Noname manuscript No.
(will be inserted by the editor)

Stein's Method and Zero Bias Transformation for CDOs tranche pricing

N. El Karoui · Y. Jiao

the date of receipt and acceptance should be inserted later

Abstract We propose an original approximation method, which is based on the Stein's method and the zero bias transformation, to calculate CDO tranches in the general factor framework. We establish first-order correction terms for the Gaussian and the Poisson approximations respectively and we estimate the approximation errors. The application to the CDOs pricing consists of combining the two approximations.

1 Introduction

Stein's method, since its introduction by Stein [Ste72] in 1972, has proved to be highly efficient in dealing with approximation and estimation problems, and in particular, for sum of random variables. The method can be applied to a wide range of distributions. The main technique consists of choosing a suitable operator for the reference distribution, notably the normal distribution, the first case treated by Stein himself and the Poisson distribution, treated by Chen [Che75] in 1975. More precisely, the reference law μ is characterized by an operator $\mathcal{A}_\mu$ such that for any random variable $Z \sim \mu$,

$$E[Zf(Z)] - E[\mathcal{A}_\mu f(Z)] = 0. \tag{1.1}$$

In the normal case, $\mathcal{A}_N f(x) = \sigma^2 f'(x)$. In the Poisson case, $\mathcal{A}_P f(x) = \lambda f(x+1)$.

N. El Karoui
Centre de Mathématiques Appliquées, Ecole Polytechnique, Palaiseau 91128, France
E-mail: elkaroui@cmapx.polytechnique.fr

Y. Jiao
Laboratoire de probabilités et modèles aléatoires, Université Paris 7
E-mail: jiao@math.jussieu.fr
The work is partially supported by *Fondation de risque.*

The operator $\mathcal{A}_\mu$ is different from the usual Stein operator T_0, defined by $T_0 f = h - \int h d\mu$, which is related to the generator of a Markov process with stationary distribution. By using $\mathcal{A}_\mu$, we here choose the framework of the zero bias transformation introduced by Goldstein and Reinert [GR97]. For a given random variable X satisfying certain conditions, X^* is said to have the X-zero biased distribution of X if

$$E[Xf(X)] = E[\mathcal{A}_\mu f(X^*)]$$

for any function f such that both sides of the above equality are well defined.

The proximity between an arbitrary distribution and the reference distribution and the estimations concerned have been studied in a large literature. The main idea of Stein is to introduce the so-called *Stein's equation*, for a given function h as below:

$$h(x) - \int h d\mu = xf(x) - \mathcal{A}_\mu f(x). \tag{1.2}$$

Hence, for any random variable X, the approximation error between $E[h(X)]$ and $E[h(Z)]$ where $Z \sim \mu$ can be calculated, using the solution f_h of the Stein's equation (1.2), by

$$E[h(X)] - \int h d\mu = E[Xf_h(X)] - E[\mathcal{A}_\mu f_h(X)].$$

Combining the zero bias transformation and the Stein's equation, the approximation error can be rewritten as

$$E[h(X)] - \int h d\mu = E[\mathcal{A}_\mu f_h(X^*) - \mathcal{A}_\mu f_h(X)]. \tag{1.3}$$

To obtain efficient error estimation, it is crucial to estimate the difference between X and X^*, and some superior norm of $\mathcal{A}_\mu f_h$ and its derivatives.

In this paper, we propose an efficient approximation method, by adopting the the framework of the Stein's method and the zero bias transformation, to evaluate the credit portfolio product CDOs. For such products of large size, it's important to find rapid and robust method to calculate the tranche prices. The main term to compute is $E[(L_T - K)^+]$ where $L_T = \sum_{i=1}^n L_i \, I_{\{\tau_i \le T\}}$ is the cumulative loss on a portfolio of financial assets susceptible to default risks. Here τ_i is the default time of each name, L_i is the loss given default of τ_i and K is the attachment or the detachment point of the tranche. Under the standard convention of the market, the defaults are supposed to be conditionally independent given some random variable U, which represents the common market factor. So the conditional loss can be written as a sum of independent random variables and we are concerned with the classical approximation problem for the expectation of functions of such sum variables.

In the credit context, the default probabilities are in general small. Moreover, the conditional probability, being a function of the factor U, can take values in the interval $(0, 1)$. So the sum of independent random variables may converge to the Gaussian or the Poisson distributions. For the finance concern, direct Gaussian approximation has been applied to loan portfolios by Vasicek [Vas91]. Further Gaussian approximation by Gram-Charlier expansion has been used by Tanaka, Yamada and Watanabe [TYW05] to study interest rate derivatives. Since the Stein's method can be applied in both Gaussian and Poisson contexts, we propose to combine the two approximations. Furthermore, we propose a first order corrector in each case.

Such high-order approximations are related to the asymptotic expansions of $E[h(W)]$, which have been studied among others, by Hipp [Hip77], Götze and Hipp [GH78] using the Fourier methods and by Barbour [Bar86], [Bar87] using the Stein's method. For the CDOs, $h(x) = (x-k)^+$ is the call function in finance, which deserves special attention because of its regularity. We shall establish approximation estimations for the call function by using as main tool the concentration inequality due to Chen and Shao [CS01].

The paper is organized as follows. We first present the framework of the Stein's method and the zero bias transformation in Section 2. Some useful estimation results are given concerning the zero biased variables. In Section 3, we propose the first-order Gaussian approximation by establishing an explicit correction term and we estimate the corrected approximation error, especially for the call function. The Poisson approximation is presented in parallel in Section 4. The framework and the results are similar but the techniques are different since we are concerned with discrete random variables. We apply in Section 5 these two approximations to CDOs tranche pricing, and we propose an empirical threshold between the two approximations. Finally, some explicit estimations are given in Appendix.

2 Stein's method and zero bias transformation

2.1 Preliminaries

Stein [Ste72] has observed that a random variable (r.v.) Z follows the central Gaussian distribution $N(0,\sigma^2)$ if and only if

$$E[Zf(Z)] = \sigma^2 E[f'(Z)] \tag{2.1}$$

for any regular enough function f. In a more general context, Goldstein and Reinert [GR97] propose, for any square integrable mean zero random variable X, the zero biased distribution as follows.

Definition 2.1 (Goldstein and Reinert) Let X be a mean zero r.v. with finite variance $\sigma^2 > 0$. We say that a r.v. X^* has the X-*zero biased distribution* if

$$E[Xf(X)] = \sigma^2 E[f'(X^*)] \tag{2.2}$$

for any absolutely continuous function f such that (2.2) is well defined.

By Stein's observation, the central Gaussian distribution is characterized by the fact that Z^* and Z have the same distribution.

Hence, the distance between an arbitrary distribution and the central Gaussian one can be measured by the distance between this distribution and its zero biased one. The *Stein's equation* for the Gaussian distribution is

$$h(x) - \Phi_\sigma(h) = xf(x) - \sigma^2 f'(x) \tag{2.3}$$

where $\Phi_\sigma(h) = E[h(Z)]$ and $Z \sim N(0, \sigma^2)$. Combining (2.2) and (2.3), the error of the Gaussian approximation of $E[h(X)]$ is

$$E[h(X)] - \Phi_\sigma(h) = E[Xf_h(X) - \sigma^2 f_h'(X)] = \sigma^2 E[f_h'(X^*) - f_h'(X)] \tag{2.4}$$

where f_h is the solution of (2.3).

The Stein's equation can be solved explicitly. If $h(t)\exp(-\frac{t^2}{2\sigma^2})$ is integrable on $\mathbb{R}$, then one solution of (2.3) is given by

$$f_h(x) = \frac{1}{\sigma^2 \phi_\sigma(x)} \int_x^\infty \bar{h}(t)\, \phi_\sigma(t) dt, \tag{2.5}$$

where $\phi_\sigma(x)$ is the density function of $N(0, \sigma^2)$ and $\bar{h}(t) = h(t) - \Phi_\sigma(h)$.

Observe that f_h is one more order differentiable than h. Stein and other authors have established estimations comparing f_h and its derivatives with respect to the function h. For example (e.g. [Ste86], [CS05]), if h is an absolutely continuous function, then we have the inequality $\|f_h''\| \leq 2\|h'\|/\sigma^2$.

The equivalent expectation form of (2.5) is given by Barbour [Bar86] as below:

$$f_h(x) = \frac{\sqrt{2\pi}}{\sigma} E\big[\bar{h}(Z + x) e^{-\frac{Zx}{\sigma^2}} I_{\{Z>0\}}\big]. \tag{2.6}$$

2.2 Properties concerning the zero bias transformation

In this section, we present some useful properties and estimations concerning the zero bias transformation of one random variable. Usually, the symbol Z is used for central Gaussian variable.

Recall that X represents a mean zero r.v. with variance $\sigma^2 > 0$. The existence of the zero biased distribution is established in [GR97]. The distribution of X^* is unique and is characterized by the density function $p_{X^*}(x) = \sigma^{-2} E[X I_{\{X>x\}}]$.

In the context of zero bias transformation, the variable X is required to be mean zero. We here present a useful example.

Example 2.2 (Asymmetric Bernoulli) Let X be a mean zero asymmetric Bernoulli r.v. taking two values $\alpha = q = 1 - p$ and $\beta = -p$ $(0 < p, q < 1)$ in $(-1, 1)$, with probabilities $P(X = q) = p$ and $P(X = -p) = q$ respectively. Then the first two moments of X are $E(X) = 0$ and $\mathrm{Var}(X) = pq$. We call this distribution the asymmetric Bernoulli distribution and denote it by $\mathcal{B}(q, -p)$. Direct calculation gives that its zero biased distribution is the uniform distribution on $[-p, q]$.

More generally, any mean zero asymmetric Bernoulli r.v. can be written as a dilation of $\mathcal{B}(q,-p)$ by letting $\alpha = \gamma\, q$ and $\beta = -\gamma\, p$, which we denote by $\mathcal{B}_\gamma(q,-p)$. If X follows $\mathcal{B}_\gamma(q,-p)$, then $\mathrm{Var}(X) = \gamma^2\, pq$. In addition, its zero biased distribution is the uniform distribution on $[-\gamma p, \gamma q]$.

2.2.1 Some estimations

By definition, for any $k \in \mathbb{N}$, if X has $(k+2)^{\text{th}}$ order moment, then X^* has k^{th} order moment. Furthermore,

$$E\big[(X^*)^k\big] = \frac{E[X^{k+2}]}{\sigma^2(k+1)} \quad \text{and} \quad E\big[|X^*|^k\big] = \frac{E[|X|^{k+2}]}{\sigma^2(k+1)}. \tag{2.7}$$

We are interested in the difference $X - X^*$. The estimations are easy when X and X^* are independent by using a symmetrical term X^s.

Proposition 2.3 *Assume that X and X^* are independent. Let f be a locally integrable even function and F be its primitive function defined by $F(x) = \int_0^x f(t)dt$. Then*

$$E\big[f(X^* - X)\big] = \frac{1}{2\sigma^2} E\big[X^s F(X^s)\big] \tag{2.8}$$

where $X^s = X - \widetilde{X}$ and $\widetilde{X}$ is an independent duplicate of X. In particular,

$$E\big[|X^* - X|\big] = \frac{1}{4\sigma^2} E\big[|X^s|^3\big], \quad E\big[|X^* - X|^k\big] = \frac{1}{2(k+1)\sigma^2} E\big[|X^s|^{k+2}\big]. \tag{2.9}$$

Proof By definition, for any real number K, we have $\sigma^2 E[f(X^* - K)] = E[XF(X-K)]$. Since X^* is independent of X, let $\widetilde{X}$ be a r.v. having the same distribution and independent of X, then

$$E[f(X^* - X)] = \frac{1}{\sigma^2} E[\widetilde{X} F(\widetilde{X} - X)].$$

As f is an even function, F is an odd function, then

$$E[\widetilde{X} F(\widetilde{X} - X)] = E[XF(X - \widetilde{X})] = -E[XF(\widetilde{X} - X)],$$

which implies (2.8). To obtain (2.9), it suffices to let $f(x) = |x|$ and $f(x) = |x|^k$.

Proposition 2.3 provides an equality to estimate $|X^* - X|$. Similar calculation yields estimation for $P(|X^* - X| \leq \epsilon)$, giving a measure of spread between X and X^*.

Corollary 2.4 *Let X and X^* be independent. Then, for any $\varepsilon > 0$,*

$$P(|X - X^*| \leq \varepsilon) \leq \frac{\varepsilon}{\sqrt{2}\,\sigma} \wedge 1, \quad P(|X - X^*| \geq \varepsilon) \leq \frac{1}{4\sigma^2\,\varepsilon} E\big[|X^s|^3\big] \tag{2.10}$$

Proof Let us observe that the second inequality is immediate from the classical Markov inequality

$$P(|X - X^*| \geq \varepsilon) \leq \frac{1}{\varepsilon} E\big[|X - X^*|\big].$$

To obtain the first inequality, we apply Proposition 2.3 to the even function $g(x) = I_{\{|x| \leq \varepsilon\}}$ and whose primitive function is $G(x) = \mathrm{sign}(x)\,\big(|x| \wedge \varepsilon\big)$. So,

$$P(|X - X^*| \leq \varepsilon) = \frac{1}{2\sigma^2} E\big[|X^s|\big(|X^s| \wedge \varepsilon\big)\big].$$

Since $|X^s| \wedge \varepsilon \leq \varepsilon$ and $E\big[|X^s|\big]^2 \leq E\big[|X^s|^2\big] = 2\sigma^2$, we get

$$P(|X - X^*| \leq \varepsilon) \leq \frac{\varepsilon}{2\sigma^2}(2\sigma^2)^{1/2} = \frac{\varepsilon}{\sqrt{2}\,\sigma}$$

The first inequality of Corollary 2.4 makes sense when ε is small, otherwise, the probability is always bounded by 1.

2.2.2 Estimation bounds for Gaussian approximation

The Stein's method has been applied to a large class of approximation problems. For the Gaussian approximation, one can find a good survey in Chen and Shao [CS05] and Raic [Rai03]. The approximation error estimations are in general based on comparison of expectations, for example, under Wasserstein distance for uniformly Lipschitz functions and under Kolmogorov distance for indicator functions.

In the context of zero bias transformation, we have, by using (2.9), a direct zero-order estimation result of Stein.

Proposition 2.5 *(Stein) If h is an absolutely continuous function, then*

$$\big|E[h(X)] - \Phi_\sigma(h)\big| \leq \frac{\|h'\|}{2\sigma^2}\, E\big[|X^s|^3\big]. \tag{2.11}$$

Proof In fact, since

$$\big|E[h(X)] - \Phi_\sigma(h)\big| = \sigma^2 E\big[|f_h'(X^*) - f_h'(X)|\big] \leq \sigma^2 \big\|f_h''\big\| E\big[|X^* - X|\big],$$

then (2.11) follows directly with previous estimations.

The upper bound of (2.11) depends on the estimation of $E[|X - X^*|]$, for which we have supposed that X^* is independent of X in Proposition 2.3.

In the general case where X^* and X are not independent, the lower bound of this term is given (see [Gol07]) by

$$\inf E[|X^* - X|] = \|F - F^*\|_1 := \int_{-\infty}^{\infty} |F(t) - F^*(t)| dt \tag{2.12}$$

where F and F^* are the distribution functions of X and X^* respectively. The equality is satisfied when $X = F^{-1}(U)$ and $X^* = (F^*)^{-1}(U)$ where

$U \sim U(0,1)$ is some uniform random variable, which means that the variables X and X^* are rather strongly correlated.

Goldstein [Gol07] has established the L^1 bound for the Gaussian approximation by using this dual form of L^1 distance between F and F^*.

Example 2.6 Let $X \sim \mathcal{B}(q,-p)$ and X^* be a r.v. having the X-zero biased distribution. If X^* is independent of X, then by direct calculation, $E[|X - X^*|] = \frac{1}{2}$.

We can calculate the lower bound of this expectation by using (2.12). In fact, we have $F(x) = qI_{\{-p \le x < q\}} + I_{\{x > q\}}$ and $F^*(x) = x + p$. Then $\inf E[|X - X^*|] = \frac{1}{2}(1-2pq)$.

2.2.3 Sum of independent random variables

A typical example which concerns the sum of independent random variables deserves special attention. The problem is relatively simple when we restrict to the most classical version where all variables are identically distributed. However, this elementary case can be extended to non-identically distributed variables.

Goldstein and Reinert [GR97] give an interesting construction by using an arbitrary index and replacing each single summand by its independent zero biased r.v.. Such construction is informative since W and W^* are only partially independent.

Proposition 2.7 *(Goldstein and Reinert) Let $X_1, \cdots, X_n$ be independent mean zero random variables with finite variance $\sigma_i^2 > 0$ and let X_i^* have the X_i-zero biased distribution. We assume that $(\bar{X}, \bar{X}^*) = (X_1, \cdots, X_n, X_1^*, \cdots, X_n^*)$ are independent r.v.. Let $W = X_1 + \cdots + X_n$ and denote by σ_W^2 its variance. We also use the notation $W^{(i)} = W - X_i$. Let us introduce a random choice I of the index i such that $P(I = i) = \sigma_i^2/\sigma_W^2$, and assume I independent of $(\bar{X}, \bar{X}^*)$. Then $W^* = W^{(I)} + X_I^*$ has the W-zero biased distribution.*

Proof We prove by verification. Let f be a continuous function with compact support and F be a primitive function of f. Since X_i is independent of $W^{(i)}$, then

$$E[WF(W)] = \sum_{i=1}^n E[X_i F(W)] = \sum_{i=1}^n E[X_i F(W^{(i)} + X_i)] = \sum_{i=1}^n \sigma_i^2 E[f(W^{(i)} + X_i^*)]. \tag{2.13}$$

Since I is independent of W, the last term of the right-hand side of (2.13) equals in fact $\sigma_W^2 E[f(W^{(I)} + X_I^*)]$, which implies that $W^* = W^{(I)} + X_I^*$ has the W-zero biased distribution.

If W is the sum of i.i.d. mean zero asymmetric Bernoulli r.v. $\mathcal{B}_\gamma(q,-p)$, then W follows the asymmetric binomial distribution. In addition, the dilatation parameter is $\gamma = \frac{\sigma_W}{\sqrt{npq}}$.

We now extend the estimation results in the previous subsection to the sum variables.

Corollary 2.8 *With the notation of Proposition 2.7, we have*

$$E[X_I^*] = \frac{1}{2\sigma_W^2}\sum_{i=1}^n E[X_i^3], \quad E[(X_I^*)^2] = \frac{1}{3\sigma_W^2}\sum_{i=1}^n E[X_i^4]$$

and

$$E\big[|W^* - W|\big] = \frac{1}{4\sigma_W^2}\sum_{i=1}^n E[|X_i^s|^3], \quad E\big[|W^* - W|^k\big] = \frac{1}{2(k+1)\sigma_W^2}\sum_{i=1}^n E\big[|X_i^s|^{k+2}\big]. \tag{2.14}$$

In particular, when W follows the asymmetric binomial distribution, we have

$$E\big[|W^* - W|^k\big] = \frac{1}{k+1}\Big(\frac{\sigma_W}{\sqrt{np(1-p)}}\Big)^k.$$

Proof In fact, the above results are direct consequences of (2.7) and (2.9), together with Proposition 2.7.

Corollary 2.9 *For any positive constant ε, we have*

$$P(|W^* - W| \le \varepsilon) \le \Big(\frac{\varepsilon}{\sqrt{2}\sigma_W^2}\sum_{i=1}^n \sigma_i\Big)\wedge 1, \quad P(|W^* - W| \ge \varepsilon) \le \frac{1}{4\sigma_W^2\varepsilon}\sum_{i=1}^n E[|X_i^s|^3]. \tag{2.15}$$

We can calculate the conditional expectations of X_I and X_I^* given $(\bar{X}, \bar{X}^*)$, which enables us to obtain some useful estimations. For example,

$$E[X_I|X_1, \cdots, X_n] = \sum_{i=1}^n \frac{\sigma_i^2}{\sigma_W^2}X_i, \quad E[E[X_I|X_1, \cdots, X_n]^2] = \sum_{i=1}^n \frac{\sigma_i^6}{\sigma_W^4}.$$

Notice that, in the homogeneous asymmetric Bernoulli case, $E[E[X_I|X_1, \cdots, X_n]^2]$ is of order $O\big(\frac{1}{n^2}\big)$, while $E[X_I^2]$ is only of order $O\big(\frac{1}{n}\big)$. This observation shows that the use of the random index I in the construction of W^* is efficient.

By applying the conditional expectation technique, we obtain the following result which is still valid when we replace W by W^*.

Proposition 2.10 *Let $f : R \to R$ and $g : R^2 \to R$ be two functions such that the variance of $f(W)$ exists, and that for any $i = 1, \cdots, n$, the variance of $g(X_i, X_i^*)$ exists, then*

$$\big|\mathrm{Cov}\big[f(W), g(X_I, X_I^*)\big]\big| \le \frac{\mathrm{Var}[f(W)]^{\frac{1}{2}}}{\sigma_W^2}\Big(\sum_{i=1}^n \sigma_i^4 \mathrm{Var}[g(X_i, X_i^*)]\Big)^{\frac{1}{2}}. \tag{2.16}$$

In particular, for any $\varepsilon \ge 0$,

$$\big|\mathrm{Cov}\big[I_{\{a\le W\le b\}}, I_{\{|X_I - X_I^*|\le\varepsilon\}}\big]\big| \le \frac{1}{4}\Big(\sum_{i=1}^n \frac{\sigma_i^4}{\sigma_W^4}\Big)^{\frac{1}{2}}. \tag{2.17}$$

Proof We first notice that $\mathrm{Cov}[f(W), g(X_I, X_I^*)] = \mathrm{Cov}\big[f(W), E[g(X_I, X_I^*)|\bar{X}, \bar{X}^*]\big]$. Since (X_i, X_i^*) are mutually independent, we have

$$\begin{aligned}\mathrm{Cov}\big[f(W), E[g(X_I, X_I^*)|\bar{X}, \bar{X}^*]\big] &\leq \mathrm{Var}[f(W)]^{\frac{1}{2}}\mathrm{Var}\big[E[g(X_I, X_I^*)|\bar{X}, \bar{X}^*]\big]^{\frac{1}{2}} \\ &\leq \frac{1}{\sigma_W^2}\mathrm{Var}[f(W)]^{\frac{1}{2}}\Big(\sum_{i=1}^{n}\sigma_i^4\mathrm{Var}[g(X_i, X_i^*)]\Big)^{\frac{1}{2}}.\end{aligned}$$

At last, it remains to observe that $\mathrm{Var}[I_{\{a\leq W\leq b\}}] \leq \frac{1}{4}$ and $\mathrm{Var}[I_{\{|X_i - X_i^*|\leq\varepsilon\}}] \leq \frac{1}{4}$.

Remark 2.11 1. The inequality (2.17) gives the order of $\mathrm{Cov}\big[I_{\{a\leq W\leq b\}}, I_{\{|X_I - X_I^*|\leq\varepsilon\}}\big]$, which is essential for estimations in Theorem 3.1 and Proposition 3.3.
2. To prove (2.16), X_i^* is required to be independent of $(X_1, \cdots, X_{i-1}, X_{i+1}, \cdots, X_n)$. However, X_i^* and X_i are not necessarily independent.

3 First-order Gaussian approximation

Classically, the expectation $E[h(W)]$ where W is the sum of independent random variables can be approximated by the Gaussian approximation $\Phi_{\sigma_W}(h)$. The error of this direct approximation $\epsilon_0(h, W) = E[h(W)] - \Phi_{\sigma_W}(h)$ is of order $O\big(\frac{1}{\sqrt{n}}\big)$ in the binomial-normal approximation, except in the symmetric case where Diener and Diener [DD04] have established the convergence order $O\big(\frac{1}{n}\big)$.

In this section, we shall improve this approximation by proposing a correction term C_h such that the corrected error $\epsilon_1(h, W) = E[h(W)] - \Phi_{\sigma_W}(h) - C_h$ is of order $O(\frac{1}{n})$ even in the asymmetric case. Some regularity condition is required to establish approximation error estimations. Notably, the call function, not possessing second order derivative, is difficult to analyze. We first present the general results for regular functions and then treat the call function.

3.1 First order correction for normal approximation

Theorem 3.1 *Let $X_1, \cdots, X_n$ be independent mean zero random variables such that $E[X_i^4]$ $(i = 1, \cdots, n)$ exist. If the function h is Lipschitz and if f_h has bounded third order derivative, then the normal approximation $\Phi_{\sigma_W}(h)$ of $E[h(W)]$ has corrector*

$$C_h = \frac{1}{\sigma_W^2}E[X_I^*]\Phi_{\sigma_W}\left(\Big(\frac{x^2}{3\sigma_W^2} - 1\Big)xh(x)\right). \tag{3.1}$$

Recall that $E[X_I^*] = \frac{1}{2\sigma_W^2}\sum_{i=1}^n E[X_i^3]$. *The corrected approximation error is bounded by*

$$\Big|E[h(W)] - \Phi_{\sigma_W}(h) - C_h\Big|$$
$$\leq \|f_h^{(3)}\|\left(\frac{1}{12}\sum_{i=1}^n E\big[|X_i^s|^4\big] + \frac{1}{4\sigma_W^2}\Big|\sum_{i=1}^n E[X_i^3]\Big|\sum_{i=1}^n E\big[|X_i^s|^3\big] + \frac{1}{\sigma_W}\sqrt{\sum_{i=1}^n \sigma_i^6}\right).$$

Proof We apply (2.4) to W and take the first order Taylor expansion at W to obtain

$$E[h(W)] - \Phi_{\sigma_W}(h) = \sigma_W^2 E[f_h'(W^*) - f_h'(W)]$$
$$= \sigma_W^2 E[f_h''(W)(W^*-W)] + \sigma_W^2 E\Big[f_h^{(3)}\big(\xi W + (1-\xi)W^*\big)\xi(W^*-W)^2.\Big] \tag{3.2}$$

where ξ is a uniform variable on $[0,1]$ independent of all X_i and X_i^*. Since $f_h^{(3)}$ is bounded, the remaining term of (3.2) is of order $E[(W-W^*)^2]$ and we have by Corollary 2.8

$$\sigma_W^2\Big|E\Big[f_h^{(3)}\big(\xi W + (1-\xi)W^*\big)\xi(W^*-W)^2\Big]\Big| \leq \frac{\|f_h^{(3)}\|}{12}\sum_{i=1}^n E\big[|X_i^s|^4\big]. \tag{3.3}$$

For the term $E[f_h''(W)(W^*-W)]$ in (3.2), we make the following decomposition where we approximate $f_h''(W)$ by $\Phi_{\sigma_W}(f_h'')$. Notice in addition that X_I^* is independent of W, so

$$E[f_h''(W)(W^*-W)] = E[f_h''(W)]E[X_I^* - X_I] + \mathrm{Cov}[f_h''(W), X_I^* - X_I]$$
$$= \Phi_{\sigma_W}(f_h'')E[X_I^*] + \big(E[f_h''(W)] - \Phi_{\sigma_W}(f_h'')\big)E[X_I^*] - \mathrm{Cov}[f_h''(W), X_I].$$

By Proposition 2.5,

$$\Big|E[X_I^*]\Big(E[f_h''(W)] - \Phi_{\sigma_W}(f_h'')\Big)\Big| \leq \frac{\|f_h^{(3)}\|}{4\sigma_W^4}\Big|\sum_{i=1}^n E[X_i^3]\Big|\sum_{i=1}^n E\big[|X_i^s|^3\big]. \tag{3.4}$$

By Proposition 2.10,

$$\big|\mathrm{Cov}[f_h''(W), X_I]\big| \leq \frac{1}{\sigma_W^2}\sqrt{\mathrm{Var}[f_h''(W)]}\sqrt{\sum_{i=1}^n \sigma_i^6} \leq \frac{\|f_h^{(3)}\|}{\sigma_W}\sqrt{\sum_{i=1}^n \sigma_i^6}, \tag{3.5}$$

where the last inequality is because $\mathrm{Var}[f_h''(W)] = \mathrm{Var}[f_h''(W) - f_h''(0)] \leq E[(f_h''(W) - f_h''(0))^2] \leq \|f_h^{(3)}\|^2\sigma_W^2$. Combining (3.3), (3.4) and (3.5), we deduce the error bound.

Finally, we use (2.1) and the Stein's equation to obtain

$$\sigma_W^2\Phi_{\sigma_W}(f_h'') = \Phi_{\sigma_W}(xf_h') = \frac{1}{\sigma_W^2}\Phi_{\sigma_W}\Big(\big(\frac{x^2}{3\sigma_W^2} - 1\big)xh(x)\Big).$$

The corrector C_h contains two parts. On one hand, $E[X_I^*]$ depends on the third moments of $X_1, \cdots, X_n$. On the other hand, the normal expectation term depends on the function h. Both terms are easy to calculate, even in the non-homogeneous case.

In the symmetric case, $E[X_I^*] = 0$, then $C_h = 0$ for any function h. Therefore, the corrector C_h can be viewed as an asymmetric corrector in the sense that after the correction, the asymmetric approximation has the same approximation order as in the symmetric case.

When X_i's are i.i.d. asymmetric Bernoulli random variables, the corrector is of order $O(\frac{1}{\sqrt{n}})$ and the corrected approximation error bound is of order $O(\frac{1}{n})$. If, in addition, $E[X_i^3] = 0$ for any $i = 1, \cdots, n$, then the error of the approximation without correction is automatically of order $O(\frac{1}{n})$. This result has been mentioned by Feller [Fel71] concerning the Edgeworth expansion and has been discussed in [GR97].

3.2 Call function

We now concentrate on the call function $C_k(x) = (x-k)^+$, which is a Lipschitz function with $C_k'(x) = I_{\{x>k\}}$. Notice that C_k'' exists only in distribution sense. So the condition in Theorem 3.1 is not satisfied and we can no longer bound the error via the norm $\|f_h^{(3)}\|$.

In the following, we shall prove that the corrector given by (3.1) remains valid for the call function. We first study the zero order estimation for the indicator function. Then by a similar method, we deduce the first order estimation for the call function. Here it is important to estimate the norm of the solution of Stein's equation f_h and its derivatives. We shall postpone the explicit calculations to the Appendix.

3.2.1 Zero order estimation for the indicator function

In Proposition 2.5, the zero order approximation has been applied to an absolutely continuous function. For the indicator function, there also exists estimation known as the Berry-Esseen inequality.

We now introduce an estimation method based on the zero bias transformation. The key tool is a concentration inequality (see [CS01],[CS05]), which is also essential for later estimations for the call function.

We begin by introducing some estimations for the zero biased variable W^*. In fact, the zero bias transformation enables us to work with more regular functions.

Lemma 3.2 *Let a and b be two real numbers such that $a \le b$. Then*

1. *$P(a \le W^* \le b) \le \frac{b-a}{2\sigma_W}$;*
2. *Denote by $I_b(x) = I_{\{x\le b\}}$. Then*

$$\big|P(W^* \le b) - N_{\sigma_W}(b)\big| \le c_b\, E\big[|W^* - W|\big]$$

where N_σ is the distribution function of $N(0,\sigma^2)$ and c_b is a constant defined by $c_b = \|f_{I_b}\| + \|xf'_{I_b}\|$.

Proof 1) Let $f'(x) = I_{[a,b]}(x)$ and $f(x)$ be the primitive function given by $f(x) = \int_{(a+b)/2}^{x} f'(t)dt$. Then $|f(x)| \leq \frac{b-a}{2}$. By definition (2.1), we have

$$\sigma_W^2 E[I_{[a,b]}(W^*)] = E[Wf(W)] \leq \sigma_W\big(\frac{b-a}{2}\big).$$

2) Consider the primitive function of I_b given by $G_I(x) = -(b-x)^+$. We denote by $\widetilde{G}_I(x) = xG_I(x)$, then

$$\sigma_W^2\big(E[I_b(W^*)] - \Phi_{\sigma_W}(I_b)\big) = E[\widetilde{G}_I(W)] - \Phi_{\sigma_W}(\widetilde{G}_I) = \sigma_W^2 E[f'_{\widetilde{G}_I}(W^*) - f'_{\widetilde{G}_I}(W)],$$

which implies that

$$\big|P(W^* \leq b) - N_{\sigma_W}(b)\big| \leq \|f''_{\widetilde{G}_I}\|\, E\big[|W^* - W|\big].$$

Notice that $f'_{\widetilde{G}_I} = xf_{I_b}$. Hence $\|f''_{\widetilde{G}_I}\| \leq \|f_{I_b}\| + \|xf'_{I_b}\| = c_b$. We shall give the estimation of c_b in Appendix.

The concentration inequality (see e.g. [CS01], [CS05]) shows that $P(a \leq W \leq b)$ can be bounded from above by a term which is linear to $b-a$ and some other terms which, in the i.i.d. asymmetric Bernoulli case, are of order $O(\frac{1}{\sqrt{n}})$. We shall give a proof of this inequality, which is coherent in our context. The idea is to bound $P(a \leq W \leq b)$ by $P(a - \varepsilon \leq W^* \leq b + \varepsilon)$ up to a small error with a suitable ε. Our objective here is not to find the optimal estimation constant.

Proposition 3.3 *For all real numbers a and b such that $a \leq b$, we have*

$$P(a \leq W \leq b) \leq \frac{b-a}{\sigma_W} + \frac{\sum_{i=1}^n E\big[|X_i^s|^3\big]}{\sigma_W^3} + \frac{\Big(\sum_{i=1}^n \sigma_i^4\Big)^{\frac{1}{2}}}{2\sigma_W^2}. \qquad (3.6)$$

Proof By 1) of Lemma 3.2, for any $\varepsilon > 0$, W^* satisfies the following concentration inequality

$$P(a - \varepsilon \leq W^* \leq b + \varepsilon) \leq \frac{1}{\sigma_W}\big(\varepsilon + \frac{b-a}{2}\big) := C_\varepsilon.$$

On the other hand,

$$\begin{aligned} P(a - \varepsilon \leq W^* \leq b + \varepsilon) &\geq P(a \leq W \leq b, |X_I - X_I^*| \leq \varepsilon) \\ &= P(a \leq W \leq b)P(|X_I^* - X_I| \leq \varepsilon) + \mathrm{Cov}(I_{\{a \leq W \leq b\}}, I_{\{|X_I^* - X_I| \leq \varepsilon\}}). \end{aligned}$$

By Proposition 2.10,

$$\mathrm{Cov}(I_{\{a \leq W \leq b\}}, I_{\{|X_I^* - X_I| \leq \varepsilon\}}) \geq -\frac{1}{4}\big(\sum_{i=1}^n \frac{\sigma_i^4}{\sigma_W^4}\big)^{\frac{1}{2}} := -B.$$

Observe that B does not depend on a, b and ε. By the Markov inequality,

$$A_\varepsilon := P(|X_I^* - X_I| \leq \varepsilon) \geq 1 - \frac{1}{\varepsilon} E\big[|W - W^*|\big].$$

So $P(a \leq W \leq b)A_\varepsilon \leq B + C_\varepsilon$. Finally, we choose ε such that $\varepsilon = 2E\big[|W - W^*|\big]$. Then $A_\varepsilon \geq \frac{1}{2}$ and

$$C_\varepsilon = \frac{b-a}{2\sigma_W} + \frac{2E\big[|W - W^*|\big]}{\sigma_W},$$

which implies (3.6).

In the following, we use sometimes the upper bound of $P(a \leq W^{(i)} \leq b)$. Since $W^{(i)}$ is also a sum of several random variables, (3.6) can be applied directly to $W^{(i)}$ by removing the variate i in the sum terms of the upper bound. However, for the simplicity of writing, it is convenient to keep all the summand terms. For this reason, we present the following inequality, which is a consequence of (3.6) by using the independence between $W^{(i)}$ and (X_i, X_i^*). Here again, our objective is not to obtain the optimal estimation.

Corollary 3.4 *For all real numbers a and b such that $a \leq b$, we have*

$$P(a \leq W^{(i)} \leq b) \leq 2\Big(\frac{b-a}{\sigma_W} + \frac{\big(\sum_{i=1}^n \sigma_i^4\big)^{\frac{1}{2}}}{2\sigma_W^2} + \frac{\sum_{i=1}^n E\big[|X_i^s|^3\big]}{\sigma_W^3}\Big) + \frac{8\sigma_i}{\sigma_W}.$$

Proof For any $\varepsilon > 0$, $P(a \leq W^{(i)} \leq b, |X_i| \leq \varepsilon) \leq P(a - \varepsilon \leq W \leq b + \varepsilon)$. Since $W^{(i)}$ and X_i are independent, we have by the Markov's inequality that

$$P(a \leq W^{(i)} \leq b) \leq \frac{P(a - \varepsilon \leq W \leq b + \varepsilon)}{1 - \frac{E[|X_i|]}{\varepsilon}}.$$

We choose $\varepsilon = 2E[|X_i|]$ and apply Proposition 3.3 to end the proof.

We now give the estimation for the indicator function. We shall use the "nearness" between W and W^*, together with the concentration inequality.

Proposition 3.5 *Let $I_b(x) = I_{\{x \leq b\}}$, then*

$$\begin{aligned} |P(W \leq b) - N_{\sigma_W}(b)| \leq \frac{c_b}{4\sigma_W^2} \sum_{i=1}^n E[|X_i^s|^3] + \sum_{i=1}^n \frac{2\sigma_i^3}{\sigma_W^3}\Big(\frac{E[|X_i^s|^3]}{4\sigma_i^3} + 4\Big) \\ + \frac{\big(\sum_{i=1}^n \sigma_i^4\big)^{\frac{1}{2}}}{\sigma_W^2} + \frac{2\sum_{i=1}^n E\big[|X_i^s|^3\big]}{\sigma_W^3} \end{aligned} \tag{3.7}$$

where $c_b = \|f_{I_b}\| + \|x f_{I_b}'\|$.

Proof We decompose $I_b(W) - N_{\sigma_W}(b)$ as the sum of two terms

$$I_b(W) - N_{\sigma_W}(b) = \big(I_b(W) - I_b(W^*)\big) + \big(I_b(W^*) - N_{\sigma_W}(b)\big).$$

The second term on the right-hand side has been estimated in Lemma 3.2. For the first term, since

$$|I_{\{x+y\leq b\}} - I_{\{x+z\leq b\}}| = I_{\{b-\max(y,z)<x\leq b-\min(y,z)\}},$$

then $E[|I_b(W^{(i)} + X_i) - I_b(W^{(i)} + X_i^*)|] = P(b - \max(X_i, X_i^*) < W^{(i)} \leq b - \min(X_i, X_i^*))$. Since $W^{(i)}$ and (X_i, X_i^*) are independent, using Corollary 3.4, we obtain (3.7).

3.2.2 Call function

We now estimate the first order approximation error for the call function. As mentioned previously, the second order derivative of C_k and thus $f_{C_k}^{(3)}$ do not exist. So we shall rewrite f''_{C_k} using more regular functions by Stein's equation.

Proposition 3.6 *Let $X_1, \cdots, X_n$ be independent mean zero random variables such that $E[X_i^4]$ $(i = 1, \cdots, n)$ exist. Then*

$$\begin{aligned}
&\Big|E[(W-k)^+] - \Phi_{\sigma_W}((x-k)^+) - \frac{1}{3}E[X_I^*]k\phi_{\sigma_W}(k)\Big| \\
&\leq \frac{1}{2\sigma_W^2}\Big|\sum_{i=1}^n E[X_i^3]\Big|\Big(\frac{c_k}{2\sigma_W^2}\sum_{i=1}^n E\big[|X_i^s|^3\big] + B(W,k)\Big) + \mathrm{Var}[f''_{C_k}(W)]^{\frac{1}{2}}(\sum_{i=1}^n \sigma_i^6)^{\frac{1}{2}} \\
&+ \frac{c_k}{6\sigma_W^2}\sum_{i=1}^n E[|X_i^s|^4] + \frac{1}{\sigma_W^3}\sum_{i=1}^n\Big(\frac{E[|X_i^s|^4]}{3} + 2\sigma_i E[|X_i^s|^3]\Big) \\
&+ \frac{1}{4\sigma_W^2}\sum_{i=1}^n E[|X_i^s|^3]\Big(\frac{2\sum_{i=1}^n E\big[|X_i^s|^3\big]}{\sigma_W^3} + \frac{(\sum_{i=1}^n \sigma_i^4)^{\frac{1}{2}}}{\sigma_W^2}\Big).
\end{aligned} \tag{3.8}$$

where $c_k = 2\|f'_{C_k}\| + \|xf''_{C_k}\|$ and $B(W,k)$ is the upper bound of the zero-order normal approximation for the indicator function $I_{\{W\leq k\}}$.

Proof Similar as in the proof of Theorem 3.1, we decompose the corrected approximation error as the sum of three terms. Here we replace the third order derivative term by (3.11), i.e.

$$\begin{aligned}
&E[C_k(W)] - \Phi_{\sigma_W}(C_k) - \sigma_W^2 E[X_I^*]\Phi_{\sigma_W}(f''_{C_k}) \\
&= \sigma_W^2 E[X_I^*]\big(E[f''_{C_k}(W)] - \Phi_{\sigma_W}(f''_{C_k})\big) && (3.9)\\
&- \sigma_W^2 \mathrm{Cov}[f''_{C_k}(W), X_I] && (3.10)\\
&+ \sigma_W^2 E\big[f'_{C_k}(W^*) - f'_{C_k}(W) - f''_{C_k}(W)(X_I^* - X_I)\big]. && (3.11)
\end{aligned}$$

We shall estimate each term respectively.

We begin by (3.9). Since f_h'' is not differentiable, we use the Stein's equation $f_h'(x) = \big(xf_h(x) - h(x) + \Phi_{\sigma_W}(h)\big)/\sigma_W^2$ to obtain $f_{C_k}'' = (f_{C_k}(x) + xf_{C_k}'(x) - C_k'(x))/\sigma_W^2$. Let $g = f_{C_k} + xf_{C_k}'$, then

$$\begin{aligned}\sigma_W^2 \left|E[f_{C_k}''(W)] - \Phi_{\sigma_W}(f_{C_k}'')\right| &\le |E[g(W)] - \Phi_{\sigma_W}(g)| + |E[I_{\{W\le k\}}] - N_{\sigma_W}(k)| \\ &\le \frac{\|g'\|}{2\sigma_W^2}\sum_{i=1}^n E[|X_i^s|^3] + B(W,k)\end{aligned}$$

where $B(W,k)$ is the error bound in (3.7). Note that $\|g'\| \le 2\|f_{C_k}'\| + \|xf_{C_k}''\| := c_k$. We shall give the estimation of c_k in Corollary A.7 in Appendix.

By Proposition 2.10, (3.10) is estimated by

$$\sigma_W^2 \left|\mathrm{Cov}[f_{C_k}''(W), X_I]\right| \le \mathrm{Var}[f_{C_k}''(W)]^{\frac{1}{2}} \Big(\sum_{i=1}^n \sigma_i^6\Big)^{\frac{1}{2}}.$$

Since f_{C_k}'' is bounded, so is $\mathrm{Var}[f_{C_k}''(W)]$.

For (3.11), we use again the Stein's equation to rewrite f_{C_k}'. Denote by $G(x) = xf_{C_k}(x)$ and notice that $G' = g$, then

$$\begin{aligned}&\sigma_W^2 E[|f_{C_k}'(W^*) - f_{C_k}'(W) - f_{C_k}''(W)(X_I^* - X_I)|] \\ &\le E[|G(W^*) - G(W) - g(W)(X_I^* - X_I)|] + E[|C_k(W^*) - C_k(W) - C_k'(W)(X_I^* - X_I)|].\end{aligned} \tag{3.12}$$

The first term of (3.12) is bounded by $\frac{1}{2}c_k E[(X_I^* - X_I)^2]$. For the second term of (3.12), notice that the call function satisfies

$$|C_k(x+a) - C_k(x+b) - (a-b)C_k'(x+b)| \le I_{\{k-\max(a,b)\le x\le k-\min(a,b)\}}\,|a-b|.$$

Hence, we have by using the concentration inequality

$$\begin{aligned}&E\big[|C_k(W^*) - C_k(W) - C_k'(W)(X_I^* - X_I)|\big] \\ &\le E\left[|X_I^* - X_I| I_{\{k-\max(X_I^*,X_I)\le W^{(I)}\le k-\min(X_I^*,X_I)\}}\right] \\ &= \sum_{i=1}^n \frac{\sigma_i^2}{\sigma_W^2} E\left[|X_i^* - X_i|\, E[I_{\{k-\max(X_i^*,X_i)\le W^{(i)}\le k-\min(X_i^*,X_i)\}}|\bar{X}, \bar{X}^*]\right] \\ &\le \frac{1}{\sigma_W^3}\sum_{i=1}^n \Big(\frac{E[|X_i^s|^4]}{3} + 2\sigma_i E[|X_i^s|^3]\Big) + \frac{1}{4\sigma_W^2}\sum_{i=1}^n E[|X_i^s|^3]\Big(\frac{2\sum_{i=1}^n E\big[|X_i^s|^3\big]}{\sigma_W^3} + \frac{(\sum_{i=1}^n \sigma_i^4)^{\frac{1}{2}}}{\sigma_W^2}\Big).\end{aligned}$$

Finally, it remains to combine all the above estimations to obtain (3.8).

In the case where X_i's are i.i.d asymmetric Bernoulli variables, the error bound is of order $O(\frac{1}{n})$. When the strike $k = 0$, the correction term $\frac{1}{3}E[X_I^*]k\phi_{\sigma_W}(k)$ vanishes, which means that the error bound for the normal approximation of $h(x) = x^+$ is automatically of order $O(\frac{1}{n})$. In addition, the corrector attains its extremal values when $k = \pm\sigma$.

The symmetric case when $p = \frac{1}{2}$ has been studied for the binomial tree model. Diener and Diener [DD04] have proved the convergence order $O(\frac{1}{n})$ towards the Black-Scholes price. Proposition 3.6 provides another proof since in the symmetric case, $E[X_I^*]$=0.

4 First order Poisson approximation

We present the Poisson approximation in a similar framework. A Poisson distributed random variable takes non-negative integer values, so we are concerned with discrete variables, where the regularity of the call function is easier to treat.

4.1 Stein's method and zero bias transformation

Chen [Che75] has observed that a non-negative integer valued random variable Λ with expectation λ follows the Poisson distribution if and only if $E[\Lambda g(\Lambda)] = \lambda E[g(\Lambda+1)]$ for any bounded function g. Let Y be a r.v. taking non-negative integer values such that $E[Y] = \lambda < \infty$. A r.v. Y^* is said to have the *Y-Poisson zero biased distribution* if

$$E[Yg(Y)] = \lambda E[g(Y^*+1)] \tag{4.1}$$

for any function g such that $E[Yg(Y)]$ exists. If Y follows the $0-1$ Bernoulli distribution, then the distribution of Y^* is the Dirac measure δ_0.

The Stein's equation in the Poisson setting is also introduced in [Che75]:

$$h(y) - \mathcal{P}_\lambda(h) = yg(y) - \lambda g(y+1) \tag{4.2}$$

where $\mathcal{P}_\lambda(h) = E[h(\Lambda)]$ and $\Lambda \sim P(\lambda)$. Then, the error of the Poisson approximation can be written as

$$E[h(Y)] - \mathcal{P}_\lambda(h) = E\big[Yg_h(Y) - \lambda g_h(Y+1)\big] = \lambda E\big[g_h(Y^*+1) - g_h(Y+1)\big] \tag{4.3}$$

where g_h is the solution of (4.2) and is given by

$$g_h(k) = \frac{(k-1)!}{\lambda^k}\sum_{i=k}^{\infty}\frac{\lambda^i}{i!}\big(h(i) - \mathcal{P}_\lambda(h)\big). \tag{4.4}$$

The solution g_h is unique except at $k=0$. However, the value $g_h(0)$ does not enter into calculations.

Similar as in the Gaussian case, we now give some estimation results. For any integer $l \geq 1$, if Y has up to $(l+1)^{\text{th}}$ order moments, then

$$E\big[|Y^* - Y|^l\big] = \frac{1}{\lambda}E\big[Y|Y^s - 1|^l\big]$$

where $Y^s = Y - \widetilde{Y}$ and $\widetilde{Y}$ is an independent copy of Y. Denote by $\Delta g(x) = g(x+1) - g(x)$, Chen has established $\|\Delta g_h\| \leq 6\|h\| \min\big(\lambda^{-\frac{1}{2}}, 1\big)$, which implies

$$|E[h(Y)] - \mathcal{P}_\lambda(h)| \leq 6\|h\| \min\Big(\frac{1}{\sqrt{\lambda}}, 1\Big)E\big[Y|Y^s - 1|\big]. \tag{4.5}$$

There also exist other estimations for the error bound (e.g. Barbour and Eagleson [BE83]). Here we are more interested in the order than in the optimal constant.

We consider the sum of independent random variables. The following result is similar with Proposition 2.7.

Proposition 4.1 *For any $i = 1, \cdots, n$, let Y_i be an independent non-negative integer valued random variable with expectation λ_i and let Y_i^* have the Y_i-Poisson zero biased distribution. Denote by $(\bar{Y}, \bar{Y}^*) = (Y_1, \cdots, Y_n, Y_1^*, \cdots, Y_n^*)$ and suppose that they are mutually independent random variables. Set $V = Y_1 + \cdots + Y_n$ and denote by $\lambda_V = E[V]$. Let I be a random index independent of $(\bar{Y}, \bar{Y}^*)$ satisfying $P(I = i) = \lambda_i / \lambda_V$. Then $V^{(I)} + Y_I^*$ has the V-Poisson zero biased distribution where $V^{(i)} = V - Y_i$.*

4.2 First order correction

The first order corrector for the Poisson approximation is given below. The difference between the Gaussian approximation is that V is a non-negative random variable. However, if we make expansion around V, the difference $V^* - V$ is not necessarily positive. So we shall work with the partial sum variables $V^{(i)}$.

Theorem 4.2 *With the notation of Proposition 4.1, for any bounded function h defined on $\mathbb{N}_+$, the Poisson approximation $\mathcal{P}_{\lambda_V}(h)$ of $E[h(V)]$ has the following corrector*

$$C_h^{\mathcal{P}} = \frac{\lambda_V}{2} \mathcal{P}_{\lambda_V}(\Delta^2 h)\, E[Y_I^* - Y_I]. \tag{4.6}$$

The corrected approximation error is bounded by

$$\begin{aligned} \big|E[h(V)] - \mathcal{P}_{\lambda_V}(h) - C_h^{\mathcal{P}}\big| &\le \|\Delta^2 g_h\| \sum_{i=1}^n \frac{\lambda_i}{2} E\Big[|Y_i^* - Y_i|\big(|Y_i^* - Y_i| + 1\big)\Big] \\ &+ \mathrm{Var}\big[\Delta g_h(V+1)\big]^{\frac{1}{2}} \Big(\sum_{i=1}^n \lambda_i^2 \mathrm{Var}[Y_i^* - Y_i]\Big)^{\frac{1}{2}} + 6\|\Delta g_h\| \Big(\sum_{i=1}^n E[Y_i|Y_i^s - 1|]\Big)^2. \end{aligned} \tag{4.7}$$

Proof Let us first recall the discrete Taylor formula. For any integer x and any positive integer $k \ge 1$,

$$g(x+k) = g(x) + k\Delta g(x) + \sum_{j=0}^{k-1} (k-1-j) \Delta^2 g(x+j).$$

We apply the above formula to the right-hand side of $E[h(V)] - \mathcal{P}_{\lambda_V}(h) = \lambda_V E[g_h(V^*+1) - g_h(V+1)]$ and obtain by taking expansions at $V^{(i)}+1$ that

$$\begin{aligned}
&E\big[g_h(V^*+1) - g_h(V+1) - \Delta g_h(V+1)(V^*-V)\big] \\
&= \sum_{i=1}^n \frac{\lambda_i}{\lambda_V}\Big(E\big[g_h(V^{(i)}+1) + Y_i^*\Delta g_h(V^{(i)}+1) + \sum_{j=0}^{Y_i^*-1}(Y_i^*-1-j)\Delta^2 g_h(V^{(i)}+1+j)\big] \\
&\qquad - E\big[g_h(V^{(i)}+1) + Y_i\Delta g_h(V^{(i)}+1) + \sum_{j=0}^{Y_i-1}(Y_i-1-j)\Delta^2 g_h(V^{(i)}+1+j)\big] \\
&\qquad - E\big[\Delta g_h(V^{(i)}+1)(Y_i^*-Y_i) + \sum_{j=0}^{Y_i-1}(Y_i^*-Y_i)\Delta^2 g_h(V^{(i)}+1+j)\big]\Big).
\end{aligned} \tag{4.8}$$

Observe that the zero order and the first order terms in (4.8) vanish. By combining the second order terms, (4.8) is bounded by

$$\begin{aligned}
&\Big|E\big[g_h(V^*+1) - g_h(V+1) - \Delta g_h(V+1)(V^*-V)\big]\Big| \\
&\leq \|\Delta^2 g_h\| \sum_{i=1}^n \frac{\lambda_i}{2} E\Big[|Y_i^*-Y_i|\big(|Y_i^*-Y_i|+1\big)\Big]
\end{aligned}$$

We then make the following decomposition

$$\begin{aligned}
E\big[\Delta g_h(V+1)(V^*-V)\big] &= \mathcal{P}_{\lambda_V}(\Delta g_h(x+1))E[Y_I^*-Y_I] + \mathrm{cov}\big(Y_I^*-Y_I, \Delta g_h(V+1)\big) \\
&\quad + \big(E[\Delta g_h(V+1)] - \mathcal{P}_{\lambda_V}(\Delta g_h(x+1))\big)\,E[Y_I^*-Y_I].
\end{aligned} \tag{4.9}$$

The first term of the right-hand side of (4.9) is the candidate for the corrector. For the second term, we use the conditional expectation technique similar as in Proposition 2.10 to obtain

$$\mathrm{Cov}\big[\Delta g_h(V+1), Y_I^*-Y_I\big] \leq \frac{1}{\lambda_V}\mathrm{Var}\big[\Delta g_h(V+1)\big]^{\frac{1}{2}}\Big(\sum_{i=1}^n \lambda_i^2 \mathrm{Var}[Y_i^*-Y_i]\Big)^{\frac{1}{2}}.$$

For the last term, by the zero order estimation,

$$\big(E[\Delta g_h(V+1)] - \mathcal{P}_{\lambda_V}(\Delta g_h(x+1))\big)E[Y_I^*-Y_I] \leq \frac{6\|\Delta g_h\|}{\lambda_V}\Big(\sum_{i=1}^n E[Y_i|Y_i^s-1|]\Big)^2.$$

Combining all the estimations, we obtain the error estimation (4.7). At last, it remains to observe that $\mathcal{P}_{\lambda_V}(\Delta g_h(x+1)) = \frac{1}{2}\mathcal{P}_{\lambda_V}(\Delta^2 h)$.

The Poisson corrector $C_h^{\mathcal{P}}$ is of similar form with the Gaussian one and contains two terms as well: one term depending on the moments of Y_i and a Poisson expectation.

If $h(x) = (x-k)^+$ where k is now an integer, then $\Delta^2 h(x) = I_{\{x=k-1\}}$. So the Poisson corrector for the call function is given by

$$C_h^{\mathcal{P}} = \frac{\sigma_V^2 - \lambda_V}{2(k-1)!} e^{-\lambda_V} \lambda_V^{k-1}. \tag{4.10}$$

The corrector $C_h^{\mathcal{P}}$ vanishes when the expectation and the variance of V are equal.

To compare the order of the Poisson corrector and the Gaussian one, it needs a normalization on the call function. Interested reader may refer to [EKJK07].

The boundedness condition on h can be relaxed in Theorem 4.2. In fact, we observe from the proof that to obtain the estimation, it suffices that g_h is bounded, while the function h is not necessarily bounded itself. This is the case for the call function. The following result shows that for any h with linear increasing speed, g_h is bounded. We leave the proof to Appendix.

Lemma 4.3 *Let h be a function defined on $\mathbb{N}_+$. If $|h(i)| \le ai$ where a is some constant, then $\|g_h\| \le a\big(2e^{\lambda} - 1\big)$.*

5 Application to CDOs tranches

In this section, we apply the Gauss and the Poisson approximations to calculate the CDOs tranches. We first describe the payoff of a CDO tranche. Suppose that the underlying portfolio is of size n and that the weight of each name is identical. The percentage cumulative loss up to time T is given by

$$l_T = \frac{1}{n} \sum_{i=1}^{n} (1 - R_i) I_{\{\tau_i \le T\}}$$

where τ_i denotes the default time of the i^{th} name and R_i is its recovery rate.

For each CDO tranche, there exist an upper barrier k_u and a lower barrier k_l. The loss on the tranche is defined as the call spread $E[(l_T - k_l)^+] - E[(l_T - k_u)^+]$. Since the cash flow of a CDO tranche is determined by the loss on the tranche, the pricing problem can be reduced to calculating the expectation of a call function, i.e., $E[(l_T - k)^+]$ where $0 \le k \le 1$.

The market adopts the factor model framework to describe the correlations between default times for the CDOs. To be more precise, the defaults are supposed to be correlated through a common market factor U, and conditional on U, the indicator default variables $I_{\{\tau_i \le T\}}$'s are independent. Suppose in addition that R_i's are mutually independent, then l_T is a sum of conditionally independent random variables. To compute $E[h(l_T)]$, we consider firstly its conditional expectation given the factor U, i.e. $H(U) = E[h(l_T)|U]$. Secondly we integrate with respect to the law of the factor and obtain $E[h(l_T)] = \int_R H(u) p_U(u) du$ where $p_U(u)$ is the probability density of U.

In practice, the conditional distribution of each default τ_i is given as a function $F^i(u)$ such that $\int_R F^i(u) du = p_i$ where p_i is the default probability of τ_i up to T, i.e., $p_i = P(\tau_i \le T)$. The function F^i depends on p_i, and also on the correlation parameter. For example, in the standard homogeneous normal factor model,

$$F^i(u) = N\Big(\frac{N^{-1}(p_i) - \sqrt{\rho}N^{-1}(u)}{\sqrt{1-\rho}}\Big)$$

where N is the distribution function of $N(0,1)$ and ρ is the correlation between two default times. The function F^i is often defined in a non-parametric way such that the correlation parameters can be calibrated from market prices for each tranche.

We now combine the Gauss and the Poisson approximations described in previous sections to calculate the conditional expectations.

In the Gauss approximation, to apply Theorem 3.1, we first need to centralize the summand variables. Let $\xi_i = \frac{1}{n}(1-R_i)I_{\{\tau_i \le T\}}$ and denote by μ_i and σ_i its expectation and standard deviation respectively. Let $X_i = \xi_i - \mu_i$ and $W = \sum_{i=1}^n X_i$. Clearly, W is of mean zero and its variance is $\sigma_W^2 = \sum_{i=1}^n \sigma_i^2$. Then $E[(l_T - k)^+] = E[(W - \widetilde{k})^+]$ where $\widetilde{k} = k - \sum_{i=1}^n \mu_i$. If X_i's are mutually independent, which is true in the conditional setting given the factor U, we can apply Proposition 3.6 to obtain the first-order approximation

$$\Phi_{\sigma_W}\big((x-\widetilde{k})^+\big) + \frac{1}{6\sigma_W^2}\sum_{i=1}^n E[X_i^3]\,\widetilde{k}\phi_{\sigma_W}(\widetilde{k}).$$

In this conditional case, the moments of X_i and the strike $\widetilde{k}$ depend on the factor. We can also treat stochastic recovery rates without extra difficulty. If R_i's are random variables which are independent with X_i, it suffices to calculate the moments of R_i up to the third order.

In the Poisson approximation, since the summand variables are integer-valued. So we can apply Theorem 4.2 directly to the indicator variables $I_{\{\tau_i \le T\}}$. However, the recovery rates R_i's are limited to be identical or proportional constants.

Classically, it is well known that the binomial distribution $B(n,p)$ can be approximated by a Gaussian distribution when $np > 10$, otherwise, it approaches a Poisson distribution. The choice between the two first-order approximations is also determined by the values of np.

In Figure 1 are presented the numerical results for various values of np. The two curves in each graph present respectively the errors of the corrected Gauss and Poisson approximations for the call function. We observe that the Poisson approximation outperforms the Gaussian one for small values of np, the threshold is around 15, which is superior than 10 because of the correction term. For each value of np, the minimal error of the two approximations, that is, the error of the mixed approximation is inferior than 1 bp (10^{-4}). Comparing all the graphes, the errors are relatively larger when np is around the threshold. Therefore, the mixed approximation method is less precise in the overlapping area of the two approximations.

The Gauss approximation errors are oscillating. In addition, they attain the maximal value when the strike equals the cumulative loss. The Poisson

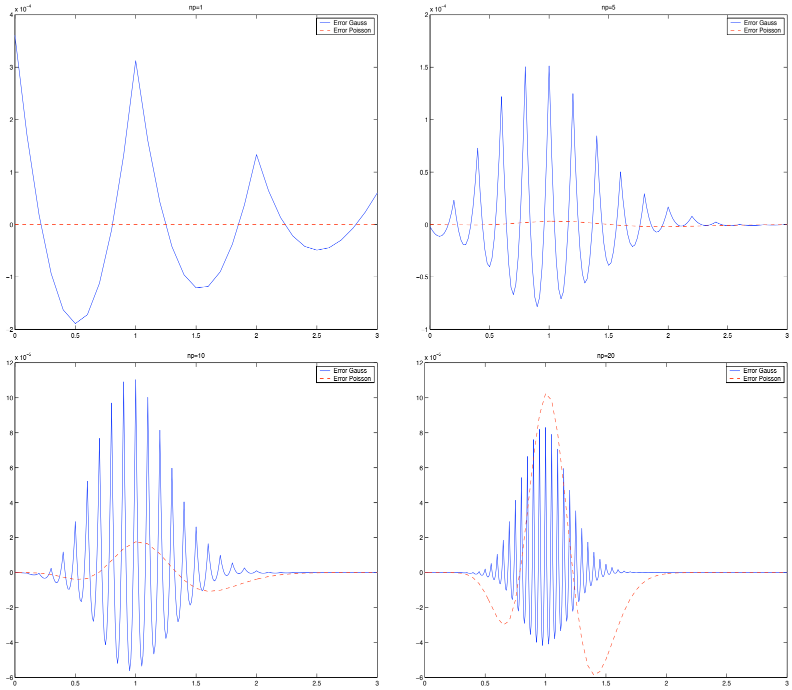

Fig. 5.1 First-order Gauss and Poisson approximation errors as a function of the strike over the expected loss for different values of np .

errors are smooth. This phenomenon is related to the fact that the binomial distribution is discrete and has been discussed in [DD04].

Let us come back to the CDOs tranches evaluation. To calculate the conditional expectation of the payoff function, we apply the mixed approximation. The choice between the two approximations depends on the average level of the conditional default probabilities, i.e., $\bar{p}(u) = \frac{1}{n}\sum_{i=1}^{n} F^i(u)$. If $n\,\bar{p}(u) > 15$, we apply the Gauss approximation, otherwise, we choose the Poisson one. The final pricing results are obtained by taking integration with respect to the law of the common factor.

We compare our method with the widely used recursive method (see [HW04]). Thanks to the closed-form formula, the Gauss-Poisson approximation method is much more rapid. For the calculation precision, we present in Figure 2 the differences for the quoted tranches prices (break even) expressed in bp(10^{-4}) between the two methods. The maximal error occurs for the junior tranche where the barriers values are 3% and 6%, This is because empirically, this tranche corresponds to the overlapping zone of the two approximations when taking into account the default probabilities of the underlying names. We remark finally that the maximal error for all tranches is less than 1.20 bp, which

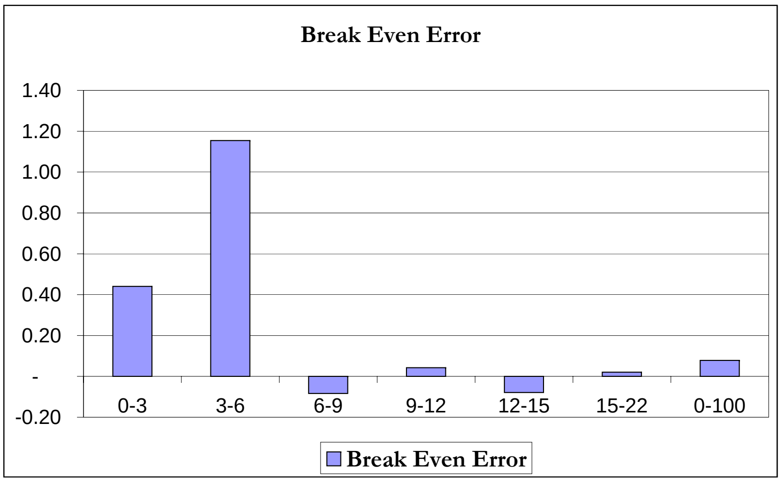

Fig. 5.2 Break even error for the quoted tranches expressed in bp.

is below the market bid-ask uncertainty and the method is hence very satisfactory.

A Estimations concerning f_h for indicator and call functions

We now give estimations for the indicator function in Lemma 3.2 and for the call function in Proposition 3.6.

We shall work with the solution of Stein's equation and its derivatives. In (2.5), the integrand function $\bar{h}$ is centered under the Gaussian distribution. However, it's no longer the case when taking derivatives. Therefore, we introduce an auxiliary function

$$\widetilde{f}_h(x) = \begin{cases} \frac{1}{\sigma^2\phi_\sigma(x)} \int_x^\infty h(t)\phi_\sigma(t)dt, & x > 0 \\ -\frac{1}{\sigma^2\phi_\sigma(x)} \int_{-\infty}^x h(t)\phi_\sigma(t)dt, & x < 0. \end{cases} \tag{A.1}$$

We also give the expectation form

$$\widetilde{f}_h(x) = \begin{cases} \frac{\sqrt{2\pi}}{\sigma} E\big[h(Z+x)e^{-\frac{Zx}{\sigma^2}} I_{\{Z>0\}}\big], & x > 0 \\ -\frac{\sqrt{2\pi}}{\sigma} E\big[h(Z+x)e^{-\frac{Zx}{\sigma^2}} I_{\{Z<0\}}\big], & x < 0 \end{cases} \tag{A.2}$$

where $Z \sim N(0, \sigma^2)$. In general, $\widetilde{f}_h$ can not be extended as a continuous function on R. if $E[|h(Z)|] < +\infty$, we have $\widetilde{f}_h(0+) = \frac{\sqrt{2\pi}}{\sigma} E[h(Z)I_{\{Z>0\}}]$ and $\widetilde{f}_h(0-) = -\frac{\sqrt{2\pi}}{\sigma} E[h(Z)I_{\{Z<0\}}]$. The two limits are equal if and only if $\Phi_\sigma(h) = E[h(Z)] = 0$. Furthermore, if $\Phi_\sigma(h) = 0$, then $\widetilde{f}_h$ coincides with the solution f_h of Stein's equation.

We first give some simple properties of the function $\widetilde{f}_h$.

Lemma A.1 *If $|h(x)| \le g(x)$ and if $g(x)/|x|$ is decreasing when $x > 0$ and is increasing when $x < 0$, then*

$$|\widetilde{f}_h(x)| \le \frac{g(x)}{|x|}.$$

Proof When $x > 0$, we have

$$|\widetilde{f}_h(x)| \le \frac{1}{\sigma^2 \phi_\sigma(x)} \int_x^\infty g(t)\phi_\sigma(t)dt = -\frac{1}{\phi_\sigma(x)} \int_x^\infty \frac{g(t)}{t} d\phi_\sigma(t).$$

Since $\frac{g(t)}{t}$ is decreasing, we get the inequality. When $x < 0$, the proof is similar.

We now give estimations for $\widetilde{f}_h$ and $\widetilde{f}'_h$ when h is a bounded function. The indicator function $I_\alpha(x) = I_{\{x \le \alpha\}}$ satisfies the boundedness condition with $c_0 = 1$.

Proposition A.2 *Let h be a bounded function and let $c_0 = \|h\|$, then for any $x \in \mathbb{R} \setminus \{0\}$,*

$$|\widetilde{f}_h(x)| \le \sqrt{2\pi} c_0 / 2\sigma, \qquad |\widetilde{f}'_h(x)| \le 2c_0/\sigma^2.$$

Proof Since $\lim_{x \to 0+} \widetilde{f}_1(x) = \sqrt{2\pi}/2\sigma$, $\lim_{x \to 0-} \widetilde{f}_1(x) = -\sqrt{2\pi}/2\sigma$, and $\lim_{|x| \to +\infty} \widetilde{f}_1(x) = 0$, we only need to prove that $\widetilde{f}_1$ is decreasing when $x > 0$ and when $x < 0$ respectively. By letting $h = g \equiv 1$ in Lemma A.1, we obtain $|x\widetilde{f}_1(x)| \le 1$. Hence

$$\widetilde{f}'_1(x) = \begin{cases} \frac{1}{\sigma^2}(x\widetilde{f}_1(x) - 1) \le 0, & x > 0 \\ -\frac{1}{\sigma^2}(1 - x\widetilde{f}_1(x)) \le 0 & x < 0. \end{cases}$$

So $|\widetilde{f}_1(x)| \le \frac{\sqrt{2\pi}}{2\sigma}$ for any $x \in \mathbb{R}$, which implies the first inequality. For the second inequality, by using $\widetilde{f}'_h(x) = \frac{1}{\sigma^2}(x\widetilde{f}_h(x) - h(x))$, together with estimations $|x\widetilde{f}_h(x)| \le |c_0 x \widetilde{f}_1(x)| \le c_0$ and $|h(x)| \le c_0$, we obtain $|\widetilde{f}'_h(x)| \le 2c_0/\sigma^2$.

The following lemma is useful to estimate $\|xf'_{I_\alpha}\|$. The argument is that the polynomial functions increase slower than the exponential functions.

Lemma A.3 *Let $Z \sim N(0, \sigma^2)$. Then*

$$xE[I_{\{Z>0\}} e^{-\frac{Zx}{\sigma^2}}] \le \frac{\sigma}{\sqrt{2\pi}} \quad \textit{if} \quad x > 0, ; \qquad |x|E[I_{\{Z<0\}} e^{-\frac{Zx}{\sigma^2}}] \le \frac{\sigma}{\sqrt{2\pi}} \quad \textit{if} \quad x \le 0. \quad \text{(A.3)}$$

For any non-negative integers l, m satisfying $l \le m$,

$$\begin{cases} E\big[I_{\{Z>0\}} x^l Z^m e^{-\frac{Zx}{\sigma^2}}\big] & \le \frac{1}{2}\big(\frac{l\sigma^2}{e}\big)^l E\big[|Z|^{m-l}\big], \quad x > 0; \\ E[I_{\{Z<0\}} |x|^l |Z|^m e^{-\frac{Zx}{\sigma^2}}] & \le \frac{1}{2}\left(\frac{l\sigma^2}{e}\right)^l E[|Z|^{m-l}] \quad x \le 0. \end{cases} \quad \text{(A.4)}$$

Proof When $x > 0$, it suffices to observe $xE[I_{\{Z>0\}} e^{-\frac{Zx}{\sigma^2}}] = \frac{\sigma}{\sqrt{2\pi}} x \widetilde{f}_1(x)$ to prove (A.3). For (A.4), consider the function $f(y) = y^l e^{-\frac{y}{\sigma^2}}$, it attains the maximum value at $y = l\sigma^2$ and $|f(y)| \le (\frac{l\sigma^2}{e})^l$. Then the lemma follows immediately. The case $x < 0$ is obtained by symmetry.

Proposition A.4 *For any real number $\beta \in [0, 1]$,*

$$|x\widetilde{f}'_{I_\alpha - \beta}(x)| \le \frac{\sqrt{2\pi}}{2\sigma e} + \frac{|\alpha|}{\sigma^2}. \quad \text{(A.5)}$$

Proof We only prove for $x>0$. By definition,

$$\widetilde{f}_{I_\alpha-\beta}(x)=\frac{\sqrt{2\pi}}{\sigma}E[I_{\{Z>0\}}(I_\alpha(x+Z)-\beta)e^{-\frac{Zx}{\sigma^2}}].$$

Then

$$\widetilde{f}'_{I_\alpha-\beta}(x)=-\frac{\sqrt{2\pi}}{\sigma^3}E[I_{\{Z>0\}}(I_\alpha(x+Z)-\beta)Ze^{-\frac{Zx}{\sigma^2}}]+I_{\{x\leq\alpha\}}\frac{1}{\sigma^2}e^{\frac{x^2-\alpha^2}{2\sigma^2}}.$$

Using Lemma A.3 with $l=m=1$ and the fact that $\|I_\alpha-\beta\|\leq 1$, we get

$$\Big|xE[I_{\{Z>0\}}(I_\alpha(x+Z)-\beta)Ze^{-\frac{Zx}{\sigma^2}}]\Big|\leq\frac{\sigma^2}{2e}.$$

In addition, $xI_{\{x\leq\alpha\}}e^{\frac{x^2-\alpha^2}{2\sigma^2}}\leq|\alpha|$. Then combining the two inequalities, we obtain (A.5).

We can now deduce the estimations concerning f_{I_α} by using the auxiliary function $\widetilde{f}_{I_\alpha}$.

Corollary A.5 *Let* $I_\alpha(x)=I_{\{x\leq\alpha\}}$. *Then*

$$\|f_{I_\alpha}\|\leq\frac{\sqrt{2\pi}}{2\sigma},\qquad\|f'_{I_\alpha}\|\leq\frac{2}{\sigma^2},\qquad\|xf'_{I_\alpha}\|\leq\frac{\sqrt{2\pi}}{2\sigma e}+\frac{|\alpha|}{\sigma^2}.$$

Proof Notice that $f_{I_\alpha}=\widetilde{f}_{\bar{I}_\alpha}$ where $\bar{I}_\alpha=I_\alpha-N_\sigma(\alpha)$. Since $\|\bar{I}_\alpha\|\leq 1$, we can apply Proposition A.2 to obtain the first two inequalities. Proposition A.4 implies the third one.

In the following, we consider functions with bounded derivatives, whose increasing speed is at most linear. The call function satisfies this property.

Proposition A.6 *Let h be an absolutely continuous function on* $\mathbb{R}$.

1) Let $c_1=|h(0)|$ *and suppose that* $c_0=\|h'\|<+\infty$, *then*

$$|\widetilde{f}_h(x)|\leq\frac{\sqrt{2\pi}c_1}{2\sigma}+2c_0,\quad|\widetilde{f}'_h(x)|\leq\frac{\sqrt{2\pi}c_0}{\sigma}\Big(1+\frac{1}{2e}\Big)+\frac{c_1}{\sigma^2}.$$

2) If, in addition, h' is locally of finite variation and has finite number of jumps. Let $h'=g_1+g_2$, where g_1 is the continuous part of h' and g_2 is the pure jump part of the following form

$$g_2(x)=\sum_{i=1}^{N}\epsilon_i(I_{\mu_i}-\beta_i).$$

We assume that $c_3=\|g'_1\|<+\infty$ *and* $c_4=\|g_1\|$, *then*

$$|x\widetilde{f}''_h(x)|\leq c_3+\frac{\sqrt{2\pi}c_4}{2\sigma e}+\sum_{i=1}^{N}|\varepsilon_i|\Big(\frac{\sqrt{2\pi}}{2\sigma e}+\frac{|\mu_i|}{\sigma^2}\Big)+\frac{1}{e\sigma}\Big(\frac{c_1}{\sigma}+\sqrt{2\pi}c_0+\frac{2\sqrt{2\pi}c_0}{e}\Big).$$

Proof Clearly we have $|h(x)|\leq c_1+c_0|x|$ for any $x\in\mathbb{R}$. By a symmetric argument it suffices to prove the inequalities for $x>0$.

1) By (A.3) and (A.4), we have

$$\begin{aligned}|\widetilde{f}_h(x)|&\leq\frac{\sqrt{2\pi}}{\sigma}E[I_{\{Z>0\}}(c_1+c_0Z+c_0x)e^{-\frac{Zx}{\sigma^2}}]\\&\leq\frac{\sqrt{2\pi}}{\sigma}\Big[\frac{c_1}{2}+\frac{c_0}{2}E[|Z|]+\frac{c_0\sigma}{\sqrt{2\pi}}\Big]\leq\frac{\sqrt{2\pi}c_1}{2\sigma}+2c_0\end{aligned}$$

since $E[|Z|]=\frac{2\sigma}{\sqrt{2\pi}}$. Taking the derivative,

$$\widetilde{f}_h'(x)=\frac{\sqrt{2\pi}}{\sigma}E[I_{\{Z>0\}}h'(Z+x)e^{-\frac{Zx}{\sigma^2}}]-\frac{\sqrt{2\pi}}{\sigma^3}E[I_{\{Z>0\}}Zh(Z+x)e^{-\frac{Zx}{\sigma^2}}]. \tag{A.6}$$

Using similar argument as above, we have

$$\begin{aligned}|\widetilde{f}_h'(x)| &\le \frac{\sqrt{2\pi}c_0}{2\sigma}+\frac{\sqrt{2\pi}}{\sigma^3}E[I_{\{Z>0\}}Z(c_0Z+c_0x+c_1)e^{-\frac{Zx}{\sigma^2}}]\\ &\le \frac{\sqrt{2\pi}c_0}{2\sigma}+\frac{\sqrt{2\pi}}{\sigma^3}\Big(\frac{c_0\sigma^2}{2}+\frac{c_0\sigma^2}{2e}+\frac{c_1\sigma}{\sqrt{2\pi}}\Big)=\frac{\sqrt{2\pi}c_0}{\sigma}\Big(1+\frac{1}{2e}\Big)+\frac{c_1}{\sigma^2}.\end{aligned}$$

2) First, we have by (A.6),

$$\widetilde{f}_h''=\widetilde{f}_{h'}'-\frac{\sqrt{2\pi}}{\sigma^3}E[I_{\{Z>0\}}Zh'(Z+x)e^{-\frac{Zx}{\sigma^2}}]+\frac{\sqrt{2\pi}}{\sigma^5}E[I_{\{Z>0\}}Z^2h(Z+x)e^{-\frac{Zx}{\sigma^2}}].$$

By the linearity of $\widetilde{f}_h$ with respect to h, we know that

$$\widetilde{f}_{h'}'=\widetilde{f}_{g_1}'+\sum_{i=1}^N\epsilon_i\widetilde{f}_{I_{\mu_i}-\beta_i}'.$$

So (A.6) and Proposition A.4 imply that

$$|x\widetilde{f}_{h'}'(x)|\le|x\widetilde{f}_{g_1}'(x)|+\sum_{i=1}^N|\epsilon_i|\Big(\frac{\sqrt{2\pi}}{2\sigma e}+\frac{|\mu_i|}{\sigma^2}\Big)\le c_3+\frac{\sqrt{2\pi}c_4}{2\sigma e}+\sum_{i=1}^N|\epsilon_i|\Big(\frac{\sqrt{2\pi}}{2\sigma e}+\frac{|\mu_i|}{\sigma^2}\Big).$$

The other two terms are estimated similarly,

$$\Big|xE[I_{\{Z>0\}}Zh'(Z+x)e^{-\frac{Zx}{\sigma^2}}]\Big|\le\frac{c_0\sigma^2}{2e}$$

$$\Big|xE[I_{\{Z>0\}}Z^2h(Z+x)e^{-\frac{Zx}{\sigma^2}}]\Big|\le\frac{c_0\sigma^4}{2e}+\frac{2c_0\sigma^4}{e^2}+\frac{c_1\sigma^3}{\sqrt{2\pi}e}.$$

So we get finally

$$|x\widetilde{f}_h''(x)|\le c_3+\frac{\sqrt{2\pi}c_4}{2\sigma e}+\sum_{i=1}^N|\varepsilon_i|\Big(\frac{\sqrt{2\pi}}{2\sigma e}+\frac{|\mu_i|}{\sigma^2}\Big)+\frac{1}{e\sigma}\Big(\frac{c_1}{\sigma}+\sqrt{2\pi}c_0+\frac{2\sqrt{2\pi}c_0}{e}\Big).$$

For the call function, we apply directly the above Proposition.

Corollary A.7 *Let $C_k=(x-k)^+$, then*

$$\|f_{C_k}\|\le 2+\frac{\sqrt{2\pi}}{2\sigma}c_1$$

where $c_1=|(-k)^+-\bar{c}|$ and $\bar{c}=\Phi_\sigma((x-k)^+)=\sigma^2\phi_\sigma(k)-k(1-\Phi_\sigma(k))$.

$$\|f_{C_k}'\|\le\frac{\sqrt{2\pi}}{\sigma}\Big(1+\frac{1}{2e}\Big)+\frac{c_1}{\sigma^2}$$

and

$$|xf_{C_k}''|\le\frac{c_1}{e\sigma^2}+\frac{|k|}{\sigma^2}+\frac{2\sqrt{2\pi}}{\sigma e}\Big(1+\frac{1}{e}\Big).$$

Proof We have $f_{C_k}=\widetilde{f}_{\overline{C}_k}$ where $\overline{C}_k=C_k-\bar{c}$. In addition, $\|\overline{C}_k'\|=1$ and $c_1=|\overline{C}_k(0)|=|(-k)^+-\bar{c}|$. Applying Proposition A.6, we get the first inequalities. And it suffices to notice $c_3=0$ and $c_4=1$ to end the proof.

B Proof of Lemma 4.3

Proof Using the solution of Stein's Poisson equation (4.4), we have for any integer $k \geq 0$,

$$|g_h(k)| \leq \frac{(k-1)!}{\lambda^k} \sum_{i=k}^{\infty} \frac{\lambda^i}{i!} \big|h(i) - \mathcal{P}_\lambda(h)\big| \leq \frac{(k-1)!}{\lambda^k} \sum_{i=k}^{\infty} \frac{\lambda^i}{i!}(ai + a\lambda).$$

For the first term,

$$\frac{(k-1)!}{\lambda^k} \sum_{i=k}^{\infty} \frac{\lambda^i}{(i-1)!} = \sum_{i=k}^{\infty} \frac{\lambda^{i-k}}{k(k+1)\cdots(i-1)} = \sum_{j\geq 0} \frac{\lambda^j}{k(k+1)\cdots(k+j-1)} \leq \sum_{j\geq 0} \frac{\lambda^j}{j!} = e^\lambda.$$

For the second term,

$$\frac{(k-1)!}{\lambda^k} \sum_{i=k}^{\infty} \frac{\lambda^i}{i!} = \sum_{i=k}^{\infty} \frac{\lambda^{i-k}}{k(k+1)\cdots i} = \sum_{j\geq 0} \frac{\lambda_j}{k(k+1)\cdots(k+j)} \leq \sum_{j\geq 0} \frac{\lambda^j}{(j+1)!} = \frac{1}{\lambda}(e^\lambda - 1).$$

Hence we have by combining the two terms

$$\|g_h\| \leq a(2e^\lambda - 1).$$

References

[Bar86] A. D. Barbour. Asymptotic expansions based on smooth functions in the central limit theorem. *Probability Theory and Related Fields*, 72:289–303, 1986.

[Bar87] A. D. Barbour. Asymptotic expansions in the poisson limit theorem. *Annals of Probability*, 15:748–766, 1987.

[BE83] A. D. Barbour and G. K. Eagleson. Poisson approximation for some statistics based on exchangeable trials. *Advances in Applied Probability*, 15(3):585–600, 1983.

[Che75] L. H. Y. Chen. Poisson approximation for dependent trials. *Annals of Probability*, 3:534–545, 1975.

[CS01] L. H. Y. Chen and Q.-M. Shao. A non-uniform Berry-Esseen bound via Stein's method. *Probability Theory and Related Fields*, 120:236–254, 2001.

[CS05] L. H. Y. Chen and Q.-M. Shao. Stein's method for normal approximation. In *An Introduction to Stein's Method*, volume 4 of *Lecture Notes Series, IMS, National University of Singapore*, pages 1–59. Singapore University Press and World Scientific Publishing Co. Pte. Ltd., 2005.

[DD04] F. Diener and M. Diener. Asymptotics of price oscillations of a european Call option in a tree model. *Mathematical Finance*, 14(2):271–293, 2004.

[EKJK07] N. El Karoui, Y. Jiao, and D. Kurtz. Gauss and poisson approximation: applications to CDOs tranche pricing. to appear in Journal of Computational Finance, 2007.

[Fel71] W. Feller. *An Introduction to Probability and Its Applications, 2nd ed.* Wiley, New York, 1971.

[GH78] F. Götze and C. Hipp. Asymptotic expansions in the central limit theorem under moment conditions. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 42(1):67–87, 1978.

[Gol07] L. Goldstein. L^1 bounds in normal approximation. *Annals of Probability*, 35(5):1888–1930, 2007.

[GR97] L. Goldstein and G. Reinert. Stein's method and the zero bias transformation with application to simple random sampling. *Annals of Applied Probability*, 7:935–952, 1997.

[Hip77] C. Hipp. Edgeworth expansions for integrals of smooth functions. *Ann. Probability*, 5(6):1004–1011, 1977.

[HW04] J. Hull and A White. Valuation of a CDO and an n^{th}-to-default CDS without Monte Carlo simulation. Working paper, University of Toronto, 2004.

[Rai03] M. Raic. Normal approximations by Stein's method. In Andrej Mrvar, editor, *Proceedings of the seventh young statisticians meeting*, pages 71–97. 2003.

[Ste72] C. Stein. A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. In *Proc, Sixty Berkeley Symp. Math. Statist. Probab.*, pages 583–602. Univ. California Press, Berkeley, 1972.

[Ste86] C. Stein. *Approximate Computation of Expectations.* IMS, Hayward, CA., 1986.

[TYW05] K. Tanaka, T. Yamada, and T. Watanabe. Approximation of interest rate derivatives' prices by gram-charlier expansion and bond moments. Working paper, Bank of Japan, 2005.

[Vas91] O. Vasicek. Limiting loan loss probability distribution. Working paper, Moody's KMV, 1991.