiss model. Ann. Probab.*, **21**(2), pp. 464–483.

[9] Chen, L. H. Y. and Shao, Q.-M. (2004). *Normal approximation under local dependence. Ann. Probab.*, **32**(3), pp. 1985–2028.

[10] Chen, L. H. Y. and Shao, Q.-M. (2007). *Normal approximation for nonlinear statistics using a concentration inequality approach. Bernoulli*, **13**(2), pp. 581–599.

[11] Chen, L. H. Y., Fang, X. and Shao, Q.-M. (2013). *From Stein identities to moderate deviations. Ann. Probab.*, **41**(1), pp. 262–293.

[12] Chen, L. H. Y., Goldstein, L. and Shao, Q.-M. (2011). *Normal approximation by Stein's method*. Probability and its Applications. Springer, Heidelberg.

[13] Chen, L. H. Y. and Röllin, A. (2010). *Stein couplings for normal approximation. ArXiv: 1003.6039*.

[14] Chen, W.-K. (2016). *Limit theorems in the imitative monomer-dimer mean-field model via Stein's method. J. Math. Phys.*, **57**(8), 083302.

[15] Diaconis, P. (1977). *Finite forms of de Finetti's theorem on exchangeability. Synthese*, **36**(2), pp. 271–281.

[16] Fowler, R. H. and Rushbrooke, G. S. (1937). *An attempt to extend the statistical theory of perfect solutions. Trans. Faraday Soc.*, **33**, pp. 1272–1294.

[17] Goldstein, L. and Reinert, G. (1997). *Stein's method and the zero bias transformation with application to simple random sampling. Ann. Probab.*, **7**(4), pp. 935–952.

[18] Linnik, Y. V. (1961). *On the probability of large deviations for the sums of independent variables*. In *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, **2**, pp. 289–306. University of California Press, Berkeley.

[19] Nourdin, I. and Peccati, G. (2009). *Stein's method on Wiener chaos. Probab. Theory Relat. Fields*, **145**(1-2), pp. 75–118.

[20] Nourdin, I. and Peccati, G. (2012). *Normal approximations with Malliavin calculus: from Stein's method to universality*. Cambridge University Press, Cambridge.

[21] Petrov, V. V. (1975). *Sums of independent random variables*. Springer-Verlag, New York-Heidelberg, translated from the Russian by A. A. Brown, Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 82.

[22] Roberts, J. K. (1938). *Some properties of mobile and immobile adsorbed films. Proceedings of the Cambridge Philosophical Society*, **34**, p. 399.

[23] Shao, Q.-M. and Zhang, Z.-S. (2016). *Identifying the limiting distribution by a general approach of Stein's method. Sci. China-Math.*, **59**(12), pp. 2379–2392.

[24] Shao, Q.-M. and Zhang, Z.-S. (2018). *Berry–Esseen bounds of normal and nonnormal Approximation for unbounded exchangeable pairs*. To appear in *Ann. Probab.*

[25] Simon, B. and Griffiths, R. B. (1973). *The* $(\phi^4)_2$ *field theory as a classical Ising model. Comm. Math. Phys.*, **33**(2), pp. 145–164.

[26] 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*, **2**, pp. 583–602. University of California Press, Berkeley.

[27] Stein, C. (1986). *Approximate Computation of Expectations. Institute of Mathematical Statistics Lecture Notes—Monograph Series,* **7**. Institute of Mathematical Statistics, Hayward, CA.

Department of Statistics
The Chinese University of Hong Kong
Shatin, N.T.
Hong Kong
E-mail: qmshao@cuhk.edu.hk

Department of Mathematics
Hong Kong University of Science and Technology
Clear Water Bay
Kowloon, Hong Kong
E-mail: mzhangag@connect.ust.hk

Department of Statistics
The Chinese University of Hong Kong
Shatin, N.T.
Hong Kong
E-mail: zszhang.stat@gmail.com