Zero bias transformation and asymptotic expansions

Ying Jiao

Laboratoire de probabilités et modèles aléatoires, Université Paris VII, Paris, France. E-mail: jiao@math.jussieu.fr

Abstract. Let W be a sum of independent random variables. We apply the zero bias transformation to deduce recursive asymptotic expansions for $\mathbb{E}[h(W)]$ in terms of normal expectations, or of Poisson expectations for integer-valued random variables. We also discuss the estimates of remaining errors.

Résumé. Soit W une somme de variables aléatoires indépendants. On applique la transformation zéro biais pour obtenir de façon recursive des développements asymptotiques de $\mathbb{E}[h(W)]$ en terme d'espérances par rapport à la loi normale, ou à la loi de Poisson si les variables aléatoires sont à valeurs entières. On discute aussi les bornes des termes d'erreur.

MSC: 60G50; 60F05

Keywords: Normal and Poisson approximations; Zero bias transformation; Stein's method; Reverse Taylor formula; Concentration inequality

1. Introduction

The zero bias transformation has been introduced by Goldstein and Reinert [15] in the framework of Stein's method. By the fundamental works of Stein [24,25], we know that a random variable (r.v.) Z with mean zero follows the normal distribution $N(0,\sigma^2)$ if and only if $\mathbb{E}[Zf(Z)] = \sigma^2\mathbb{E}[f'(Z)]$ for any Borel function f such that both sides of the equality are well defined. More generally, for any r.v. X with mean zero and finite variance $\sigma^2 > 0$, a r.v. X^* is said to have the zero biased distribution of X if the equality

$$\mathbb{E}\big[Xf(X)\big] = \sigma^2\mathbb{E}\big[f'\big(X^*\big)\big] \tag{1}$$

holds for any differentiable function f such that (1) is well defined. By *Stein's equation*

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

where h is a given function and $\Phi_\sigma(h)$ denotes the expectation of h under the distribution $N(0,\sigma^2)$,

$$\mathbb{E}\big[h(X)\big] - \Phi_\sigma(h) = \mathbb{E}\big[Xf_h(X) - \sigma^2 f_h'(X)\big] = \sigma^2\mathbb{E}\big[f_h'\big(X^*\big) - f_h'(X)\big], \tag{3}$$

where f_h is the solution of Stein's equation and is given by

$$f_h = \frac{1}{\sigma^2\phi_\sigma(x)}\int_x^\infty \big(h(t) - \Phi_\sigma(h)\big)\phi_\sigma(t)\,\mathrm{d}t. \tag{4}$$

An important remark is that X^* needs not be independent of X ([15], see also [13]). Let $W = X_1 + \cdots + X_n$ where X_i $(i = 1,\ldots,n)$ are independent random variables with mean zero and variance $\sigma_i > 0$, $\sigma_W^2 := \mathrm{Var}[W] = \sigma_1^2 + \cdots + \sigma_n^2$. Goldstein and Reinert have proposed the construction $W^* = W^{(I)} + X_I^*$ where $W^{(i)} := W - X_i$, X_i^* is independent of $W^{(i)}$, and I is a random index taking values in $\{1,\dots,n\}$ and independent of $(X_1,\dots,X_n,X_1^*,\dots,X_n^*)$ such that $\mathbb{P}(I=i)=\sigma_i^2/\sigma_W^2$. This construction of the zero bias transformation is similar to Lindeberg's method except that we here consider an average of punctual substitutions of X_i by X_i^*, while in Lindeberg's method, X_i's are progressively substituted by a central normal r.v. with the same variance.

The asymptotic expansion of $\mathbb{E}[h(W)]$ is a classical topic related to the central limit theorems. Using Stein's method, Barbour [2,3] has obtained a full expansion of $\mathbb{E}[h(W)]$ for sufficiently regular functions. Compared to the classical Edgeworth expansion (see [21], Chapter V, also [22]), the results of [2] do not require that the distributions of X_i's be smooth. However, as a price paid, we need some regularity conditions on the function h. The result of [2] can also be compared to those in [17,18] using Fourier transform. The key point of Barbour's method is a Taylor type formula with cumulant coefficients, which allows to write the difference $\mathbb{E}[Wf(W)]-\sigma_W^2\mathbb{E}[f'(W)]$ as a series containing the cumulants of W. The W-expectations are then replaced by the normal expectations in an iterative way until the desired order. It has been pointed out in [22] that the key formula of Barbour can also be obtained by Fourier transform.

The zero bias transformation has been used in [11] to obtain a first-order correction term for the normal approximation of $\mathbb{E}[h(W)]$, where the motivation was to find a rapid numerical method for large-sized credit derivatives. The function of interest is the so-called *call function* in finance: $h(x)=(x-k)_+$ where k is a real number. Since such h is only absolutely continuous, the function f_h is not regular enough to have a third-order derivative. To achieve the estimates, we used a conditional expectation technique, together with a concentration inequality due to Chen and Shao [8,9].

The difficulty for generalizing the result in [11] to obtain a full expansion of $\mathbb{E}[h(W)]$ is that W and W^*-W are not independent. If we consider the Taylor expansion of $f_h'(W^*)$ at W and then take the expectation, the terms of the form $\mathbb{E}[f_h^{(l)}(W)(W^*-W)^k]$ appear, where $f^{(l)}$ denotes the lth derivative of f. For the first-order expansion in [11], a conditional expectation argument allowed us to replace $\mathbb{E}[f_h''(W)(W^*-W)]$ by $\mathbb{E}[f_h''(W)]\mathbb{E}[W^*-W]$ and to put the covariance in the error term. However, in the higher-order expansions, the error term can no longer contain such covariances. An alternative way is to consider the Taylor expansion of $\mathbb{E}[f_h'(W^*)-f_h'(W)]$ at $W^{(i)}$. Since X_i^* is independent of $W^{(i)}$, there is no crossing term. However, the expectations of the form $\mathbb{E}[f_h^{(l)}(W^{(i)})]$ appear, which makes it difficult to apply the recurrence procedure. To overcome this difficulty, we shall propose a so-called reverse Taylor formula which will enable us to replace $\mathbb{E}[f_h^{(l)}(W^{(i)})]$ by $\mathbb{E}[f_h^{(l)}(W)]$ up to an error term.

The main result of this paper is an expansion formula $\mathbb{E}[h(W)]=C_N(h)+e_N(h)$, $C_N(h)$ being the principal term of the Nth-order expansion and $e_N(h)$ being the remaining error. We shall precise the conditions under which the expansion formula is well defined. In particular, in order to include some irregular functions, we shall consider the continuous and the jump parts of h separately. Moreover, we also need to show that $e_N(h)$ is "small" enough as an error term. In our results, $e_N(h)$ is given in a recursive way so that it is actually a linear combination of the remaining terms of Taylor and reverse Taylor expansions and can be thus estimated. A key ingredient for the estimates is a concentration inequality which provides an upper bound for $\mathbb{P}(a\le W\le b)$ involving some power $(b-a)^\alpha$ of the interval length $(b-a)$ with $0<\alpha\le 1$. This allows to obtain estimates under relatively mild moment conditions on X_i's. It is interesting to point out that the same method can be used in a discrete setting to obtain a Poisson asymptotic expansion for integer-valued summand random variables.

The rest of the paper is organized as follows. In Section 2, we give the normal expansion by using the reverse Taylor formula and we discuss the conditions on h and on X_i's. Section 3 is devoted to error estimates. We present the Poisson case in Section 4. Finally, some technical proofs are left in Appendix.

2. Expansions in normal case

2.1. *Reverse Taylor formula*

Let N be a positive integer, X and Y be two independent random variables such that Y has up to Nth moments, and f be an N times differentiable function such that $f^{(k)}(X)$ and $f^{(k)}(X+Y)$ are integrable for any $k=0,\dots,N$.

We use the notation $m_Y^{(k)} := \mathbb{E}[Y^k]/k!$. Denote by $\delta_N(f, X, Y)$ the error term in the Nth-order Taylor expansion of $\mathbb{E}[f(X+Y)]$ centered at X, i.e.

$$\delta_N(f, X, Y) := \mathbb{E}\big[f(X+Y)\big] - \sum_{k=0}^{N} m_Y^{(k)} \mathbb{E}\big[f^{(k)}(X)\big]. \tag{5}$$

Recall that for any $N \geq 1$,

$$\delta_N(f, X, Y) = \frac{1}{(N-1)!} \int_0^1 (1-t)^{N-1} \mathbb{E}\big[\big(f^{(N)}(X+tY) - f^{(N)}(X)\big) Y^N\big] \, \mathrm{d}t \tag{6}$$

provided that the right-hand side is well defined.

The so-called reverse Taylor formula gives an expansion of $\mathbb{E}[f(X)]$ using the expectations of functions of $X+Y$ and the moments of Y. We specify some notation and conventions. First of all, $\mathbb{N}_* := \mathbb{N} \setminus \{0\}$ denotes the set of strictly positive integers. For any integer $d \geq 1$ and any $\mathbf{J} = (j_l)_{l=1}^d \in \mathbb{N}_*^d$, $|\mathbf{J}|$ is defined as $j_1 + \cdots + j_d$, and $m_Y^{(\mathbf{J})} := m_Y^{(j_1)} \cdots m_Y^{(j_d)}$. By convention, $\mathbb{N}_*^0$ denotes the set $\{\varnothing\}$ of the empty vector, $|\varnothing| = 0$ and $m_Y^{(\varnothing)} = 1$. We note that in the right–hand side of the expansion formula (7), the variables $X+Y$ and Y are not independent.

Proposition 2.1 (Reverse Taylor formula). *With the above notation, the equality*

$$\mathbb{E}\big[f(X)\big] = \sum_{d \geq 0} (-1)^d \sum_{\mathbf{J} \in \mathbb{N}_*^d, |\mathbf{J}| \leq N} m_Y^{(\mathbf{J})} \mathbb{E}\big[f^{(|\mathbf{J}|)}(X+Y)\big] + \varepsilon_N(f, X, Y) \tag{7}$$

holds, where $\varepsilon_N(f, X, Y)$ is defined as

$$\varepsilon_N(f, X, Y) = -\sum_{d \geq 0} (-1)^d \sum_{\mathbf{J} \in \mathbb{N}_*^d, |\mathbf{J}| \leq N} m_Y^{(\mathbf{J})} \delta_{N-|\mathbf{J}|}\big(f^{(|\mathbf{J}|)}, X, Y\big). \tag{8}$$

Proof. We replace $\mathbb{E}[f^{(|\mathbf{J}|)}(X+Y)]$ on the right-hand side of (7) by

$$\sum_{k=0}^{N-|\mathbf{J}|} m_Y^{(k)} \mathbb{E}\big[f^{(|\mathbf{J}|+k)}(X)\big] + \delta_{N-|\mathbf{J}|}\big(f^{(|\mathbf{J}|)}, X, Y\big)$$

and observe that the sum of terms containing δ vanishes with $\varepsilon_N(f, X, Y)$. Hence the right-hand side of (7) equals

$$\sum_{d \geq 0} (-1)^d \sum_{\mathbf{J} \in \mathbb{N}_*^d, |\mathbf{J}| \leq N} m_Y^{(\mathbf{J})} \sum_{k=0}^{N-|\mathbf{J}|} m_Y^{(k)} \mathbb{E}\big[f^{(|\mathbf{J}|+k)}(X)\big].$$

If we split the terms where $k = 0$ and where $1 \leq k \leq N - |\mathbf{J}|$, the above formula can be written as

$$\sum_{d \geq 0} (-1)^d \sum_{\mathbf{J} \in \mathbb{N}_*^d, |\mathbf{J}| \leq N} m_Y^{(\mathbf{J})} \mathbb{E}\big[f^{(|\mathbf{J}|)}(X)\big] + \sum_{d \geq 0} (-1)^d \sum_{\mathbf{J} \in \mathbb{N}_*^d, |\mathbf{J}| \leq N} m_Y^{(\mathbf{J})} \sum_{k=1}^{N-|\mathbf{J}|} m_Y^{(k)} \mathbb{E}\big[f^{(|\mathbf{J}|+k)}(X)\big]. \tag{9}$$

We make the index changes $\mathbf{J}' = (\mathbf{J}, k)$ and $u = d+1$ in the second part, and observe that it is just

$$\sum_{u \geq 1} (-1)^{u-1} \sum_{\mathbf{J}' \in \mathbb{N}_*^u, |\mathbf{J}'| \leq N} m_Y^{(\mathbf{J}')} \mathbb{E}\big[f^{(|\mathbf{J}'|)}(X)\big].$$

Thus, the terms in the first and the second parts of (9) cancel out except the one where $d = 0$ in the first part. So (9) equals $\mathbb{E}[f(X)]$, which proves the proposition. □

2.2. *Admissible function space*

In this subsection, we precise the functions for which we can make the Nth-order expansion. We need conditions on both the regularity and the increasing speed at infinity of the function.

Recall ([20], Chapter VI) that any function g on $\mathbb{R}$ which is locally of finite variation can be uniquely decomposed into the sum of a function of pure jump and a continuous function locally of finite variation and vanishing at the origin. That is, $g = g_c + g_d$ where g_c is called the *continuous part* of g and g_d is the *purely discontinuous part*.

Let $\alpha \in (0, 1]$ and $p \geq 0$ be two real numbers. For any function f on $\mathbb{R}$, the following quantity has been defined by Barbour in [2]:

$$\|f\|_{\alpha,p} := \sup_{x \neq y} \frac{|f(x) - f(y)|}{|x - y|^{\alpha}(1 + |x|^p + |y|^p)}. \tag{10}$$

The finiteness of $\|f\|_{\alpha,p}$ implies that the function f is locally α-Lipschitz, and that the increasing speed of f at infinity is at most of order $|x|^{\alpha+p}$. All functions f such that $\|f\|_{\alpha,p} < +\infty$ form a vector space over $\mathbb{R}$, and $\|\cdot\|_{\alpha,p}$ is a norm on it. We list below several properties of $\|\cdot\|_{\alpha,p}$ and we leave the proofs in Appendix A.

Lemma 2.2. *Let f be a function on $\mathbb{R}$, $\alpha, \beta \in (0, 1]$ and $p, q \geq 0$.*

(1) *If $p \leq q$, then $\|f\|_{\alpha,p} < +\infty$ implies $\|f\|_{\alpha,q} < +\infty$.*
(2) *If $\alpha \leq \beta$, then $\|f\|_{\beta,p} < +\infty$ implies $\|f\|_{\alpha,p+\beta-\alpha} < +\infty$.*
(3) *If P is a polynomial of degree d, then $\|f\|_{\alpha,p} < +\infty$ implies $\|Pf\|_{\alpha,p+d} < +\infty$.*
(4) *Assume that F is a primitive function of f, then $\|f\|_{\alpha,p} < +\infty$ implies $\|F\|_{1,p+\alpha} < +\infty$. (Hence $\|F\|_{\alpha,p+1} < +\infty$ by 2.)*

Inspired by [2], we introduce the following function space.

Definition 2.3. *Let $N \geq 0$ be an integer, and $\alpha \in (0, 1]$, $p \geq 0$ be two real numbers. Denote by $\mathcal{H}^N_{\alpha,p}$ the vector space of all Borel functions h on $\mathbb{R}$ verifying the following conditions*:

(a) *h has Nth derivative which is locally of finite variation and has finitely many jumps*,
(b) *the continuous part of $h^{(N)}$ satisfies $\|h_c^{(N)}\|_{\alpha,p} < +\infty$.*

Condition (a) implies that the pure jump part of $h^{(N)}$ is bounded. Condition (b) implies that $h_c^{(N)}$ has at most polynomial increasing speed at infinity, and so does h. These conditions allow us to include some irregular functions such as the indicator functions. Let k be a real number and $I_k(x) = \mathbb{1}_{\{x \leq k\}}$. Then $\|I_k\|_{\alpha,p}$ is clearly not finite. However, $\|I_{k,c}\|_{\alpha,p} = 0$, which means that for any $\alpha \in (0, 1]$ and any $p \geq 0$, $I_k(x) \in \mathcal{H}^0_{\alpha,p}$. Note that any function h in $\mathcal{H}^0_{\alpha,p}$ can be decomposed as $h = h_c + h_d$, where h_c satisfies $\|h_c\|_{\alpha,p} < +\infty$ and h_d is a linear combination of indicator functions of the form $\mathbb{1}_{\{x \leq k\}}$ plus a constant (so that $h_c(0) = 0$).

Proposition 2.4. *Let $N \geq 0$ be an integer, $\alpha, \beta \in (0, 1]$ and $p, q \geq 0$ be real numbers. Then the following assertions hold*:

(1) *when $N \geq 1$, $h \in \mathcal{H}^N_{\alpha,p}$ if and only if $h' \in \mathcal{H}^{N-1}_{\alpha,p}$;*
(2) *if $p \leq q$, then $\mathcal{H}^N_{\alpha,p} \subset \mathcal{H}^N_{\alpha,q}$; if $\alpha \leq \beta$, then $\mathcal{H}^N_{\beta,p} \subset \mathcal{H}^N_{\alpha,p+\beta-\alpha}$;*
(3) *when $N \geq 1$, $\mathcal{H}^N_{\alpha,p} \subset \mathcal{H}^{N-1}_{1,\alpha+p} \subset \mathcal{H}^{N-1}_{\alpha,p+1}$;*
(4) *if $h \in \mathcal{H}^N_{\alpha,p}$ and if P is a polynomial of degree d, then $Ph \in \mathcal{H}^N_{\alpha,p+d}$.*

Proof. Assertion (1) results from the definition. Assertions (2)–(4) are consequences of Lemma 2.2. □

The following result is important and its proof which is rather technical is postponed to Appendix B.

Proposition 2.5. *If $h \in \mathcal{H}^N_{\alpha,p}$, then the solution of Stein's equation satisfies $f_h \in \mathcal{H}^{N+1}_{\alpha,p}$.*

2.3. *Normal expansion*

We consider a family of independent random variables X_i $(i = 1, \dots, n)$ with mean zero and finite variance $\sigma_i^2 > 0$. Let $W = X_1 + \cdots + X_n$ and $\sigma_W^2 = \mathrm{Var}(W)$. Denote by X_i^* a random variable which follows the zero-biased distribution of X_i and is independent of $W^{(i)} := W - X_i$.

Theorem 2.6. *Let $N \geq 0$ be an integer, $\alpha \in (0,1]$ and $p \geq 0$. Assume that $h \in \mathcal{H}_{\alpha,p}^N$. Let $X_1, \dots, X_n$ be independent random variables with mean zero and up to $(N + \max(\alpha + p, 2))$th moments. Then $\mathbb{E}[h(W)]$ can be written as the sum of two terms $C_N(h)$ and $e_N(h)$ where $C_0(h) = \Phi_{\sigma_W}(h)$ and $e_0(h) = \mathbb{E}[h(W)] - \Phi_{\sigma_W}(h)$, and recursively for $N \geq 1$,*

$$C_N(h) = C_0(h) + \sum_{i=1}^n \sigma_i^2 \sum_{d\geq 1} (-1)^{d-1} \sum_{\mathbf{J}\in\mathbb{N}_*^d, |\mathbf{J}|\leq N} m_{X_i}^{(\mathbf{J}^\circ)} \big(m_{X_i^*}^{(\mathrm{J}^\dagger)} - m_{X_i}^{(\mathrm{J}^\dagger)}\big) C_{N-|\mathbf{J}|}\big(f_h^{(|\mathbf{J}|+1)}\big), \tag{11}$$

$$\begin{aligned} e_N(h) = \sum_{i=1}^n \sigma_i^2 \Bigg[& \sum_{d\geq 1} (-1)^{d-1} \sum_{\mathbf{J}\in\mathbb{N}_*^d, |\mathbf{J}|\leq N} m_{X_i}^{(\mathbf{J}^\circ)} \big(m_{X_i^*}^{(\mathrm{J}^\dagger)} - m_{X_i}^{(\mathrm{J}^\dagger)}\big) e_{N-|\mathbf{J}|}\big(f_h^{(|\mathbf{J}|+1)}\big) \\ & + \sum_{k=0}^N \varepsilon_{N-k}\big(f_h^{(k+1)}, W^{(i)}, X_i\big) m_{X_i^*}^{(k)} + \delta_N\big(f_h', W^{(i)}, X_i^*\big) \Bigg], \end{aligned} \tag{12}$$

where for any integer $d \geq 1$, and any $\mathbf{J} \in \mathbb{N}_*^d$, $\mathrm{J}^\dagger \in \mathbb{N}_*$ denotes the last coordinate of $\mathbf{J}$, and $\mathbf{J}^\circ$ denotes the vector in $\mathbb{N}_*^{d-1}$ obtained from $\mathbf{J}$ by omitting the last coordinate.

Remark 2.7. *The moments of X_i^* in (11) and (12) can be written as moments of X_i. By definition of the zero bias transformation, for any $k \in \mathbb{N}$ such that X_i^* has kth moment, one has*

$$\mathbb{E}\big[(X_i^*)^k\big] = \frac{\mathbb{E}[X_i^{k+2}]}{\sigma_i^2(k+1)}.$$

Hence $m_{X_i^*}^{(\mathrm{J}^\dagger)} = \sigma_i^{-2}(\mathrm{J}^\dagger + 2) m_{X_i}^{(\mathrm{J}^\dagger+2)}$ where $\mathrm{J}^\dagger$ is the strictly positive integer figuring in the last coordinate of $\mathbf{J}$.

Proof. We first prove that all terms in (11) and (12) are well defined. When $N = 0$, $h \in \mathcal{H}_{\alpha,p}^0$ and $h(x) = \mathrm{O}(|x|^{\alpha+p})$. Hence $\mathbb{E}[h(W)]$ and $\Phi_{\sigma_W}(h)$ are well defined. Assume that we have proved for $0, \dots, N-1$. Let $h \in \mathcal{H}_{\alpha,p}^N$. Then by Proposition 2.4, $h(x) \in \mathcal{H}_{\alpha,p}^N \subset \mathcal{H}_{\alpha,p+1}^{N-1} \subset \cdots \subset \mathcal{H}_{\alpha,p+N}^0$, so $h(x) = \mathrm{O}(|x|^{\alpha+p+N})$. By Proposition 2.5, $f_h \in \mathcal{H}_{\alpha,p}^{N+1}$ and by Proposition 2.4(1), for any $|\mathbf{J}| = 1, \dots, N$, $f_h^{(|\mathbf{J}|+1)} \in \mathcal{H}_{\alpha,p}^{N-|\mathbf{J}|}$. So the induction hypothesis implies that $C_{N-|\mathbf{J}|}(f_h^{(|\mathbf{J}|+1)})$ and $e_{N-|\mathbf{J}|}(f_h^{(|\mathbf{J}|+1)})$ are well defined. Furthermore, since $f_h^{(k+1)}(x) = \mathrm{O}(|x|^{\alpha+p+N-k})$ for any $k = 0, \dots, N$, the terms ε_{N-k} and δ_N in (12) are well defined. Finally, all moments figuring in (11) and (12) exist.

We then prove the equality $\mathbb{E}[h(W)] = C_N(h) + e_N(h)$. In the case $N = 0$, the equality holds by definition. In the following, we assume that the equality $\mathbb{E}[h(W)] = C_k(h) + e_k(h)$ has been verified for any $k = 0, \dots, N-1$. By Stein's equation (3), $\mathbb{E}[h(W)] - C_0(h)$ is equal to

$$\sigma_W^2 \mathbb{E}\big[f_h'(W^*) - f_h'(W)\big] = \sum_{i=1}^n \sigma_i^2 \big(\mathbb{E}\big[f_h'(W^{(i)} + X_i^*)\big] - \mathbb{E}\big[f_h'(W)\big]\big). \tag{13}$$

Consider the following Taylor expansion

$$\mathbb{E}\big[f_h'(W^{(i)} + X_i^*)\big] = \sum_{k=0}^N m_{X_i^*}^{(k)} \mathbb{E}\big[f_h^{(k+1)}(W^{(i)})\big] + \delta_N\big(f_h', W^{(i)}, X_i^*\big).$$

By replacing $\mathbb{E}[f_h^{(k+1)}(W^{(i)})]$ in the above formula by its $(N-k)$th reverse Taylor expansion, we obtain that $\mathbb{E}[f_h'(W^{(i)}+X_i^*)]$ equals

$$\sum_{k=0}^{N} m_{X_i^*}^{(k)}\biggl[\sum_{d\geq 0}(-1)^d \sum_{\mathbf{J}\in\mathbb{N}_*^d|\mathbf{J}|\leq N-k} m_{X_i}^{(\mathbf{J})}\mathbb{E}\bigl[f_h^{(|\mathbf{J}|+k+1)}(W)\bigr]+\varepsilon_{N-k}\bigl(f_h^{(k+1)},W^{(i)},X_i\bigr)\biggr]+\delta_N\bigl(f_h',W^{(i)},X_i^*\bigr).$$

Note that the term with indexes $k=d=0$ in the sum inside the bracket is $\mathbb{E}[f_h'(W)]$. Therefore $\mathbb{E}[f_h'(W^{(i)}+X_i^*)]-\mathbb{E}[f_h'(W)]$ can be written as the sum of the following three parts

$$\sum_{k=1}^{n} m_{X_i^*}^{(k)}\sum_{d\geq 0}(-1)^d \sum_{\mathbf{J}\in\mathbb{N}_*^d,|\mathbf{J}|\leq N-k} m_{X_i}^{(\mathbf{J})}\mathbb{E}\bigl[f_h^{(|\mathbf{J}|+k+1)}(W)\bigr], \tag{14}$$

$$\sum_{d\geq 1}(-1)^d \sum_{\mathbf{J}\in\mathbb{N}_*^d,|\mathbf{J}|\leq N} m_{X_i}^{(\mathbf{J})}\mathbb{E}\bigl[f_h^{(|\mathbf{J}|+1)}(W)\bigr], \tag{15}$$

$$\sum_{k=0}^{N} m_{X_i^*}^{(k)}\varepsilon_{N-k}\bigl(f_h^{(k+1)},W^{(i)},X_i\bigr)+\delta_N\bigl(f_h',W^{(i)},X_i^*\bigr). \tag{16}$$

By interchanging summations and then making the index changes $\mathbf{K}=(\mathbf{J},k)$ and $u=d+1$, we obtain

$$(14)=\sum_{u\geq 1}(-1)^{u-1}\sum_{\mathbf{K}\in\mathbb{N}_*^u,|\mathbf{K}|\leq N} m_{X_i}^{(\mathbf{K}^\circ)} m_{X_i^*}^{(\mathrm{K}^\dagger)}\mathbb{E}\bigl[f_h^{(|\mathbf{K}|+1)}(W)\bigr].$$

As the equality $m_{X_i}^{(\mathbf{J})}=m_{X_i}^{(\mathbf{J}^\circ)}m_{X_i}^{(\mathrm{J}^\dagger)}$ holds for any $\mathbf{J}$, $(14)+(15)$ equals

$$\sum_{d\geq 1}(-1)^{d-1}\sum_{\mathbf{J}\in\mathbb{N}_*^d,|\mathbf{J}|\leq N} m_{X_i}^{(\mathbf{J}^\circ)}\bigl(m_{X_i^*}^{(\mathrm{J}^\dagger)}-m_{X_i}^{(\mathrm{J}^\dagger)}\bigr)\mathbb{E}\bigl[f_h^{(|\mathbf{J}|+1)}(W)\bigr].$$

By the hypothesis of induction, we have $\mathbb{E}[f_h^{(|\mathbf{J}|+1)}(W)]=C_{N-|\mathbf{J}|}(f_h^{(|\mathbf{J}|+1)})+e_{N-|\mathbf{J}|}(f_h^{(|\mathbf{J}|+1)})$. So the equality $\mathbb{E}[h(W)]=C_N(h)+e_N(h)$ follows. □

Denote by $\mu_i^{(k)}$ the kth moment of X_i. We give the first few explicit expansion formulas below:

$$C_1(h)=\Phi_{\sigma_W}(h)+C_0\bigl(f_h^{(2)}\bigr)\sum_{i=1}^{n}\frac{\mu_i^{(3)}}{2};$$

$$C_2(h)=\Phi_{\sigma_W}(h)+C_1\bigl(f_h^{(2)}\bigr)\sum_{i=1}^{n}\frac{\mu_i^{(3)}}{2}+C_0\bigl(f_h^{(3)}\bigr)\sum_{i=1}^{n}\biggl(\frac{\mu_i^{(4)}}{6}-\frac{(\mu_i^{(2)})^2}{2}\biggr);$$

$$C_3(h)=\Phi_{\sigma_W}(h)+C_2\bigl(f_h^{(2)}\bigr)\sum_{i=1}^{n}\frac{\mu_i^{(3)}}{2}+C_1\bigl(f_h^{(3)}\bigr)\sum_{i=1}^{n}\biggl(\frac{\mu_i^{(4)}}{6}-\frac{(\mu_i^{(2)})^2}{2}\biggr)+C_0\bigl(f_h^{(4)}\bigr)\sum_{i=1}^{n}\biggl(\frac{\mu_i^{(5)}}{24}-\frac{5\mu_i^{(2)}\mu_i^{(3)}}{12}\biggr).$$

Note that in $C_N(h)$, there appear normal expectations of the operators which are of the form $h\mapsto f_h^{(l)}$. As pointed out by Barbour [2], p. 294, such expectations can be expressed as those of h multiplied by a Hermite polynomial, see also remark below.

Remark 2.8. *Let us compare our result with Barbour's. In the main theorem of* [2], *the condition on the function h for the Nth-order expansion is: h is N times differentiable and $\|h^{(N)}\|_{\alpha,p}<\infty$. The condition $h\in\mathcal{H}_{\alpha,p}^N$ in Theorem* 2.6 *is weaker since we separate the continuous and the jump parts of h. An example in $\mathcal{H}_{\alpha,p}^1$ which does not satisfy the* *condition of* [2] *is* $h(x)=(x-k)^+$ *since* $\|h'\|_{\alpha,p}$ *is infinite. Concerning the summand variables, up to* $(N+\max(\alpha+p,2))$*th moments of* X_i*'s are needed for the* N*th expansion and up to* $(N+\alpha+p+2)$*th moments, which are similar to* [2], *are required to achieve the error estimates* (*see Section* 3.2 *for details*). *We consider now the expansion formula. Let* $r\geq 1$ *be an integer. By Remark* 2.7 *and the fact that* X_i*'s are with mean zero, the moment term figuring in* (11) *satisfies*

$$\begin{aligned}
&\sigma_i^2\sum_{d\geq 1}(-1)^{d-1}\sum_{\mathbf{J}\in\mathbb{N}_*^d,|\mathbf{J}|=r} m_{X_i}^{(\mathbf{J}^\circ)}\big(m_{X_i^*}^{(\mathrm{J}^\dagger)}-m_{X_i}^{(\mathrm{J}^\dagger)}\big)\\
&\quad=\sum_{d\geq 1}(-1)^{d-1}\sum_{\mathbf{J}\in\mathbb{N}_*^d,|\mathbf{J}|=r} m_{X_i}^{(\mathbf{J}^\circ)}\big((\mathrm{J}^\dagger+2)m_{X_i}^{(\mathrm{J}^\dagger+2)}-2m_{X_i}^{(\mathrm{J}^\dagger)}m_{X_i}^{(2)}\big)\\
&\quad=\sum_{d\geq 1}(-1)^{d-1}\sum_{\mathbf{J}\in\mathbb{N}_*^d,|\mathbf{J}|=r}\big(\mathrm{J}^\dagger+2\big)m_{X_i}^{(\mathbf{J}^\circ)}m_{X_i}^{(\mathrm{J}^\dagger+2)}-\sum_{d\geq 1}(-1)^{d-1}\sum_{\mathbf{J}\in\mathbb{N}_*^d,|\mathbf{J}|=r}2m_{X_i}^{(\mathbf{J}^\circ)}m_{X_i}^{(\mathrm{J}^\dagger)}m_{X_i}^{(2)}\\
&\quad=\sum_{d\geq 1}(-1)^{d-1}\sum_{\substack{\mathbf{K}\in\mathbb{N}_*^d,|\mathbf{K}|=r+2\\ \mathrm{K}^\dagger\geq 3}}\mathrm{K}^\dagger m_{X_i}^{(\mathbf{K})}+\sum_{d'\geq 1}(-1)^{d'-1}\sum_{\substack{\mathbf{K}\in\mathbb{N}_*^{d'},|\mathbf{K}|=r+2\\ \mathrm{K}^\dagger=2}}\mathrm{K}^\dagger m_{X_i}^{(\mathbf{K})}\\
&\quad=\sum_{\substack{\mathbf{K}\in\bigcup_{d\geq 1}\mathbb{N}_*^d\\ |\mathbf{K}|=r+2}}(-1)^{\ell(\mathbf{K})-1}\mathrm{K}^\dagger m_{X_i}^{(\mathbf{K})},
\end{aligned}$$

where in the third equality, we use the index changes $\mathbf{K}=(\mathbf{J}^\circ,\mathrm{J}^\dagger+2)$ *for the first term and* $\mathbf{K}=(\mathbf{J},2)$, $d'=d+1$ *for the second term, and in the last equality,* $\ell(\mathbf{K})$ *denotes the length of the vector* $\mathbf{K}$. *Denote by* $\kappa_i^{(r)}$ *the* r*th cumulant of* X_i. *By the recurrence formula which relates cumulants and moments* (*e.g.*, [23]):

$$\kappa_i^{(r)}=\mu_i^{(r)}-\sum_{j=1}^{r-1}\binom{r-1}{j}\mu_i^{(j)}\kappa_i^{(r-j)},\tag{17}$$

we obtain

$$\sigma_i^2\sum_{d\geq 1}(-1)^{d-1}\sum_{\mathbf{J}\in\mathbb{N}_*^d,|\mathbf{J}|=r} m_{X_i}^{(\mathbf{J}^\circ)}\big(m_{X_i^*}^{(\mathrm{J}^\dagger)}-m_{X_i}^{(\mathrm{J}^\dagger)}\big)=\frac{\kappa_i^{(r+2)}}{(r+1)!}.\tag{18}$$

In fact, the formula (17) *can also be written as*

$$\frac{\kappa_i^{(r)}}{(r-1)!}=rm_{X_i}^{(r)}-\sum_{j=1}^{r-1}m_{X_i}^{(j)}\frac{\kappa_i^{(r-j)}}{(r-j-1)!}.$$

Denote by $\eta_i^{(r)}$ *the sum* $\sum_{|\mathbf{K}|=r}(-1)^{\ell(\mathbf{K})-1}\mathrm{K}^\dagger m_{X_i}^{(\mathbf{K})}$. *A direct verification shows* $\eta_i^{(1)}=\kappa_i^{(1)}=0$ *and for any* $r\geq 2$,

$$\begin{aligned}
rm_{X_i}^{(r)}-\sum_{j=1}^{r-1}m_{X_i}^{(j)}\eta_i^{(r-j)}&=rm_{X_i}^{(r)}-\sum_{j=1}^{r-1}m_{X_i}^{(j)}\sum_{|\mathbf{K}|=r-j}(-1)^{\ell(\mathbf{K})-1}\mathrm{K}^\dagger m_{X_i}^{(\mathbf{K})}\\
&=rm_{X_i}^{(r)}-\sum_{j=1}^{r-1}\sum_{|\mathbf{K}|=r-j}(-1)^{\ell(\mathbf{K})-1}\mathrm{K}^\dagger m_{X_i}^{(j,\mathbf{K})}\\
&=rm_{X_i}^{(r)}+\sum_{|\mathbf{J}|=r,\ell(\mathbf{J})\geq 2}(-1)^{\ell(\mathbf{J})-1}\mathrm{J}^\dagger m_{X_i}^{(\mathbf{J})}=\eta_i^{(r)},
\end{aligned}$$

where $\mathbf{J}=(j,\mathbf{K})$ *in the third equality. By induction, we obtain* $\eta_i^{(r)}=\kappa_i^{(r)}/(r-1)!$ *for any* $r\in\mathbb{N}_*$ *and hence* (18). *So the expansion formula* (11) *becomes*

$$C_N(h)=C_0(h)+\sum_{r=1}^{N}C_{N-r}\big(f_h^{(r+1)}\big)\frac{\kappa^{(r+2)}(W)}{(r+1)!},\tag{19}$$

where $\kappa^{(r)}(W)$ *denotes the* r*th cumulant of* W. *We thus recover the result of Barbour.*

3. Estimates of remaining errors

3.1. *Concentration inequalities*

We shall prove a concentration inequality which gives an upper bound for $\mathbb{P}(a\le W\le b)$, a and b being two real numbers. The special point here is to consider a parameter $\alpha\le 1$ and to give a bound depending sub-linearly on $b-a$. The case where $\alpha=1$ has been studied in [8,9,11,14]. The following result, measuring the nearness between X and X^*, is useful to prove the concentration inequality.

Lemma 3.1. *Let* $\alpha\in(0,1]$ *and* X *be a r.v. with mean zero, finite variance* $\sigma^2>0$ *and up to* $(\alpha+2)$*th moments. Let* X^* *have the zero biased distribution of* X *and be independent of* X. *Then, for any* $\varepsilon>0$,

$$\mathbb{P}\big(\big|X-X^*\big|>\varepsilon\big)\le\frac{1}{2\varepsilon^\alpha(\alpha+1)\sigma^2}\mathbb{E}\big[\big|X^s\big|^{\alpha+2}\big],$$

where $X^s=X-\widetilde{X}$ *and* $\widetilde{X}$ *is an independent copy of* X.

Proof. Similarly to the Markov inequality, the following inequality holds:

$$\mathbb{P}\big(\big|X-X^*\big|>\varepsilon\big)\le\frac{1}{\varepsilon^\alpha}\mathbb{E}\big[\big|X-X^*\big|^\alpha\big].$$

For any locally integrable even function f, by definition of the zero bias transformation and the fact that X and X^* are independent,

$$\mathbb{E}\big[f\big(X^*-X\big)\big]=\frac{1}{2\sigma^2}\mathbb{E}\big[X^sF\big(X^s\big)\big],$$

where $F(x)=\int_0^x f(t)\,\mathrm{d}t$. In particular,

$$\mathbb{E}\big[\big|X-X^*\big|^\alpha\big]=\frac{1}{2(\alpha+1)\sigma^2}\mathbb{E}\big[\big|X^s\big|^{\alpha+2}\big],$$

which implies the lemma. □

Lemma 3.2. *We keep the notation of Theorem* 2.6 *and suppose that* $f:\mathbb{R}\to\mathbb{R}$ *and* $g:\mathbb{R}^2\to\mathbb{R}$ *are two functions such that the variance of* $f(W)$ *and* $g(X_i,X_i^*)$ *exist for* $i=1,\dots,n$. *Then*

$$\big|\mathrm{Cov}\big[f(W),g\big(X_I,X_I^*\big)\big]\big|\le\frac{\mathrm{Var}[f(W)]^{1/2}}{\sigma_W^2}\left(\sum_{i=1}^{n}\sigma_i^4\,\mathrm{Var}\big[g\big(X_i,X_i^*\big)\big]\right)^{1/2}.\tag{20}$$

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

$$\big|\mathrm{Cov}[\mathbb{1}_{\{a\le W\le b\}},\mathbb{1}_{\{|X_I-X_I^*|\le\varepsilon\}}]\big|\le\frac{1}{4}\left(\sum_{i=1}^{n}\frac{\sigma_i^4}{\sigma_W^4}\right)^{1/2}.\tag{21}$$

Proof. Denote by $(\overrightarrow{X}, \overrightarrow{X}^*) = (X_1, \ldots, X_n, X_1^*, \ldots, X_n^*)$. Note that

$$\mathrm{Cov}\big[f(W), g\big(X_I, X_I^*\big)\big] = \mathrm{Cov}\big[f(W), \mathbb{E}\big[g\big(X_I, X_I^*\big)\big|\big(\overrightarrow{X}, \overrightarrow{X}^*\big)\big]\big].$$

Since (X_i, X_i^*) are mutually independent, we have

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

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

We present below the concentration inequality. One of the referees has pointed out that (22) can also be deduced, up to some constant, from the inequality (5.2) (with a suitable choice of δ) of Chen and Shao [10].

Proposition 3.3. *Let $\alpha \in (0,1]$. For real numbers a, b such that $a \le b$,*

$$\mathbb{P}(a \le W \le b) \le 2\left(\frac{b-a}{2\sigma_W}\right)^{\alpha} + \frac{2}{\alpha+1}\sum_{i=1}^n \mathbb{E}\left[\left|\frac{X_i^s}{\sigma_W}\right|^{\alpha+2}\right] + \frac{1}{2\sigma_W^2}\left(\sum_{i=1}^n \sigma_i^4\right)^{1/2}. \tag{22}$$

Proof. Consider the indicator function $I_{[a,b]}(x) = 1$ if $x \in [a,b]$ and $I_{[a,b]}(x) = 0$ otherwise. Its primitive function $f(x) := \int_{(a+b)/2}^x I_{[a,b]}(t)\,\mathrm{d}t$ satisfies $|f(x)| \le (b-a)/2$. Then

$$\mathbb{E}\big[I_{[a,b]}\big(W^*\big)\big] = \frac{1}{\sigma_W^2}\mathbb{E}\big[Wf(W)\big] \le \min\left(\frac{b-a}{2\sigma_W}, 1\right).$$

Note that for any $u \ge 0$ and any $\alpha \in (0,1]$, $\min(u,1) \le u^\alpha$. Then for any $\varepsilon > 0$,

$$\mathbb{P}\big(a-\varepsilon \le W^* \le b+\varepsilon\big) \le \left(\frac{b-a+2\varepsilon}{2\sigma_W}\right)^{\alpha} \le \left(\frac{b-a}{2\sigma_W}\right)^{\alpha} + \left(\frac{\varepsilon}{\sigma_W}\right)^{\alpha},$$

where the last inequality is because $(u+v)^\alpha \le u^\alpha + v^\alpha$ for any u and v positive. On the other hand, by Lemma 3.2,

$$\begin{aligned}\mathbb{P}\big(a-\varepsilon \le W^* \le b+\varepsilon\big) &\ge \mathbb{P}\big(a \le W \le b, \big|X_I - X_I^*\big| \le \varepsilon\big) \\ &\ge \mathbb{P}(a \le W \le b)\mathbb{P}\big(\big|X_I^* - X_I\big| \le \varepsilon\big) - \frac{1}{4}\left(\sum_{i=1}^n \frac{\sigma_i^4}{\sigma_W^4}\right)^{1/2}.\end{aligned}$$

In this proof, we assume exceptionally that X_i^* is also independent of X_i. By Lemma 3.1,

$$\begin{aligned}\mathbb{P}\big(\big|X_I^* - X_I\big| \le \varepsilon\big) &= 1 - \sum_{i=1}^n \frac{\sigma_i^2}{\sigma_W^2}\mathbb{P}\big(\big|X_i^* - X_i\big| > \varepsilon\big) \\ &\ge 1 - \frac{1}{2\sigma_W^2(\alpha+1)\varepsilon^\alpha}\sum_{i=1}^n \mathbb{E}\big[\big|X_i^s\big|^{\alpha+2}\big].\end{aligned}$$

Finally, the inequality (22) follows by taking

$$\varepsilon = \left(\frac{1}{\sigma_W^2(\alpha+1)}\sum_{i=1}^n \mathbb{E}\big[\big|X_i^s\big|^{\alpha+2}\big]\right)^{1/\alpha}.$$

□

Corollary 3.4. *Let $\alpha \in (0,1]$. For $a, b \in \mathbb{R}$ such that $a \le b$ and $i = 1, \dots, n$,*

$$\mathbb{P}\big(a \le W^{(i)} \le b\big) \le 4\left(\frac{b-a}{2\sigma_W}\right)^{\alpha} + \frac{4}{\alpha+1}\sum_{j=1}^{n}\mathbb{E}\left[\left|\frac{X_j^s}{\sigma_W}\right|^{\alpha+2}\right] + \frac{1}{\sigma_W^2}\left(\sum_{j=1}^{n}\sigma_j{}^4\right)^{1/2} + 4\left(\frac{2\sigma_i}{\sigma_W}\right)^{\alpha}.$$

Proof. Let $\varepsilon > 0$ be a real number, then

$$\mathbb{P}\big(a \le W^{(i)} \le b, |X_i| \le \varepsilon\big) \le \mathbb{P}(a - \varepsilon \le W \le b + \varepsilon).$$

Note that $W^{(i)}$ and X_i are independent and

$$\mathbb{P}\big(|X_i| \le \varepsilon\big) = 1 - \mathbb{P}\big(|X_i| > \varepsilon\big) \ge 1 - \frac{E[|X_i|]}{\varepsilon}.$$

By Proposition 3.3 and taking $\varepsilon = 2\mathbb{E}[|X_i|]$, we obtain the inequality. □

3.2. *Error estimates*

In this section, we shall estimate the error term $e_N(h)$ in Theorem 2.6. The recursive formulas (12) and (8) allow us to reduce the problem to estimating the classical Taylor expansion errors.

For any positive random variable Y and any real number $\beta \ge 0$, we introduce the notation $m_Y^{(\beta)} := \mathbb{E}[Y^\beta]/\Gamma(\beta+1)$, where Γ is the Gamma function

$$\Gamma(x) = \int_0^\infty t^{x-1}\mathrm{e}^{-t}\,\mathrm{d}t.$$

This notation generalizes the one introduced in Section 2.1 since when $\beta \in \mathbb{N}$, $\Gamma(\beta+1) = \beta!$.

Proposition 3.5. *Let $N \ge 0$ be an integer, $\alpha \in (0,1]$ and $p \ge 0$. Suppose that X is a r.v. having up to $(N+\alpha+p)$th moments and satisfying the concentration inequality*

$$\mathbb{P}(a \le X \le b) \le c(b-a)^\alpha + r \quad \forall a, b \in \mathbb{R}, a \le b,$$

where c and r are two constants. Let Y be a r.v. independent of X and having up to $(N+\alpha+p)$th moments. Then, for any function $g \in \mathcal{H}_{\alpha,p}^N$ and any $k = 0, \dots, N$,

$$\begin{aligned}\big|\delta_{N-k}\big(g^{(k)}, X, Y\big)\big| &\le V\big(g_d^{(N)}\big)\big(cm_{|Y|}^{(N-k+\alpha)} + rm_{|Y|}^{(N-k)}\big)\\ &\quad + \big\|g_c^{(N)}\big\|_{\alpha,p}\big(u_{\alpha,p,X}m_{|Y|}^{(N-k+\alpha)} + v_{\alpha,p}m_{|Y|}^{(N-k+\alpha+p)}\big),\end{aligned} \tag{23}$$

where $V(g_d^{(N)})$ denotes the total variation of $g_d^{(N)}$, the coefficients $u_{\alpha,p,X}$ and $v_{\alpha,p}$ are defined by $u_{\alpha,p,X} = (1+(1+2^p)\mathbb{E}[|X|^p])\Gamma(\alpha+1)$ and $v_{\alpha,p} = 2^p\Gamma(\alpha+p+1)$.

Proof. By (6), when $k < N$,

$$\delta_{N-k}\big(g^{(k)}, X, Y\big) = \frac{1}{(N-k-1)!}\int_0^1 (1-t)^{N-k-1}\mathbb{E}\big[\big(g^{(N)}(X+tY) - g^{(N)}(X)\big)Y^{N-k}\big]\,\mathrm{d}t.$$

Since $g \in \mathcal{H}_{\alpha,p}^N$, there exist M, ε_j and K_j $(1 \le j \le M)$ such that the function $g_d^{(N)}$ can be written as

$$g_d^{(N)}(x) = g_d^{(N)}(0) + \sum_{1\le j\le M}\varepsilon_j \mathbb{1}_{x \le K_j} - \sum_{\substack{1\le j\le M\\ K_j \ge 0}}\varepsilon_j.$$

Therefore, $g_d^{(N)}(X+tY)-g_d^{(N)}(X)=\sum_{j=1}^M \varepsilon_j \mathbb{1}_{K_j-tY_+<X\le K_j-tY_-}$, where $Y_+=\max(Y,0)$ and $Y_-=\min(Y,0)$. So the concentration inequality hypothesis implies

$$\mathbb{E}\big[\big|g_d^{(N)}(X+tY)-g_d^{(N)}(X)\big|\big|Y\big]\le\sum_{j=1}^M|\varepsilon_j|\big(ct^\alpha|Y|^\alpha+r\big).$$

Moreover,

$$\int_0^1\frac{(1-t)^{N-k-1}}{(N-k-1)!}\mathbb{E}\Bigg[\sum_{j=1}^M|\varepsilon_j|\big(ct^\alpha|Y|^\alpha+r\big)|Y|^{N-k}\Bigg]\,\mathrm{d}t\le\sum_{j=1}^M|\varepsilon_j|\big(cm_{|Y|}^{(N-k+\alpha)}+rm_{|Y|}^{(N-k)}\big)$$

by using the following equality concerning the Beta function

$$B(x,y):=\int_0^1 t^{x-1}(1-t)^{y-1}\,\mathrm{d}t=\frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)},\quad x,y>0.$$

By definition of the norm $\|\cdot\|_{\alpha,p}$,

$$\begin{aligned}\big|g_c^{(N)}(X+tY)-g_c^{(N)}(X)\big|&\le\big\|g_c^{(N)}\big\|_{\alpha,p}|tY|^\alpha\big(1+|X+tY|^p+|X|^p\big)\\&\le\big\|g_c^{(N)}\big\|_{\alpha,p}|tY|^\alpha\big(1+\big(2^p+1\big)|X|^p+2^p|tY|^p\big),\end{aligned}$$

where the last inequality results from $(a+b)^p\le 2^p(a^p+b^p)$. Note that

$$\begin{aligned}&\int_0^1\frac{(1-t)^{N-k-1}}{(N-k-1)!}\mathbb{E}\big[|tY|^\alpha\big(1+\big(2^p+1\big)|X|^p+2^p|tY|^p\big)|Y|^{N-k}\big]\,\mathrm{d}t\\&\quad=u_{\alpha,p,X}m_{|Y|}^{(N-k+\alpha)}+v_{\alpha,p}m_{|Y|}^{(N-k+\alpha+p)}.\end{aligned}$$

Thus we obtain (23). Finally, it remains to check the case $k=N$. Consider the continuous and the discontinuous parts of $g^{(N)}(X+Y)-g^{(N)}(X)$, respectively. By a similar method as above,

$$\mathbb{E}\big[\big|g_d^{(N)}(X+Y)-g_d^{(N)}(X)\big|\big]\le V\big(g_d^{(N)}\big)\big(cm_{|Y|}^{(1)}+r\big)$$

and

$$\mathbb{E}\big[\big|g_c^{(N)}(X+Y)-g_c^{(N)}(X)\big|\big]\le\big\|g_c^{(N)}\big\|_{\alpha,p}\big(\mathbb{E}\big[|Y|^\alpha\big]\big(1+\big(2^p+1\big)\mathbb{E}\big[|X|^p\big]\big)+2^p\mathbb{E}\big[|Y|^{\alpha+p}\big]\big),$$

which ends the proof. □

By Propositions 2.1 and 3.5, we obtain the error estimates for the reverse Taylor expansion.

Corollary 3.6. *With the previous notation,*

$$\begin{aligned}&\big|\varepsilon_{N-k}\big(g^{(k)},X,Y\big)\big|\\&\quad\le\sum_{d\ge0}\sum_{\mathbf{J}\in\mathbb{N}_*^d,|\mathbf{J}|\le N-k}m_{|Y|}^{(\mathbf{J})}\big[V\big(g_d^{(N)}\big)\big(cm_{|Y|}^{(N-k-|\mathbf{J}|+\alpha)}+rm_{|Y|}^{(N-k-|\mathbf{J}|)}\big)\\&\qquad+\big\|g_c^{(N)}\big\|_{\alpha,p}\big(u_{\alpha,p,X}m_{|Y|}^{(N-k-|\mathbf{J}|+\alpha)}+v_{\alpha,p}m_{|Y|}^{(N-k-|\mathbf{J}|+\alpha+p)}\big)\big].\end{aligned}\tag{24}$$

The above estimates allow us to obtain an upper bound for $e_N(h)$ by using the recursive formula (12) and the concentration inequality in Corollary 3.4. In particular, we give the order estimate of $e_N(h)$ when $X_1,\ldots,X_n$ are in addition i.i.d. random variables.

Proposition 3.7. *Suppose that $X_1, \dots, X_n$ are i.i.d. random variables with mean zero and up to $(N+2+\alpha+p)$th moments and having the same distribution as $X/\sqrt{n}$ where X is a given r.v. with mean zero and finite variance $\sigma^2 > 0$. Then for any $g \in \mathcal{H}^N_{\alpha,p}$ and any $k = 0, \dots, N$,*

$$\delta_{N-k}\big(g^{(k)}, W^{(i)}, X_i^*\big) = \mathrm{O}\left(\left(\frac{1}{\sqrt{n}}\right)^{N-k+\alpha}\right), \tag{25}$$

$$\varepsilon_{N-k}\big(g^{(k)}, W^{(i)}, X_i\big) = \mathrm{O}\left(\left(\frac{1}{\sqrt{n}}\right)^{N-k+\alpha}\right), \tag{26}$$

where the implied constants depend on $V(g_d^{(N)})$, $\|g_c^{(N)}\|_{\alpha,p}$ and up to $(N-k+2+\alpha+p)$th moments of X.

Proof. By Corollary 3.4, for $a \le b$ and $\alpha \in (0,1]$,

$$\mathbb{P}\big(a \le W^{(i)} \le b\big) \le c(b-a)^\alpha + r(n),$$

where the coefficients are given by

$$c = \frac{2^{2-\alpha}}{\sigma^\alpha}, \qquad r(n) = \frac{4}{\sigma^{2+\alpha}(\alpha+1)} \frac{\mathbb{E}[|X^s|^{\alpha+2}]}{\sqrt{n^\alpha}} + \frac{1}{\sqrt{n}} + \frac{2^{2+\alpha}}{\sqrt{n^\alpha}}.$$

By Proposition 3.5, we obtain an upper bound of $\delta_{N-k}(g^{(k)}, W^{(i)}, X_i^*)$ which is determined by a linear combination of terms (with coefficient not depending on n):

$$m_{|X_i^*|}^{(N-k+\alpha)}, \qquad r(n) m_{|X_i^*|}^{(N-k)}, \qquad \mathbb{E}\big[\big|W^{(i)}\big|^p\big] m_{|X_i^*|}^{(N-k+\alpha)} \quad \text{and} \quad m_{|X_i^*|}^{(N-k+\alpha+p)}. \tag{27}$$

Note that $r(n) = \mathrm{O}((1/\sqrt{n})^\alpha)$. For any $k = 0, \dots, N$, $\mathbb{E}[|X_i^*|^k]$ equals $\mathbb{E}[|X_i|^{k+2}]/(\sigma_i^2(k+1))$ and is of order $(1/\sqrt{n})^k$. So the first three terms in (27) are of order $(1/\sqrt{n})^{N-k+\alpha}$ and the last term is order $(1/\sqrt{n})^{N-k+\alpha+p}$, which implies (25). Then (26) follows similarly by Corollary 3.6. □

Proposition 3.8. *Let $N \ge 0$ be an integer, $\alpha \in (0,1]$ and $p \ge 0$. Let $h \in \mathcal{H}^N_{\alpha,p}$. Assume that $X_1, \dots, X_n$ are i.i.d. random variables having up to $(N+2+\alpha+p)$th moments and the same distribution as $X/\sqrt{n}$, X being a given r.v. with mean zero and finite variance. Then*

$$e_N(h) = \mathrm{O}\left(\left(\frac{1}{\sqrt{n}}\right)^{N+\alpha}\right), \tag{28}$$

where the implied constant depends on up to $(N+2+\alpha+p)$th moments of X.

Proof. We prove the proposition by induction on N. When $N = 0$,

$$e_0(h) = \sum_{i=1}^n \sigma_i^2 \big(\delta_0\big(f_h', W^{(i)}, X_i^*\big) + \varepsilon_0\big(f_h', W^{(i)}, X_i\big)\big).$$

Since $h \in \mathcal{H}^0_{\alpha,p}$, $f_h \in \mathcal{H}^1_{\alpha,p}$. Then by Proposition 3.7, $e_0(h) = \mathrm{O}((1/\sqrt{n})^\alpha)$. Assume that we have already proved the proposition for $0, \dots, N-1$. Consider $h \in \mathcal{H}^N_{\alpha,p}$ and $e_N(h)$ defined in (12). For any $\mathbf{J}$ such that $1 \le |\mathbf{J}| \le N$, $e_{N-|\mathbf{J}|} = \mathrm{O}((1/\sqrt{n})^{N-|\mathbf{J}|+\alpha})$. In addition, since $|\mathbf{J}^\circ| + |\mathbf{J}^\dagger| = |\mathbf{J}|$, $m_{X_i}^{(\mathbf{J}^\circ)}(m_{X_i^*}^{(\mathbf{J}^\dagger)} - m_{X_i}^{(\mathbf{J}^\dagger)})$ is of order $(1/\sqrt{n})^{|\mathbf{J}|}$. Since $f_h \in \mathcal{H}^{N+1}_{\alpha,p}$, $f_h' \in \mathcal{H}^N_{\alpha,p}$, so $\delta_N(f_h', W^{(i)}, X_i^*) = \mathrm{O}((1/\sqrt{n})^{N+\alpha})$. For any $k = 0, \dots, N$, $f_h^{(k+1)} \in \mathcal{H}^{N-k}_{\alpha,p}$, so $\varepsilon_{N-k}(f_h^{(k+1)}, W^{(i)}, X_i^*) = \mathrm{O}((1/\sqrt{n})^{N-k+\alpha})$. Finally we have $m_{X_i^*}^{(k)} = \mathrm{O}((1/\sqrt{n})^k)$. The proposition is hence proved. □

In Barbour [2], the Nth-order error estimates depend also on the $(N+\alpha)$th moments of the summand variables. However, one needs more regularity on the function h. Consider now an example. Let $I_k(x) = \mathbb{1}_{\{x\leq k\}}$ be the indicator function. Then $I_k \in \mathcal{H}^0_{\alpha,0}$. By Proposition 3.8, if $X_1, \ldots, X_n$ are i.i.d. random variables having $(2+\alpha)$th moment, then $e_0(h) = \mathrm{O}((1/\sqrt{n})^\alpha)$ and the coefficient depends on up to $(2+\alpha)$th moments of X_i's. This recovers the error estimates of [21], Theorem V.3.6.

4. The Poisson case

In this section, we show that the same method can be applied to obtain a Poisson expansion formula. We note in the first place that with a slight abuse of notation, we shall use the similar notation in this section as in the previous ones for the normal case. However, the meaning is different since we here are concerned with discrete random variables.

4.1. *Some reminders*

Stein's method for Poisson approximation has been introduced by Chen [7], and then developed in a series of papers such as [1,4,5] among many others, one can also consult the monograph [6] and the survey paper [12]. In particular, Barbour [3] has developed, in parallel with the normal case [2], asymptotic expansions for the sum of independent integer-valued random variables and for polynomially growing functions. The extension of the zero bias transformation in [15] for the normal case was extended in [16] to apply to distributions more generally; in particular, the transformation considered here for the Poisson is essentially the simplest special case of biasing with respect to the Charlier polynomials. To distinguish it from the normal version, we refer to it as the Poisson zero bias transformation.

Let Z be an $\mathbb{N}$-valued random variable ($\mathbb{N}$-r.v.), then Z follows the Poisson distribution with parameter λ if and only if the equality $\mathbb{E}[Zf(Z)] = \lambda\mathbb{E}[f(Z+1)]$ holds for any function $f:\mathbb{N}\to\mathbb{R}$ such that both sides of the equality are well defined. Based on this observation, Chen has proposed the following discrete Stein's equation:

$$xf(x) - \lambda f(x+1) = h(x) - \mathcal{P}_\lambda(h), \quad x \in \mathbb{N}, \tag{29}$$

where $\mathcal{P}_\lambda(h)$ denotes the expectation of h under the λ-Poisson distribution. If X is an $\mathbb{N}$-r.v. with mean λ, then $\mathbb{E}[h(X)] - \mathcal{P}_\lambda(h) = \mathbb{E}[Xf_h(X) - f_h(X+1)]$ where f_h is the solution of (29) given by

$$f_h(x) = \frac{(x-1)!}{\lambda^x}\sum_{i=x}^{\infty}\frac{\lambda^i}{i!}\big(h(i) - \mathcal{P}_\lambda(h)\big). \tag{30}$$

The value $f_h(0)$ can be arbitrary and is not used in calculations in general.

We say that an $\mathbb{N}$-r.v. X^* has Poisson X-zero biased distribution if the equality

$$\mathbb{E}\big[Xf(X)\big] = \lambda\mathbb{E}\big[f\big(X^*+1\big)\big] \tag{31}$$

holds for any function $f:\mathbb{N}\to\mathbb{R}$ such that the left-hand side of (31) is well defined. The distribution of X^* is unique: one has $\mathbb{P}(X^* = x) = (x+1)\mathbb{P}(X = x+1)/\lambda$. By Stein's equation (29) and the zero bias transformation (31),

$$\mathbb{E}\big[h(X)\big] - \mathcal{P}_\lambda(h) = \lambda\mathbb{E}\big[f_h\big(X^*+1\big) - f_h(X+1)\big]. \tag{32}$$

Recall the Newton's expansion [4], Theorem 5.1. For all $x, y \in \mathbb{N}$ and $N \in \mathbb{N}$,

$$f(x+y) = \sum_{j=0}^{N}\binom{y}{j}\Delta^j f(x) + \sum_{0\leq j_1<\cdots<j_{N+1}<y}\Delta^{N+1}f(x+j_1),$$

where $\Delta f(x) = f(x+1) - f(x)$. Let us introduce, for any $\mathbb{N}$-r.v. Y and any $k \in \mathbb{N}$ such that $\mathbb{E}[|Y|^k] < +\infty$, the following quantity

$$m_Y^{(k)} := \mathbb{E}\bigg[\binom{Y}{k}\bigg].$$

By (31), if Y^* has kth moment and $\mathbb{E}[Y]=\lambda$, then $\mathbb{E}[(Y^*)^k]=\frac{1}{\lambda}\mathbb{E}[Y(Y-1)^k]$ and

$$m_{Y^*}^{(k)}=\frac{k+1}{\lambda}m_Y^{(k+1)}.$$

Let X and Y be two independent $\mathbb{N}$-r.v.'s and $f:\mathbb{N}\to\mathbb{R}$ such that $\Delta^k f(X)$ and $\Delta^k f(X+Y)$ are integrable, then

$$\mathbb{E}\big[f(X+Y)\big]=\sum_{k=0}^{N}m_Y^{(k)}\mathbb{E}\big[\Delta^k f(X)\big]+\delta_N(f,X,Y),$$

where

$$\delta_N(f,X,Y)=\mathbb{E}\Bigg[\sum_{0\le j_1<\cdots<j_{N+1}<Y}\Delta^{N+1}f(X+j_1)\Bigg]. \tag{33}$$

Remark 4.1. *With the above notation, we can verify that the reverse Taylor formula still holds in this discrete setting if we replace $f^{(|\mathbf{J}|)}$ by $\Delta^{|\mathbf{J}|}f$ in Proposition 2.1, the proof is the same.*

Consider a family of independent $\mathbb{N}$-r.v.'s $X_i (i=1,\dots,n)$ with mean λ_i. Let $W=X_1+\cdots+X_n$ and $\lambda_W:=\mathbb{E}[W]=\lambda_1+\cdots+\lambda_n$. Let $W^{(i)}=W-X_i$ and X_i^* be an $\mathbb{N}$-r.v. having the Poisson X_i-zero biased distribution and independent of $W^{(i)}$. Finally, let I be a random index valued in $\{1,\dots,n\}$ and independent of $(X_1,\dots,X_n,X_1^*,\dots,X_n^*)$ such that $\mathbb{P}(I=i)=\lambda_i/\lambda_W$. Then, similar as in [15] and [16], the random variable $W^*:=W^{(I)}+X_I^*$ follows the Poisson W-zero biased distribution.

4.2. *Poisson expansion*

We shall give the asymptotic expansion of $\mathbb{E}[h(W)]$ in terms of Poisson expectations and we begin by specifying the conditions on the function h. Compared to the normal case, the differentiability conditions on h are no more needed and we shall concentrate on its increasing speed at infinity. This makes the study much simpler.

Let $\mathcal{H}_p$, $p\ge 0$ be the space of functions $h:\mathbb{N}\to\mathbb{R}$ such that $h(x)=\mathrm{O}(x^p)$ when $x\to\infty$. Then it's easy to verify that $\Delta h\in\mathcal{H}_p$ and that f_h is well defined. The following result can be compared to Proposition 2.5.

Proposition 4.2. *Let $p\ge 0$. If $h\in\mathcal{H}_p$, then $f_h\in\mathcal{H}_{p-1}$.*

Proof. Without loss of generality, we consider a function h with Poisson mean zero and the corresponding Stein's equation on $\mathbb{N}_*$:

$$xf(x)-\lambda f(x+1)=h(x), \tag{34}$$

whose particular solution is given by

$$f_h(x):=\frac{(x-1)!}{\lambda^x}\sum_{i=x}^{\infty}\frac{\lambda^i}{i!}h(i).$$

A general solution of (34) can be written as $f_h(x)+C(x-1)!/\lambda^x$, where C is an arbitrary constant. However, if $h\in\mathcal{H}_p$, f_h is the only solution of (34) which has polynomial increasing speed at infinity. We shall prove $f_h(x)=\mathrm{O}(x^{p-1})$.

Notice that

$$0\le\frac{x!}{\lambda^x}\sum_{i\ge x}\frac{\lambda^i}{i!}=\sum_{i\ge x}\frac{\lambda^{i-x}}{i!/x!}\le\sum_{i\ge x}\frac{\lambda^{i-x}}{(i-x)!}=\mathrm{e}^{-\lambda}.$$

Therefore, when $p \le 0$, one has

$$x f_h(x) = \frac{x!}{\lambda^x} \sum_{i=x}^{\infty} \frac{\lambda^i}{i!} h(i) = \mathrm{O}\big(x^p\big)$$

since $i^p \le x^p$ if $x \le i$. Hence $f_h(x) = \mathrm{O}(x^{p-1})$. The general case follows by induction on p by using the equality

$$f_{\tau(h)}(x) = f_h(x+1)/x,$$

where $\tau(h) : \mathbb{N} \to \mathbb{R}$ is defined by $\tau(h)(x) = h(x+1)/(x+1)$. □

The asymptotic expansion formula in the Poisson case is similar to the one given in Theorem 2.6.

Theorem 4.3. *Let $N \in \mathbb{N}$ and $p \ge 0$. Assume that $h \in \mathcal{H}_p$ and that X_i $(i = 1, \dots, n)$ are independent $\mathbb{N}$-r.v.'s with mean λ_i and up to $(N + \max(p, 1))$th moments. Let $W = X_1 + \cdots + X_n$ and $\lambda_W = \mathbb{E}[W]$. Then $\mathbb{E}[h(W)]$ can be written as the sum of $C_N(h)$ and $e_N(h)$ such that $C_0(h) = \mathcal{P}_{\lambda_W}(h)$ and $e_0(h) = \mathbb{E}[h(W)] - \mathcal{P}_{\lambda_W}(h)$, and recursively for any $N \ge 1$,*

$$C_N(h) = C_0(h) + \sum_{i=1}^{n} \lambda_i \sum_{d \ge 1} (-1)^{d-1} \sum_{\mathbf{J} \in \mathbb{N}_*^d, |\mathbf{J}| \le N} m_{X_i}^{(\mathbf{J}^\circ)} \big(m_{X_i^*}^{(\mathrm{J}^\dagger)} - m_{X_i}^{(\mathrm{J}^\dagger)}\big) C_{N-|\mathbf{J}|}\big(\Delta^{|\mathbf{J}|} f_h(x+1)\big), \tag{35}$$

$$\begin{aligned} e_N(h) = \sum_{i=1}^{n} \lambda_i \Bigg[&\sum_{d \ge 1} (-1)^{d-1} \sum_{\mathbf{J} \in \mathbb{N}_*^d, |\mathbf{J}| \le N} m_{X_i}^{(\mathbf{J}^\circ)} \big(m_{X_i^*}^{(\mathrm{J}^\dagger)} - m_{X_i}^{(\mathrm{J}^\dagger)}\big) e_{N-|\mathbf{J}|}\big(\Delta^{|\mathbf{J}|} f_h(x+1)\big) \\ &+ \sum_{k=0}^{N} m_{X_i^*}^{(k)} \varepsilon_{N-k}\big(\Delta^k f_h(x+1), W^{(i)}, X_i\big) + \delta_N\big(f_h(x+1), W^{(i)}, X_i^*\big) \Bigg], \end{aligned} \tag{36}$$

where for any integer $d \ge 1$ and any $\mathbf{J} \in \mathbb{N}_*^d$, $\mathrm{J}^\dagger \in \mathbb{N}_*$ denotes the last coordinate of $\mathbf{J}$, and $\mathbf{J}^\circ$ denotes the element in $\mathbb{N}_*^{d-1}$ obtained from $\mathbf{J}$ by omitting the last coordinate.

Proof. We shall verify that all terms in (35) and (36) are well defined. When $N = 0$, X_i has $\max(p, 1)$th moment and $h = \mathrm{O}(x^p)$, so $\mathcal{P}_{\lambda_W}(h)$ exists. Suppose that the assertion holds for $1, \dots, N-1$. Consider now the case of N. By Proposition 4.2, $f_h \in \mathcal{H}_{p-1}$, so $\Delta^{|\mathbf{J}|}(f_h(x+1)) \in \mathcal{H}_{p-1}$. Then $C_{N-|\mathbf{J}|}(\Delta^{|\mathbf{J}|} f_h(x+1))$ and $e_{N-|\mathbf{J}|}(\Delta^{|\mathbf{J}|} f_h(x+1))$ are well defined. In addition, all the moment terms exist. The proof of the equality $\mathbb{E}[h(W)] = C_N(h) + e_N(h)$ is similar to the second part of the proof of Theorem 2.6, which we omit. □

Remark 4.4. *We have $m_{X_i^*}^{(\mathrm{J}^\dagger)} = \lambda_i^{-1}(\mathrm{J}^\dagger + 1) m_{X_i}^{(\mathrm{J}^\dagger + 1)}$ where $\mathrm{J}^\dagger$ is the strictly positive integer figuring in the last coordinate of $\mathbf{J}$. By an argument similar to Remark 2.8,*

$$\lambda_i \sum_{d \ge 1} (-1)^{d-1} \sum_{\mathbf{J} \in \mathbb{N}_*^d, |\mathbf{J}| = r} m_{X_i}^{(\mathbf{J}^\circ)} \big(m_{X_i^*}^{(\mathrm{J}^\dagger)} - m_{X_i}^{(\mathrm{J}^\dagger)}\big) = \sum_{\substack{\mathbf{K} \in \bigcup_{d \ge 1} \mathbb{N}_*^d \\ |\mathbf{K}| = r+1}} (-1)^{\ell(\mathbf{K})-1} \mathrm{K}^\dagger m_{X_i}^{(\mathbf{K})} =: \eta_i^{(r+1)}.$$

One has $\eta_i^{(1)} = \lambda_i$ and $\eta_i^{(r)} = r m_{X_i}^{(r)} - \sum_{j=1}^{r-1} m_{X_i}^{(j)} \eta_i^{(r-j)}$ for any $r \ge 2$. Note that $\eta_i^{(r)}/r$ is just the coefficient of t^r in the series $\log \mathbb{E}[(1+t)^{X_i}]$ and $\sum_i \eta_i^{(r)}/r$ is the coefficient of t^r in the series $\log \mathbb{E}[(1+t)^W]$. We hence recover the result of Barbour [3].

4.3. *Error estimates*

We finally concentrate on the remaining term $e_N(h)$ in (36). The following quantity will be useful:

$$\|h\|_{N,p} := \sup_{x\in\mathbb{N}_*} \frac{|\Delta^{N+1}h(x)|}{x^p}, \tag{37}$$

which is finite for any $h \in \mathcal{H}_p$.

Lemma 4.5. *Let $N \in \mathbb{N}$, $k = 0, \dots, N$ and $p \geq 0$. Let $f \in \mathcal{H}_p$, X be an $\mathbb{N}$-r.v. with pth moment, Y be an $\mathbb{N}$-r.v. independent of X and having $(N-k+1+p)$th moment. Then*

$$\begin{aligned}&\big|\delta_{N-k}\big(\Delta^k f(x+1), X, Y\big)\big| \\ &\quad\leq \max\big(2^{p-1}, 1\big)\|f\|_{N,p}\big(\mathbb{E}\big[X^p\big]m_Y^{(N-k+1)} + m_Y^{(N-k+1),p}\big),\end{aligned} \tag{38}$$

where

$$m_Y^{(N-k+1),p} := \mathbb{E}\bigg[\binom{Y}{N-k+1}Y^p\bigg],$$

and

$$\begin{aligned}&\big|\varepsilon_{N-k}\big(\Delta^k f(x+1), X, Y\big)\big| \\ &\quad\leq \max\big(2^{p-1}, 1\big)\|f\|_{N,p}\sum_{d\geq 0}\sum_{\substack{\mathbf{J}\in\mathbb{N}_*^d\\ |\mathbf{J}|\leq N-k}} m_Y^{(\mathbf{J})}\big(\mathbb{E}\big[X^p\big]m_Y^{(N-k-|\mathbf{J}|+1)} + m_Y^{(N-k-|\mathbf{J}|+1),p}\big).\end{aligned} \tag{39}$$

Proof. By definitions (33) and (37),

$$\begin{aligned}&\big|\delta_{N-k}\big(\Delta^k f(x+1), X, Y\big)\big| \\ &\quad\leq \mathbb{E}\bigg[\sum_{0\leq j_1<\cdots<j_{N-k+1}<Y}\big|\Delta^{N+1}f(X+1+j_1)\big|\bigg] \\ &\quad\leq \|f\|_{N,p}\mathbb{E}\bigg[\sum_{0\leq j_1<\cdots<j_{N-k+1}<Y}(X+j_1+1)^p\bigg] \\ &\quad\leq \|f\|_{N,p}\mathbb{E}\bigg[\binom{Y}{N-k+1}(X+Y)^p\bigg] \\ &\quad\leq \max\big(2^{p-1}, 1\big)\|f\|_{N,p}\big(\mathbb{E}\big[X^p\big]m_Y^{(N-k+1)} + m_Y^{(N-k+1),p}\big),\end{aligned}$$

where we have used in the last inequality the estimations $(X+Y)^p \leq 2^{p-1}(X^p + Y^p)$ if $p > 1$ and $(X+Y)^p \leq X^p + Y^p$ if $p \leq 1$. Thus (38) is proved. The inequality (39) follows from (8) and (38). $\square$

Proposition 4.6. *Let $N \in \mathbb{N}$, $p \geq 0$ and $h \in \mathcal{H}_p$. Suppose that X_i $(i = 1, \dots, n)$ are independent $\mathbb{N}$-r.v.'s with mean λ_i and up to $(N+p+2)$th moments. Let $W = X_1 + \cdots + X_n$ and X_i^* be an $\mathbb{N}$-r.v. having Poisson X_i-zero biased distribution and independent of $W^{(i)} := W - X_i$. Then*

$$\big|e_0(h)\big| \leq \max\big(2^{p-1}, 1\big)\|f_h\|_{0,p}\sum_{i=1}^n \lambda_i\big(\mathbb{E}\big[\big(W^{(i)}\big)^p\big]\big(m_{X_i^*}^{(1)} + m_{X_i}^{(1)}\big) + \big(m_{X_i^*}^{(1),p} + m_{X_i}^{(1),p}\big)\big). \tag{40}$$

When $N \geq 1$,

$$\left|e_N(h)\right| \leq \sum_{i=1}^{n} \lambda_i \Bigg[\sum_{d\geq 1} \sum_{\substack{\mathbf{J}\in\mathbb{N}_*^d \\ |\mathbf{J}|\leq N}} m_{X_i}^{(\mathbf{J}^\circ)} \big(m_{X_i^*}^{(\mathbf{J}^\dagger)} + m_{X_i}^{(\mathbf{J}^\dagger)} \big) \big| e_{N-|\mathbf{J}|}\big(\Delta^{|\mathbf{J}|} f_h(x+1)\big) \big|$$

$$+ \max\big(2^{p-1}, 1\big) \| f_h \|_{N,p} \sum_{k=0}^{N} m_{X_i^*}^{(k)} \sum_{d\geq 0} \sum_{\substack{\mathbf{J}\in\mathbb{N}_*^d \\ |\mathbf{J}|\leq N-k}} m_{X_i}^{(\mathbf{J})} \big(\mathbb{E}\big[\big(W^{(i)}\big)^p \big] m_{X_i}^{(N-k-|\mathbf{J}|+1)} + m_{X_i}^{(N-k-|\mathbf{J}|+1),p} \big)$$

$$+ \max\big(2^{p-1}, 1\big) \| f_h \|_{N,p} \big(\mathbb{E}\big[\big(W^{(i)}\big)^p \big] m_{X_i^*}^{(N+1)} + m_{X_i^*}^{(N+1),p} \big) \Bigg]. \quad (41)$$

Proof. We begin by the case when $N = 0$. By (32),

$$\begin{aligned} e_0(h) &= \sum_{i=1}^{n} \lambda_i \big(\mathbb{E}\big[f_h\big(W^{(i)} + X_i^* + 1\big) \big] - \mathbb{E}\big[f_h\big(W^{(i)} + X_i + 1\big) \big] \big) \\ &= \sum_{i=1}^{n} \lambda_i \big(\delta_0\big(f_h(x+1), W^{(i)}, X_i^*\big) + \varepsilon_0\big(f_h(x+1), W^{(i)}, X_i\big) \big) \\ &\leq \max\big(2^{p-1}, 1\big) \| f_h \|_{0,p} \sum_{i=1}^{n} \lambda_i \big\{ \mathbb{E}\big[\big(W^{(i)}\big)^p \big] \big(m_{X_i^*}^{(1)} + m_{X_i}^{(1)} \big) + \big(m_{X_i^*}^{(1),p} + m_{X_i}^{(1),p} \big) \big\}, \end{aligned}$$

where the last inequality is by Lemma 4.5, so (40) follows. We obtain the estimates for $N \geq 1$ by applying the recursive formula (36). □

Proposition 4.7. *Let $N \in \mathbb{N}$, $p \geq 0$ and $h \in \mathcal{H}_p$. Suppose that X_i $(i = 1, \ldots, n)$ are i.i.d. $\mathbb{N}$-r.v.'s having the same distribution as X/n where X is a given $\mathbb{N}$-r.v. with up to $(N + p + 2)$th moments. Then $e_N(h)$ defined in* (36) *satisfies*

$$e_N(h) = \mathrm{O}\left(\left(\frac{1}{n} \right)^{N+1} \right), \quad (42)$$

where the implied constant depends on up to $(N + p + 2)$th moments of X.

Proof. For any $k = 0, \ldots, N + 1$ and any $i = 1, \ldots, n$,

$$m_{X_i}^{(k)} \leq \frac{1}{k!} \mathbb{E}\big[X_i^k\big] = \mathrm{O}\big(n^{-k}\big) \quad \text{and} \quad m_{X_i^*}^{(k)} = \frac{k+1}{\lambda_i} m_{X_i}^{(k+1)} = \mathrm{O}\big(n^{-k}\big).$$

Similarly $m_{X_i}^{(k),p} = \mathrm{O}(n^{-k-p})$ and $m_{X_i^*}^{(k),p} = \mathrm{O}(n^{-k-p})$, where the implied constants depend on up to $(N + p + 2)$th moments of X. Hence by induction on N from (40) and (41), we obtain the estimate (42). □

Appendix A: Proof of Lemma 2.2

Proof. For the first two assertions, it suffices to prove respectively the boundless of the following two functions

$$\frac{1 + |x|^p + |y|^p}{1 + |x|^q + |y|^q}, \qquad |x - y|^{\beta-\alpha} \frac{1 + |x|^p + |y|^p}{1 + |x|^{p+\beta-\alpha} + |y|^{p+\beta-\alpha}}.$$

These functions are both continuous on $\mathbb{R}^2$, therefore are bounded on any compact subset of $\mathbb{R}^2$. Thus we may assume without loss of generality that $r = \sqrt{x^2+y^2} \geq 1$. In this case, $\max\{|x|,|y|\} \geq r/\sqrt{2}$, so

$$\frac{1+|x|^p+|y|^p}{1+|x|^q+|y|^q} \leq \frac{1+2r^p}{1+(r/\sqrt{2})^q} \leq 3\cdot 2^{q/2},$$

$$|x-y|^{\beta-\alpha}\frac{1+|x|^p+|y|^p}{1+|x|^{p+\beta-\alpha}+|y|^{p+\beta-\alpha}} \leq (2r)^{\beta-\alpha}\frac{1+2r^p}{1+(r/\sqrt{2})^{p+\beta-\alpha}} \leq 3\cdot 2^{(p+3\beta-3\alpha)/2},$$

thus proving (1) and (2). For (3), we have

$$\frac{|P(x)f(x)-P(y)f(y)|}{|x-y|^\alpha(1+|x|^{p+d}+|y|^{p+d})} \leq \frac{(1+|x|^p+|y|^p)P(x)}{1+|x|^{p+d}+|y|^{p+d}}\|f\|_{\alpha,p} + \frac{|f(y)|\cdot|P(x)-P(y)|}{|x-y|^\alpha(1+|x|^{p+d}+|y|^{p+d})}.$$

By the argument in the proof of (1) and (2), the first term in the right-hand side is bounded. Since P is a polynomial of degree d, there exists a polynomial $Q(x,y)$ in two variables and of degree $d-1$, such that $Q(x,y) = (P(x)-P(y))/(x-y)$. Therefore, the second term equals

$$\frac{|Q(x,y)|\cdot|x-y|^{1-\alpha}\cdot|f(y)|}{1+|x|^{p+d}+|y|^{p+d}}$$

which is bounded by a similar argument as for proving (1) and (2).

(4) Since $\|f\|_{\alpha,p} < +\infty$, $|f(t)| \ll 1+|t|^{\alpha+p}$ for $t \in \mathbb{R}$ where $\ll$ is the Vinogradov's symbol.[1] Therefore, for $x, y \in \mathbb{R}$, $x \leq y$,

$$\big|F(x)-F(y)\big| \leq \int_x^y \big|f(t)\big|\,\mathrm{d}t \ll \int_x^y \big(1+|t|^{\alpha+p}\big)\,\mathrm{d}t \leq \big(1+|x|^{\alpha+p}+|y|^{\alpha+p}\big)|y-x|.$$

Hence

$$\frac{|F(x)-F(y)|}{|x-y|(1+|x|^{p+\alpha}+|y|^{p+\alpha})}$$

is bounded. □

Appendix B: Proof of Proposition 2.5

We now prove Proposition 2.5. Let $h \in \mathcal{H}^N_{\alpha,p}$. Then the function f_h is $N+1$ times differentiable by (4). Taking the Nth derivative on both sides of Stein's equation (2), we get

$$\big(xf_h(x)\big)^{(N)} - \sigma^2 f_h^{(N+1)}(x) = h^{(N)}(x). \tag{43}$$

The function $(xf_h(x))^{(N)}$ is continuous, so $f_h^{(N+1)}$ is locally of finite variation and has finitely many jumps as $h^{(N)}$ does. It remains to prove $\|f_{h,c}^{(N+1)}\|_{\alpha,p} < +\infty$.

Let A be an interval in $\mathbb{R}$ and f be a Borel function on A. We define for any $\alpha \in (0,1]$ and any $p \geq 0$,

$$\|f\|^A_{\alpha,p} := \sup_{\substack{x\neq y\\ x,y\in A}} \frac{|f(x)-f(y)|}{|x-y|^\alpha(1+|x|^p+|y|^p)}, \tag{44}$$

which is analogous to (10) restricted to an interval A.

[1] Let F and G be two real functions defined on a set Ω, and Ω_0 be a subset of Ω. The expression "$F(x) \ll G(x)$ for $x \in \Omega_0$" signifies that there exists a constant $C > 0$ (which may depend on Ω_0) such that $|F(x)| \leq C|G(x)|$ holds for any $x \in \Omega_0$.

Lemma B.1. *If $h \in \mathcal{H}^N_{\alpha,p}$, then $\|f^{(N+1)}_{h,c}\|^A_{\alpha,p} < +\infty$ for any bounded interval A.*

Proof. Firstly, for any bounded interval A and any Borel function g, $\|g\|^A_{\alpha,p} < +\infty$ if and only if g is α-Lipschitz on A. We examine $f^{(N+1)}_{h,c}$ using (43). Since $h \in \mathcal{H}^N_{\alpha,p}$, $h^{(N)}_c$ is locally α-Lipschitz. The function $(xf_h(x))^{(N+1)} = xf^{(N+1)}_h(x) + (N+1)f^{(N)}_h(x)$ has finitely many jumps. Hence $(xf_h(x))^{(N)}$ is a primitive function of a locally bounded function, thus is locally 1-Lipschitz. So $f^{(N+1)}_{h,c}$ is locally α-Lipschitz, which implies the lemma. □

Lemma B.2. *If $A = A_1 \cup A_2$ where A_1 and A_2 are two intervals such that $A_1 \cap A_2 \neq \emptyset$, then*

$$\sup\left\{\|f\|^{A_1}_{\alpha,p}, \|f\|^{A_2}_{\alpha,p}\right\} \le \|f\|^A_{\alpha,p} \le 2\left(\|f\|^{A_1}_{\alpha,p} + \|f\|^{A_2}_{\alpha,p}\right). \tag{45}$$

Proof. The first inequality of (45) is obvious. For the second one, we only need to prove for any $x \in A_1$ and any $y \in A_2$ that

$$\frac{|f(x) - f(y)|}{|x-y|^\alpha(1 + |x|^p + |y|^p)} \le 2\left(\|f\|^{A_1}_{\alpha,p} + \|f\|^{A_2}_{\alpha,p}\right).$$

Without loss of generality, we may suppose that $A_1 \cap A_2$ contains a single point z. Then $|f(x) - f(y)| \le |f(x) - f(z)| + |f(y) - f(z)|$. In addition, since z is between x and y, we have $|x - y| \ge \max(|x - z|, |y - z|)$ and $|x|^p + |y|^p \ge \frac{1}{2}\max(|x|^p + |z|^p, |y|^p + |z|^p)$. So

$$\frac{|f(x) - f(y)|}{|x-y|^\alpha(1 + |x|^p + |y|^p)} \le 2\left(\frac{|f(x) - f(z)|}{|x-z|^\alpha(1 + |x|^p + |z|^p)} + \frac{|f(z) - f(y)|}{|z-y|^\alpha(1 + |z|^p + |y|^p)}\right),$$

which implies the second inequality. □

Combining Lemmas B.1 and B.2, we obtain the following proposition.

Proposition B.3. *Let $A = \mathbb{R} \setminus (-1, 1)$. Then $\|f^{(N+1)}_{h,c}\|_{\alpha,p} < \infty$ if $\|f^{(N+1)}_{h,c}\|^A_{\alpha,p} < \infty$.*

When A avoids an open neighborhood of 0, the finiteness of $\|f\|^A_{\alpha,p}$ is equivalent to that of

$$\sup_{\substack{x \neq y \\ x,y \in A}} \frac{|f(x) - f(y)|}{|x-y|^\alpha(|x|^p + |y|^p)}.$$

This property does not hold for the norm $\|\cdot\|_{\alpha,p}$ defined in (10). As a consequence, we have the following result.

Lemma B.4. *Let $A \subset (-\infty, -1] \cup [1, +\infty)$ be an interval, $\alpha \in (0, 1]$ and $p \ge 0$. Let q be a real number such that $0 \le q \le p$. Then for any Borel function f defined on A, $\|f\|^A_{\alpha,p} < +\infty$ if and only if $\|f(x)/x^{p-q}\|^A_{\alpha,q} < +\infty$.*

Proof. If $\|f(x)/x^{p-q}\|^A_{\alpha,q} < +\infty$, then by similar arguments as for proving Lemma 2.2, we have $\|f\|^A_{\alpha,p} < +\infty$. We now consider the converse assertion. Firstly, there exists a constant $C > 0$ such that $|f(x)| \leqslant C|x|^{\alpha+p}$ for any $x \in A$. For any $x, y \in A$, $|x| < |y|$,

$$\frac{|f(x)x^{q-p} - f(y)y^{q-p}|}{|x-y|^\alpha(1 + |x|^q + |y|^q)} \le \left|f(x)\right|\frac{|x^{q-p} - y^{q-p}|}{|x-y|^\alpha(1 + |x|^q + |y|^q)} + |y|^{q-p}\frac{|f(x) - f(y)|}{|x-y|^\alpha(1 + |x|^q + |y|^q)}.$$

The second term is finite since

$$\frac{|f(x) - f(y)|}{|x-y|^\alpha(|y|^{p-q} + |x|^q|y|^{p-q} + |y|^p)} \le \frac{|f(x) - f(y)|}{|x-y|^\alpha(1 + |x|^p + |y|^p)} = \|f\|^A_{\alpha,p}.$$

By the mean value theorem, the first term is bounded by

$$C|x|^{\alpha+p}\frac{|x-y|\cdot|q-p|\cdot|x|^{q-p-1}}{|x-y|^{\alpha}(1+|x|^q+|y|^q)}$$

and thus by $C|q-p|$ if we assume in addition that $|y|<2|x|$. When $|y|\geqslant 2|x|$, one has

$$\big|f(x)\big|\frac{|x^{q-p}-y^{q-p}|}{|x-y|^{\alpha}(1+|x|^q+|y|^q)}\leqslant C\frac{|x|^{\alpha+p}|x|^{q-p}}{|x|^{\alpha+q}}\leqslant C.$$

□

We shall study $f_{h,c}^{(N+1)}$ on the set $\mathbb{R}\setminus(-1,1)$. Without loss of generality, we consider in the following a function h with normal mean zero and the corresponding Stein's equation

$$xf_h(x)-\sigma^2 f_h'(x)=h(x). \tag{46}$$

The particular solution of (46) is the restriction of the standard solution (4) to the complement of $(-1,1)$ and is given by

$$f_h(x):=\begin{cases}\frac{1}{\sigma^2\phi_\sigma(x)}\int_x^\infty h(t)\phi_\sigma(t)\,\mathrm{d}t, & x\geq 1,\\ \frac{1}{\sigma^2\phi_\sigma(x)}\int_{-\infty}^x h(t)\phi_\sigma(t)\,\mathrm{d}t, & x\leq -1.\end{cases} \tag{47}$$

For any $\sigma>0$, let $\mathcal{E}_\sigma$ be the space of all Borel functions h on $\mathbb{R}\setminus(-1,1)$ such that $\int_{|x|\geq 1}|h(x)P(x)|\phi_\sigma(x)\,\mathrm{d}x<\infty$ for any polynomial P. The vector space $\mathcal{E}_\sigma$ contains all Laurent polynomials (that is, polynomials in x and x^{-1}) and is stable by multiplication by Laurent polynomials. Furthermore, as shown by the lemma below, it is invariant by the operator $h\to f_h$.

Lemma B.5. *Let $h\in\mathcal{E}_\sigma$. Then f_h is well defined and $f_h\in\mathcal{E}_\sigma$. Furthermore, if H is a primitive function of h, then $H\in\mathcal{E}_\sigma$.*

Proof. Let P be an arbitrary polynomial on $\mathbb{R}$. Then

$$\begin{aligned}\int_1^\infty\big|P(x)f_h(x)\big|\phi_\sigma(x)\,\mathrm{d}x&\leq\frac{1}{\sigma^2}\int_1^\infty\mathrm{d}x\,\big|P(x)\big|\int_x^\infty\big|h(t)\big|\phi_\sigma(t)\,\mathrm{d}t\\&=\frac{1}{\sigma^2}\int_1^\infty\mathrm{d}t\,\big|h(t)\big|\phi_\sigma(t)\int_1^t\big|P(x)\big|\,\mathrm{d}x.\end{aligned}$$

There exists a polynomial Q such that $\int_1^t|P(x)|\,\mathrm{d}x\leq Q(t)$ for any $t\geq 0$. Therefore, the fact that $h\in\mathcal{E}_\sigma$ implies that $\int_1^\infty|P(x)f_h(x)|\phi_\sigma(x)\,\mathrm{d}x<+\infty$. The finiteness of the integral on $(-\infty,1]$ is similar. The second assertion can be proved by integration by part. □

If $h\in\mathcal{E}_\sigma$, then f_h is the only solution of (46) in $\mathcal{E}_\sigma$. For the derivatives of f_h, we consider, for any integer $N\geq 1$, the set $\mathcal{E}_\sigma^N$ which contains all functions h such that h is N times differentiable on $\mathbb{R}\setminus(-1,1)$ and that $h^{(N)}\in\mathcal{E}_\sigma$. It is not difficult to observe that $h\in\mathcal{E}_\sigma^N$ if and only if it is a primitive function of an element in $\mathcal{E}_\sigma^{N-1}$. The relationship between $\mathcal{E}_\sigma^N$ and $\mathcal{H}_{\alpha,p}^N$ is as follows.

Lemma B.6. *If $h\in\mathcal{H}_{\alpha,p}^N$, then the restriction of h on $\mathbb{R}\setminus(-1,1)$ is in $\mathcal{E}_\sigma^N$.*

Proof. It suffices to show that the restriction of $h^{(N)}$ on $\mathbb{R}\setminus(-1,1)$ is in $\mathcal{E}_\sigma$. This is obvious since $h_c^{(N)}$ has at most polynomial increasing speed at infinity. □

For any differentiable function h on $\mathbb{R} \setminus (-1, 1)$, define

$$\Lambda(h)(x) := \left(\frac{h(x)}{x}\right)'. \tag{48}$$

Lemma B.7. *If $h \in \mathcal{E}_\sigma^1$, then $\Lambda(h) \in \mathcal{E}_\sigma$. Furthermore,*

$$f_h'(x) = x f_{\Lambda(h)}(x). \tag{49}$$

Proof. By definition, $\Lambda(h)(x) = h'(x)/x - h(x)/x^2$, so $\Lambda(h) \in \mathcal{E}_\sigma$. To prove the equality, it suffices to verify that the function $u(x) := x^{-1} f_h'(x)$ satisfies the Eq. (46) for $\Lambda(h)$. In fact, if we divide both sides of the equation $x f_h(x) - \sigma^2 f_h'(x) = h(x)$ by x and then take the derivative, we obtain $x u(x) - \sigma^2 u'(x) = \Lambda(h)(x)$. □

Lemma B.8. *If $h \in \mathcal{E}_\sigma$ and $h(x) = \mathrm{O}(|x|^l)$ where l is a real number, then $f_h(x) = \mathrm{O}(|x|^{l-1})$.*

Proof. Recall that [11], Lemma 1, if $|h(x)| \le g(x)$ and if $g(x)/|x|$ is decreasing when $x > 0$ and is increasing when $x < 0$, then $|f_h(x)| \le g(x)/|x|$. Hence, we prove the lemma for the cases where $l < 1$. By Lemma B.7, one has

$$\begin{aligned} f_{|x|^l}(x) &= x^{-1}\big(|x|^l + \sigma^2 f_{|x|^l}'(x)\big) = \mathrm{sgn}(x)|x|^{l-1} + \sigma^2 f_{\Lambda(|x|^l)}(x) \\ &= \mathrm{sgn}(x)|x|^{l-1} + \sigma^2 (l-1) f_{|x|^{l-2}}(x). \end{aligned}$$

Thus, $f_{|x|^{l-2}} = \mathrm{O}(|x|^{l-3})$ implies $f_{|x|^l} = \mathrm{O}(|x|^{l-1})$. Hence by induction on l, we obtain the result. □

Remark B.9. *With the notation of Barbour [2], the equivalent expectation form of f_h is given by*

$$f_h(x) = \begin{cases} \frac{\sqrt{2\pi}}{\sigma}\mathbb{E}\big[h(Z+x)\mathrm{e}^{-Zx/\sigma^2}\mathbb{1}_{\{Z>0\}}\big], & x \ge 1, \\ -\frac{\sqrt{2\pi}}{\sigma}\mathbb{E}\big[h(Z+x)\mathrm{e}^{-Zx/\sigma^2}\mathbb{1}_{\{Z<0\}}\big], & x \le -1, \end{cases} \tag{50}$$

where $Z \sim N(0, \sigma^2)$. So Lemma B.8 implies that the function

$$\frac{1}{x^l}\mathbb{E}\big[\mathbb{1}_{\{Z>0\}}(Z+x)^{l+1}\mathrm{e}^{-Zx/\sigma^2}\big]$$

is bounded on $[1, +\infty)$. Then, for all $l \in \mathbb{R}$ and $m \in \mathbb{R}_+$,

$$\frac{1}{x^l}\mathbb{E}\big[\mathbb{1}_{\{Z>0\}}(Zx)^m(Z+x)^{l+1}\mathrm{e}^{-Zx/\sigma^2}\big]$$

is also bounded on $[1, +\infty)$ by the fact that $u^m \mathrm{e}^{-u/(2\sigma^2)}$ is bounded on $[0, \infty)$.

The following lemma shows the relationship between the derivatives of f_h and those of h. The proof is by induction, which we omit in this article (interest readers may refer to [19], pp. 144–145). We only remind that the first formula is a generalization of (49).

Lemma B.10. *If $h \in \mathcal{E}_\sigma^N$ with N being a strictly positive integer, then*

$$f_h^{(N)}(x) = \sum_{k=0}^{\lfloor N/2 \rfloor} \binom{N}{2k}(2k-1)!!\,x^{N-2k} f_{\Lambda^{N-k}(h)}(x); \tag{51}$$

$$\Lambda^N(h)(x) = \sum_{k=0}^{N} (-1)^k (2k-1)!! \binom{N+k}{2k} \frac{h^{(N-k)}(x)}{x^{N+k}}, \tag{52}$$

where we use the convention $(-1)!! = 1$ and $\lfloor N/2 \rfloor$ denotes the largest integer not exceeding $N/2$.

Remark B.11.

(1) *For $h \in \mathcal{E}_\sigma^N$, Lemma B.10 also holds for $f_h^{(m)}(x)$ and $\Lambda^m(h)$ where $1 \le m \le N$. Since the operator $h \to f_h$ is linear on h, we can write the derivatives of f_h as a linear combination of those of h with Laurent polynomial coefficients. This will allow us to deduce the increasing speed at infinity.*

(2) *The derivative $f_h^{(N+1)}$ has to be treated differently. In fact, we can no longer apply (52) to $\Lambda^{N+1}(h)$ since $h^{(N+1)}$ does not necessarily exist. The idea is to separate the first term where $k = 0$ in (51) from the others and then take the derivative*

$$\begin{aligned}
f_h^{(N+1)} &= x^N f'_{\Lambda^N(h)}(x) + \sum_{k=0}^{\lfloor N/2\rfloor} \binom{N}{2k}(2k-1)!!(N-2k)x^{N-2k-1} f_{\Lambda^{N-k}(h)}(x) \\
&\quad + \sum_{k=1}^{\lfloor N/2\rfloor} \binom{N}{2k}(2k-1)!!x^{N-2k} f'_{\Lambda^{N-k}(h)}(x) \\
&= x^N f'_{\Lambda^N(h)}(x) + \sum_{k=1}^{\lfloor (N+1)/2\rfloor} \binom{N+1}{2k}(2k-1)!!x^{N+1-2k} f_{\Lambda^{N+1-k}(h)}(x).
\end{aligned}$$

This is the key point in the proof of Proposition B.13.

Lemma B.12. *Let h be a Borel function on $\mathbb{R}\setminus(-1,1)$ and let $A = (-\infty,-1]$ or $[1,\infty)$. If $\|h\|_{\alpha,p}^A < +\infty$, then, for any $n \in \mathbb{N}$,*

$$\left\|x^{n+1} f_{h/x^n}\right\|_{\alpha,p}^A < +\infty, \qquad \left\|x^n f'_{h/x^n}\right\|_{\alpha,p}^A < +\infty.$$

Proof. We only prove for $A = [1,\infty)$ and the case for $(-\infty,-1]$ is by symmetry. Let $g(x) = h(x)/x^n$. Assume that x and y are two real numbers such that $1 \leqslant x < y$. Then one has

$$\begin{aligned}
&\frac{|x^{n+1} f_g(x) - y^{n+1} f_g(y)|}{|x-y|^\alpha(1+|x|^p+|y|^p)} \\
&\quad = \frac{\sqrt{2\pi}}{\sigma}\mathbb{E}\left[I_{\{Z>0\}}\frac{|(h(Z+x)x^{n+1}/(Z+x)^n)\mathrm{e}^{-Zx/\sigma^2} - (h(Z+y)y^{n+1}/(Z+y)^n)\mathrm{e}^{-Zy/\sigma^2}|}{|x-y|^\alpha(1+|x|^p+|y|^p)}\right],
\end{aligned}$$

which can be bounded from above by the sum of the following two terms:

$$\frac{\sqrt{2\pi}}{\sigma}\mathbb{E}\left[I_{\{Z>0\}}\frac{|h(Z+x)-h(Z+y)|}{|x-y|^\alpha(1+|x|^p+|y|^p)}\cdot\frac{y^{n+1}}{(Z+y)^n}\mathrm{e}^{-Zy/\sigma^2}\right], \tag{53}$$

$$\frac{\sqrt{2\pi}}{\sigma}\mathbb{E}\left[I_{\{Z>0\}}\big|h(Z+x)\big|\frac{|(x^{n+1}/(Z+x)^n)\mathrm{e}^{-Zx/\sigma^2} - (y^{n+1}/(Z+y)^n)\mathrm{e}^{-Zy/\sigma^2}|}{|x-y|^\alpha(1+|x|^p+|y|^p)}\right]. \tag{54}$$

Note that (53) is bounded from above by

$$\|h\|_{\alpha,p}^A\frac{\sqrt{2\pi}}{\sigma}y^{n+1}\mathbb{E}\left[I_{\{Z>0\}}\frac{1}{(Z+y)^n}\mathrm{e}^{-Zy/\sigma^2}\right] = \|h\|_{\alpha,p}^A y^{n+1} f_{1/|x|^n}(y).$$

By Lemma B.8, this quantity is bounded. We then consider the upper bound of (54) under thesupplementary condition $y \leqslant 2x$. As $\|h\|_{\alpha,p}^A < +\infty$, there exists a constant $C > 0$ such that $h(x) \leqslant C|x|^{\alpha+p}$. By applying the mean value theorem on the function $\frac{x^{n+1}}{(Z+x)^n}\mathrm{e}^{-Zx/\sigma^2}$ and the fact that $|x-y|^{\alpha-1}(1+|x|^p+|y|^p)\geq |x|^{\alpha-1+p}$ where $\alpha<1$,

$$
\begin{aligned}
(54) &\leqslant \frac{\sqrt{2\pi}}{\sigma}|x|^{1-\alpha-p}\mathbb{E}\bigg[I_{\{Z>0\}}C(Z+x)^{\alpha+p}\mathrm{e}^{-Zx/\sigma^2}\bigg(\frac{(n+1)(2x)^n}{(Z+x)^n}+\frac{n(2x)^{n+1}}{(Z+x)^{n+1}}+\frac{Z(2x)^{n+1}}{\sigma^2(Z+x)^n}\bigg)\bigg]\\
&\leqslant C\frac{\sqrt{2\pi}}{\sigma}|x|^{1-\alpha-p}\bigg\{\big(2^n\cdot(n+1)+2^{n+1}\cdot n\big)\mathbb{E}\big[I_{\{Z>0\}}(Z+x)^{\alpha+p}\mathrm{e}^{-Zx/\sigma^2}\big]\\
&\quad +2^{n+1}\mathbb{E}\bigg[I_{\{Z>0\}}\frac{Zx}{\sigma^2}(Z+x)^{\alpha+p}\mathrm{e}^{-Zx/\sigma^2}\bigg]\bigg\}\\
&\ll |x|^{1-\alpha-p}\bigg\{E\big[I_{\{Z>0\}}(Z+x)^{\alpha+p}\mathrm{e}^{-Zx/\sigma^2}\big]+\mathbb{E}\bigg[I_{\{Z>0\}}\frac{Zx}{\sigma^2}(Z+x)^{\alpha+p}\mathrm{e}^{-Zx/\sigma^2}\bigg]\bigg\},
\end{aligned}
$$

which is bounded (see Remark B.9). In the case where $y>2x$, one has $x^{n+1}/(Z+x)^n\leq x$ when $Z\geq 0$, and $|x-y|^\alpha(1+|x|^p+|y|^p)\geq \max(|x|^{\alpha+p},(\frac{y}{2})^\alpha\cdot y^p)$, so

$$
(54)\leqslant C\frac{\sqrt{2\pi}}{\sigma}\bigg\{\frac{1}{2|x|^{\alpha+p-1}}\mathbb{E}\big[I_{\{Z>0\}}(Z+x)^{\alpha+p}\mathrm{e}^{-Zx/\sigma^2}\big]+\frac{2^\alpha}{|y|^{\alpha+p-1}}\mathbb{E}\big[I_{\{Z>0\}}(Z+y)^{\alpha+p}\mathrm{e}^{-Zy/\sigma^2}\big]\bigg\},
$$

which is bounded. □

We now summarize the above arguments and the following proposition allows to conclude the proof of Proposition 2.5.

Proposition B.13. *Let $N\in\mathbb{N}$, $\alpha\in(0,1]$ and $p\geqslant 0$. Let h be an N times differentiable function defined on $\mathbb{R}\setminus(-1,1)$ such that $h^{(N)}$ is locally of finite variation, having finitely many jumps and verifying $\|h_c^{(N)}\|_{\alpha,p}^A<+\infty$ with $A=(-\infty,-1]$ or $[1,\infty)$. Then f_h is $N+1$ times differentiable, $f_h^{(N+1)}$ is locally of finite variation, having finitely many jumps and verifying $\|f_{h,c}^{(N+1)}\|_{\alpha,p}^A<+\infty$.*

Proof. The function f_h is $N+1$ times differentiable by (43), $f_h^{(N+1)}$ is locally of finite variation, having only finitely many jumps. By Lemmas B.1 and B.2, it suffices to prove

$$
\max\big\{\|f_{h,c}^{(N+1)}\|_{\alpha,p}^{(-\infty,-b]},\|f_{h,c}^{(N+1)}\|_{\alpha,p}^{[b,+\infty)}\big\}<+\infty
$$

for sufficiently positive number b. Therefore, without loss of generality, we may assume that $h^{(N)}$ is continuous and hence $f_h^{(N+1)}$ is also continuous.

By Remark B.11(2), the function $f_h^{(N+1)}$ can be written as a linear combination of $x^N f'_{\Lambda^N(h)}$ and terms of the form $x^{N+1-2k}f_{\Lambda^{N+1-k}(h)}(x)$ where $k=1,\ldots,\lfloor\frac{N+1}{2}\rfloor$. By (52), $x^N f'_{\Lambda^N(h)}$ itself is also a linear combination of $x^N f'_{h^{(N-i)}/x^{N+i}}$ where $i=0,\ldots,N$. As $\|h^{(N)}\|_{\alpha,p}^A<\infty$, we have, similar as in Lemma 2.2(4), that $\|h^{(N-i)}\|_{\alpha,p+i}^A<+\infty$. Hence $\|h^{(N-i)}/x^i\|_{\alpha,p}^A<\infty$ by Lemma B.4 and Lemma B.12 then implies that $\|x^N f'_{h^{(N-i)}/x^{N+i}}\|_{\alpha,p}^A<\infty$.

The terms $x^{N+1-2k}f_{\Lambda^{N+1-k}(h)}(x)$ are also, by (52) again, linear combinations of the functions of the form $x^{N+1-2k}f_{h^{(N+1-k-i)}/x^{N+1-k+i}}$. By a similar argument as above using Lemma B.4, $\|h^{(N+1-k-i)}/x^{1+k+i}\|_{\alpha,p}^A<\infty$. Finally, we apply Lemma B.12 to obtain $\|x^{N+1-2k}f_{h^{(N+1-k-i)}/x^{N+1-k+i}}\|_{\alpha,p}^A<\infty$, which completes the proof. □

Acknowledgments

The author thanks two anonymous referees for helpful comments and suggestions.

References

[1] R. Arratia, L. Goldstein and L. Gordon. Two moments suffice for Poisson approximations: The Chen–Stein method. *Ann. Probab.* **17** (1989) 9–25. MR0972770

[2] A. D. Barbour. Asymptotic expansions based on smooth functions in the central limit theorem. *Probab. Theory Related Fields* **72** (1986) 289–303. MR0836279

[3] A. D. Barbour. Asymptotic expansions in the Poisson limit theorem. *Ann. Probab.* **15** (1987) 748–766. MR0885141

[4] A. D. Barbour and V. Čekanavičius. Total variation asymptotics for sums of independent integer random variables. *Ann. Probab.* **30** (2002) 509–545. MR1905850

[5] A. D. Barbour, L. H. Y. Chen and K. P. Choi. Poisson approximation for unbounded functions. I. Independent summands. *Statist. Sinica* **5** (1995) 749–766. MR1347617

[6] A. D. Barbour, L. Holst and S. Janson. *Poisson Approximation*. Oxford Univ. Press, Oxford, 1992. MR1163825

[7] L. H. Y. Chen. Poisson approximation for dependent trials. *Ann. Probab.* **3** (1975) 534–545. MR0428387

[8] L. H. Y. Chen and Q.-M. Shao. A non-uniform Berry–Esseen bound via Stein's method. *Probab. Theory Related Fields* **120** (2001) 236–254. MR1841329

[9] L. H. Y. Chen and Q.-M. Shao. Stein's method for normal approximation. In *An Introduction to Stein's Method* 1–59. *Lecture Notes Series, IMS, National University of Singapore* **4**. Singapore Univ. Press, Singapore, 2005. MR2235448

[10] L. H. Y. Chen and Q.-M. Shao. Normal approximation for nonlinear statistics using a concentration inequality approach. *Bernoulli* **13** (2007) 581–599. MR2331265

[11] N. El Karoui and Y. Jiao. Stein's method and zero bias transformation for CDOs tranches pricing. *Finance Stoch.* **13** (2009) 151–180. MR2482050

[12] T. Erhardsson. Stein's method for Poisson and compound Poisson approximation. In *An Introduction to Stein's Method* 61–113. *Lecture Notes Series, IMS, National University of Singapore* **4**. Singapore Univ. Press, Singapore, 2005. MR2235449

[13] L. Goldstein. L^1 bounds in normal approximation. *Ann. Probab.* **35** (2007) 1888–1930. MR2349578

[14] L. Goldstein. Bounds on the constant in the mean central limit theorem. *Ann. Probab.* **38** (2010) 1672–1689. MR2663641

[15] L. Goldstein and G. Reinert. Stein's method and the zero bias transformation with application to simple random sampling. *Ann. Appl. Probab.* **7** (1997) 935–952. MR1484792

[16] L. Goldstein and G. Reinert. Distributional transformations, orthogonal polynomials, and stein characterizations. *J. Theoret. Probab.* **18** (2005) 237–260. MR2132278

[17] F. Götze and C. Hipp. Asymptotic expansions in the central limit theorem under moment conditions. *Z. Wahrsch. Verw. Gebiete* **42** (1978) 67–87. MR0467882

[18] C. Hipp. Edgeworth expansions for integrals of smooth functions. *Ann. Probab.* **5** (1977) 1004–1011. MR0455076

[19] Y. Jiao. Risque de crédit: modélisation et simulation numérique. PhD thesis, Ecole Polytechnique, 2006. Available at http://www.imprimerie.polytechnique.fr/Theses/Files/Ying.pdf.

[20] A. Kolmogolov and S. Fomine. *Éléments de la théorie des fonctions et de l'analyse fonctionnelle*. Éditions Mir., Moscow, 1974. MR0367598

[21] V. V. Petrov. *Sums of Independent Random Variables*. Springer, New York, 1975. MR0388499

[22] V. Rotar. Stein's method, Edgeworth's expansions and a formula of Barbour. In *Stein's Method and Applications* 59–84. *Lect. Notes Ser. Inst. Math. Sci. Natl. Univ. Singap.* **5**. Singapore Univ. Press, Singapore, 2005. MR2201886

[23] P. Smith. A recursive formulation of the old problem of obtaining moments from cumulants and vice versa. *Amer. Statist.* **49** (1995) 217–218. MR1347727

[24] C. Stein. A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. In *Proc. Sixth Berkeley Symp. Math. Statist. Probab.* 583–602. California Univ. Press, Berkeley, 1972. MR0402873

[25] C. Stein. *Approximate Computation of Expectations*. IMS, Hayward, CA, 1986. MR0882007