Wasserstein distance estimates for stochastic integrals by forward-backward stochastic calculus

Jean-Christophe Breton*
Univ Rennes
CNRS, IRMAR - UMR 6625
263 Avenue du Général Leclerc
F-35000 Rennes, France

Nicolas Privault†
Division of Mathematical Sciences
School of Physical and Mathematical Sciences
Nanyang Technological University
21 Nanyang Link, Singapore 637371

August 7, 2020

Abstract

We prove Wasserstein distance bounds between the probability distributions of stochastic integrals with jumps, based on the integrands appearing in their stochastic integral representations. Our approach does not rely on the Stein equation or on the propagation of convexity property for Markovian semigroups, and makes use instead of forward-backward stochastic calculus arguments. This allows us to consider a large class of target distributions constructed using Brownian stochastic integrals and pure jump martingales, which can be specialized to infinitely divisible target distributions with finite Lévy measure and Gaussian components.

Keywords: Wasserstein distance; stochastic integrals; forward-backward stochastic calculus; point processes.

Mathematics Subject Classification (2010): 60H05, 60H10, 60G57, 60G44, 60J60, 60J75.

1 Introduction

Comparison inequalities for option prices with convex payoff functions have been obtained in the literature, based on the classical Kolmogorov equation under the propagation of convexity hypothesis for Markovian semigroups. See for instance [EKJS98]

*jean-christophe.breton@univ-rennes1.fr
†nprivault@ntu.edu.sg

in the case of continuous diffusion processes, and [BJ00], [BR06], [ET07], in the case of jump-diffusion processes. In [BP08], lower and upper bounds on option prices have been obtained in one-dimensional jump-diffusion markets with point process components under different conditions.

Note however that the propagation of convexity property is not always satisfied, even in the (Markovian) jump-diffusion case, see e.g. Theorem 4.4 in [ET07]. Using different arguments based on forward-backward stochastic calculus, related convex ordering results have been obtained for exponential jump-diffusion processes in [BP08]. The case of random vectors admitting a predictable representation in terms of a Brownian motion and a non-necessarily independent jump component has also been treated in [ABP08] using forward-backward stochastic calculus, extending the one-dimensional framework of [KMP06], see also [BLP13] for the case of Itô integrals and [MP13] for additive functionals.

In [BP20b], bounds on differences in expectation have been obtained in order to estimate the distance between the distribution $\mathscr{L}(X_T)$ of the terminal value X_T of a stochastic integral process $(X_t)_{t\in[0,T]}$ on a finite time horizon $[0,T]$ and a target distribution $\mathscr{L}(Y_T)$ given by the terminal value Y_T of a jump-diffusion process $(Y_t)_{t\in[0,T]}$ solution of a stochastic differential equation (SDE). The main idea consists in expanding the difference $h(F)-h(G)$ for suitable functions $h:\mathbb{R}\to\mathbb{R}$ by use of the Itô formula and, after taking expectations, to bound the remaining terms via an adequate control of the characteristics of the related jump-diffusion processes, see e.g. [BP20a] and references therein.

In this paper we apply a different approach based on forward-backward stochastic calculus, see Theorems 3.1 and 3.3 below, from which we derive bounds on the Wasserstein distance between stochastic integrals with jumps. In contrast to [BP20b], this approach also allows us to provide distance bounds between the distribution of a pure point process stochastic integral and a Brownian stochastic integral, see Corollary 5.3. Note that convergence in the Wasserstein distance implies convergence in distribution.

Note that another fruitful approach to obtain Wasserstein bounds is the Stein method, for a short presentation see [ABD+20] and the references therein. However this method applies for some fixed target distribution and relies on the so-called Stein equation depending on this target distribution, for instance see [NP12] for the normal distribution. In contrast, our approach applies to stochastic integrals with jumps whose distributions possibly escape the scope of the Stein method. As an example, consider $(B_t)_{t\in[0,T]}$ a standard Brownian motion, $(Z_t)_{t\in[0,T]}$ a pure-jump martingale, and $(\sigma_t)_{t\in[0,T]}$ an adapted process with respect to the filtration generated by $(B_t)_{t\in[0,T]}$. In Corollary 5.1, given $(f(t))_{t\in[0,T]}$ a deterministic function we consider the Wasserstein distance $d_{\mathrm{W}}(F,G)$ between the sum

$$F = \int_0^T \sigma_t dB_t + Z_T$$

and the mixture

$$G = \int_0^T f(t) dB_t + N_T$$

of a centered Gaussian $\mathcal{N}\left(0, \int_0^T |f(t)|^2 dt\right)$ random variable and the terminal value of a (compensated) compound Poisson process $(N_t)_{t\in[0,T]}$ with deterministic intensity $\mu(t,dx)dt$. From Corollary 4.1 we show that $d_{\mathrm{W}}(F,G)$ can be bounded as

$$d_{\mathrm{W}}(F\,G) \leq \begin{cases} \sqrt{(4+\mathbb{E}[|F|]+\mathbb{E}[|G|])}\left(\mathbb{E}\left[\int_0^T d_{\mathrm{TV}}\big(\widetilde{\nu}_t, \widetilde{\mu}(t,\cdot)\big)dt\right]\right)^{1/2}, & \text{(1.1a)} \\ \frac{\sqrt[3]{9}}{2}(4+\mathbb{E}[|F|]+\mathbb{E}[|G|])^{2/3}\left(\mathbb{E}\left[\int_0^T d_{\mathrm{W}}\big(\widetilde{\nu}_t, \widetilde{\mu}(t,\cdot)\big)\ dt\right]\right)^{1/3}, & \text{(1.1b)} \\ C\ \left(\mathbb{E}\left[\int_0^T d_{\mathrm{FM}}\big(\widetilde{\nu}_t, \widetilde{\mu}(t,\cdot)\big)\ dt\right]\right)^{1/3}, & \text{(1.1c)} \end{cases}$$

for some finite constant $C > 0$. In these bounds, d_{TV} and d_{FM} denote the total variation and Fortet-Mourier distances, and we let

$$\widetilde{\nu}_t(dx) := |\sigma_t|^2\delta_0(dx) + |x|^2\nu_t(dx), \quad \widetilde{\mu}(t,dx) := |f(t)|^2\delta_0(dx) + |x|^2\mu(t,dx),$$

$t \in [0,T]$, where $\nu_t(dx)dt$ and $\mu(t,dx)dt$ denote the compensators of the pure-jump martingales $(Z_t)_{t\in[0,T]}$ and $(N_t)_{t\in[0,T]}$, see Section 2 for detailed definitions. Due to the inequality $d_{\mathrm{FM}}(\cdot,\cdot) \leq d_{\mathrm{W}}(\cdot,\cdot)$ the bound (1.1c) is better than (1.1b), however it holds when $\mathbb{E}\left[\int_0^T d_{\mathrm{FM}}\big(\widetilde{\nu}_t, \widetilde{\mu}(t,\cdot)\big)\ dt\right]$ is small enough, while (1.1b) has an explicit constant.

Such estimates are then specialized to stochastic integrals of the form

$$F := \int_0^T \sigma_t dB_t + Z_T = \int_0^T \sigma_t dB_t + \int_0^T J_t\big(dY_t - \lambda_t dt\big),$$

where $(Y_t)_{t\in[0,T]}$ is a point process with jumps of size 1 and compensator $(\lambda_t)_{t\in[0,T]}$, and N_T in G is a centered Poissonian random variable with parameter μT, $\mu > 0$. In Corollary 5.2 we show that $d_{\rm W}(F,G)$ in (1.1a)-(1.1c) can be controlled via

$$\begin{aligned}\mathbb{E}\left[\int_0^T d_{\rm FM}\big(\widetilde{\nu}_t, \widetilde{\mu}(t,\cdot)\big)dt\right] \ \leq \ & \mathbb{E}\left[\int_0^T \big||\sigma_t|^2 - |f(t)|^2\big|dt\right] + \mathbb{E}\left[\int_0^T |J_t|^2|J_t - 1|\lambda_t dt\right] \\ & + \mathbb{E}\left[\int_0^T \big||J_t|^2 - 1\big|\lambda_t dt\right] + \mathbb{E}\left[\int_0^T |\lambda_t - \mu| dt\right].\end{aligned}$$

When F is the stochastic integral with jumps

$$F = \int_0^T J_t(dY_t - \lambda dt)$$

where $(Y_t)_{t\in[0,T]}$ is a Poisson point process with intensity $\lambda > 0$, with $\sigma_t = 0$ and $N_T = 0$ in G, it follows that the distance $d_{\rm W}(F,G)$ in (1.1a)-(1.1c) of F to a centered Gaussian random variable $G \sim \mathcal{N}\left(0, \int_0^T |f(t)|^2 dt\right)$ is controlled via the bound

$$\mathbb{E}\left[\int_0^T d_{\rm FM}\big(\widetilde{\nu}_t, \widetilde{\mu}(t,\cdot)\big)dt\right] \leq \mathbb{E}\left[\int_0^T \big||f(t)|^2 - \lambda|J_t|^2\big|dt\right] + \lambda\mathbb{E}\left[\int_0^T |J_t|^3 dt\right],$$

see Corollary 5.3. Similarly, when $J_t = 0$ in F and $N_T = 0$ in G, the distance $d_{\rm W}(F,G)$ in (1.1a)-(1.1c) between the distribution of the Brownian stochastic integral

$$F = \int_0^T \sigma_t dB_t$$

and the centered Gaussian random variable $G \sim \mathcal{N}\left(0, \int_0^T |f(t)|^2 dt\right)$ can be controlled via the bound

$$\mathbb{E}\left[\int_0^T d_{\rm FM}\big(\widetilde{\nu}_t, \widetilde{\mu}(t,\cdot)\big)dt\right] \leq \mathbb{E}\left[\int_0^T \big||\sigma_t|^2 - |f(t)|^2\big|dt\right], \tag{1.2}$$

see Corollary 5.4. Observe that in this particular case, other bounds can be directly obtained by bounding the Wasserstein by the L^2 distance and using the Itô isometry, as

$$d_{\rm W}(F,G) \ \leq \ \left(\mathbb{E}\left[\left|G - F\right|^2\right]\right)^{1/2}$$

$$
\begin{aligned}
&= \left(\mathbb{E}\left[\left(\int_0^T (\sigma_t - f(t))dB_t\right)^2\right]\right)^{1/2} \\
&= \left(\mathbb{E}\left[\int_0^T \left|\sigma_t - f(t)\right|^2 dt\right]\right)^{1/2}. \qquad (1.3)
\end{aligned}
$$

However, the bound (1.3) is of a different nature, and it cannot be compared to (1.2) in general. In addition, this simple argument does not apply when mixing continuous and jump stochastic integrals.

We proceed as follows. In Section 2 we recall some background results on martingale and jump-diffusion characteristics. In Section 3 we derive bounds for the sums of forward and backward martingales, which are applied in Section 4 to Wasserstein bounds for forward-backward stochastic integrals with jumps. Applications to standard stochastic integrals are considered in Section 5, and technical lemmas are gathered in the appendix.

2 Notation

Jump-diffusion processes

Consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ equipped with an increasing filtration $(\mathcal{F}_t)_{t\in[0,T]}$, and an $(\mathcal{F}_t)$-martingale $(M_t)_{t\in[0,T]}$ having right-continuous paths with left limits. We denote by $(M_t^c)_{t\in[0,T]}$ the continuous part of $(M_t)_{t\in[0,T]}$, and by

$$
\mu(dt, dy) := \sum_{s>0} \mathbf{1}_{\{\Delta M_s \neq 0\}} \delta_{(s,\Delta M_s)}(dt, dy),
$$

its jump measure, where $\Delta M_t := M_t - M_{t^-}$ denotes jump size and $\delta_{(s,x)}$ is the Dirac measure at $(s, x) \in [0, T] \times \mathbb{R}$. The pair

$$
\big(\nu(dt, dy), \langle M^c, M^c\rangle\big),
$$

where $\nu(dt, dy)$ and $(\langle M^c, M^c\rangle_t)_{t\in[0,T]}$ denote respectively the $(\mathcal{F}_t)_{t\in[0,T]}$-dual predictable projection of $\mu(dt, dy)$ and the predictable quadratic variation of $(M_t)_{t\in[0,T]}$, is called the local characteristics of $(M_t)_{t\in[0,T]}$, see [JM76].

Forward and backward Itô integrals

Let $(\mathcal{F}_t)_{t\in[0,T]}$, resp. $(\mathcal{F}_t^*)_{t\in[0,T]}$ be a forward, resp. backward, filtration, and consider

$$(M_t)_{t\in[0,T]} \text{ an } (\mathcal{F}_t^*)_{t\in[0,T]}\text{-adapted, } (\mathcal{F}_t)_{t\in[0,T]}\text{-forward martingale} \tag{2.1}$$

with right-continuous paths and left limits, and

$$\left(M_t^*\right)_{t\in[0,T]} \text{ an } (\mathcal{F}_t)_{t\in[0,T]}\text{-adapted, } (\mathcal{F}_t^*)_{t\in[0,T]}\text{-backward martingale} \tag{2.2}$$

with left-continuous paths and right limits. Given $(X_t)_{t\in[0,T]}$, resp. $(X_t^*)_{t\in[0,T]}$, a forward (resp. backward) adapted process, the forward and backward Itô differentials d and d^* are respectively defined by the limits in probability

$$\int_0^t F(X_t)\, dM_t = \mathbb{P}\text{-}\lim_{n\to+\infty} \sum_{i=1}^{k_n} f\left(X_{t_{i-1}^n}\right)\left(M_{t_i^n} - M_{t_{i-1}^n}\right) \tag{2.3}$$

and

$$\int_0^t F(X_t^*)\, d^* M_t^* = \mathbb{P}\text{-}\lim_{n\to+\infty} \sum_{i=0}^{k_n-1} f\left(X_{t_{i+1}^n}^*\right)\left(M_{t_i^n}^* - M_{t_{i+1}^n}^*\right) \tag{2.4}$$

for all refining sequences $\{0 = t_0^n \leq t_1^n \leq \cdots \leq t_{k_n}^n = t\}$, $n \geq 1$, of partitions of $[0,t]$ tending to the identity. The following forward-backward Itô formula has been proved in [KMP06, Theorem 8.1].

Proposition 2.1 *Consider $(M_t)_{t\in[0,T]}$ an $(\mathcal{F}_t)_{t\in[0,T]}$-forward martingale and $(M_t^*)_{t\in[0,T]}$ an $(\mathcal{F}_t^*)_{t\in[0,T]}$-backward martingale satisfying (2.1)–(2.2), whose characteristics have the form*

$$\nu(dt,dx) = \nu_t(dx)dt, \qquad \nu^*(dt,dx) = \nu_t^*(dx)dt, \tag{2.5}$$

with the predictable quadratic variations

$$d\langle M^c, M^c\rangle_t = |\sigma_t|^2 dt \quad \text{and} \quad d\langle M^{*c}, M^{*c}\rangle_t = \left|\sigma_t^*\right|^2 dt, \tag{2.6}$$

where $(\sigma_t)_{t\in[0,T]}$, $\left(\sigma_t^*\right)_{t\in[0,T]}$, are predictable with respect to $(\mathcal{F}_t)_{t\in[0,T]}$ and $(\mathcal{F}_t^*)_{t\in[0,T]}$, respectively. Then, for all $f \in \mathcal{C}^2(\mathbb{R}^2,\mathbb{R})$ we have

$$f(M_t, M_t^*) - f(M_0, M_0^*)$$
$$= \int_{0^+}^t \frac{\partial f}{\partial x_1}(M_{u^-}, M_u^*) dM_u + \frac{1}{2}\int_0^t \frac{\partial^2 f}{\partial x_1^2}(M_u, M_u^*) d\langle M^c, M^c\rangle_u$$

$$
\begin{aligned}
&+\sum_{0<u\le t}\left(f(M_u,M_u^*)-f(M_{u^-},M_u^*)-\Delta M_u\frac{\partial f}{\partial x_1}(M_{u^-},M_u^*)\right)\\
&-\int_0^{t^-}\frac{\partial f}{\partial x_2}(M_u,M_{u^+}^*)d^*M_u^*-\frac{1}{2}\int_0^t\frac{\partial^2 f}{\partial x_2^2}(M_u,M_u^*)d\langle M^{*c},M^{*c}\rangle_u\\
&-\sum_{0\le u<t}\left(f(M_u,M_u^*)-f(M_u,M_{u^+}^*)-\Delta M_u^*\frac{\partial f}{\partial x_2}(M_u,M_{u^+}^*)\right).
\end{aligned}
$$

Note that in the above statements, $(M_t^c)_{t\in[0,T]}$ and $(M_t^{*c})_{t\in[0,T]}$ respectively denote the continuous parts of $(M_t)_{t\in[0,T]}$ and $(M_t^*)_{t\in[0,T]}$.

Distances between distributions

Given a set $\mathcal{H}$ of functions $h:\mathbb{R}\to\mathbb{R}$, we define the distance $d_{\mathcal{H}}$ between two probability measures μ, ν on $(\mathbb{R},\mathcal{B}(\mathbb{R}))$ by

$$
d_{\mathcal{H}}(\mu,\nu):=\sup_{h\in\mathcal{H}}\left|\int_{-\infty}^{+\infty}h\,d\mu-\int_{-\infty}^{+\infty}h\,d\nu\right|,
$$

and we write $d_{\mathcal{H}}(X,Y)=d_{\mathcal{H}}(\mu,\nu)$ when μ and ν are the probability distributions of the random variables X,Y.

The Fortet-Mourier distance d_{FM} corresponds to the choice $\mathcal{H}=\mathcal{FM}$, where $\mathcal{FM}$ is the class of functions h such that $\|h\|_{BL}=\|h\|_L+\|h\|_\infty\le 1$, where $\|\cdot\|_L$ denotes the Lipschitz semi-norm and $\|\cdot\|_\infty$ is the supremum norm.

The total variation distance d_{TV} is obtained when $\mathcal{H}$ is the set of indicator functions $\mathbf{1}_A$, $A\in\mathcal{B}(\mathbb{R})$.

The Wasserstein distance d_{W} corresponds to $\mathcal{H}=\mathrm{Lip}(1)$, where $\mathrm{Lip}(1)$ is the class of functions h such that $\|h\|_L\le 1$.

The smooth Wasserstein distance d_{W_r}, $r\ge 0$, is obtained when $\mathcal{H}:=\mathcal{H}_r$ is the set of continuous functions which are r-times continuously differentiable and such that $\|h^{(k)}\|_\infty\le 1$, for all $0\le k\le r$, where $h^{(0)}=h$, and where $h^{(k)}$, $k\ge 1$, is the *k-th* derivative of h.

It is easy to observe that $d_{\mathrm{FM}}(\cdot,\cdot)\le d_{\mathrm{W}}(\cdot,\cdot)$ and the topologies induced by d_{W} and d_{TV} are stronger than the topology of convergence in distribution which is metrized by d_{FM}.

Moreover, for the smooth Wasserstein distance d_{W_r} with, $r > 1$, an approximation argument shows that

$$d_{\mathrm{W}_r}(X,Y) = \sup_{h \in C_c^\infty(\mathbb{R}) \cap \mathcal{H}_r} |\mathbb{E}[h(X)] - \mathbb{E}[h(Y)]|, \tag{2.7}$$

where $C_c^\infty(\mathbb{R})$ is the space of compactly supported, infinitely differentiable functions on $\mathbb{R}$, see Lemma A.3 in [AH19]. Note that $d_{\mathrm{W}_{r-1}}(X,Y) \le 3\sqrt{2}\sqrt{d_{\mathrm{W}_r}(X,Y)}$ and that the smooth Wasserstein distance d_{W_r} is a weaker distance than the Wasserstein distance d_{W} since

$$d_{\mathrm{W}_r}(X,Y) \le d_{\mathrm{W}_1}(X,Y) \le d_{\mathrm{W}}(X,Y),$$

see (2.16) in [AH19] to which we refer for further details in this direction, see also [Dud02].

Moreover, recall that for a signed measure μ on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ with Jordan decomposition $\mu = \mu^+ + \mu^-$ in terms of (positive) measures μ^+, μ^-, we note $|\mu|_{\mathrm{TV}} = \mu^+ + \mu^-$ its total variation measure and $\|\mu\|_{\mathrm{TV}} = |\mu|_{\mathrm{TV}}(\mathbb{R})$ its total variation. With this notation, we have $d_{\mathrm{TV}}(\mu,\nu) = \|\mu - \nu\|_{\mathrm{TV}}$. We also let $\mathcal{C}_b^n(\mathbb{R})$, $1 \le n \le \infty$, denote the space of functions in $\mathcal{C}^n(\mathbb{R})$ with bounded derivatives of orders 1 to n.

3 Wasserstein bounds for forward and backward integrals

We begin with distance estimates for forward and backward martingales, which will be applied to stochastic integrals with jumps in Section 5. In the next Theorems 3.1 and 3.3 we derive general bounds on the Wasserstein distance between values of the sum $M_t + M_t^*$ of a forward and a backward martingale at different times. Our argument allows us to provide three bounds in terms of either the total variation, Wasserstein or Fortet-Mourier distances. Since the bounds (3.12), (3.13) and (3.14) are not directly comparable, we state each of them explicitly. In the sequel, we let $f(x) \sim_{x\to 0} g(x)$ if $\lim_{x\to 0} f(x)/g(x) = 1$. First, we have the following bounds for the smooth Wasserstein distances:

Theorem 3.1 *Consider* $(M_t)_{t\in[0,T]}$ *an* $(\mathcal{F}_t)_{t\in[0,T]}$*-forward martingale and* $(M_t^*)_{t\in[0,T]}$ *an* $(\mathcal{F}_t^*)_{t\in[0,T]}$*-backward martingale satisfying* (2.1)–(2.2)*. Assume also that the local char-*

acteristics of $(M_t)_{t\in[0,T]}$ *and* $(M_t^*)_{t\in[0,T]}$ *have the form* (2.5)–(2.6)*, and let*

$$\widetilde{\nu}_t(dx) := |\sigma_t|^2\delta_0(dx) + |x|^2\nu_t(dx), \qquad \widetilde{\nu}_t^*(dx) := |\sigma_t^*|^2\delta_0(dx) + |x|^2\nu_t^*(dx). \tag{3.1}$$

Then, the following bounds hold true for the smooth Wasserstein distances for $s \le t$*:*

$$d_{\mathrm{W}_2}(M_s + M_s^*, M_t + M_t^*) \;\le\; \frac{1}{2}\,\mathbb{E}\left[\int_s^t d_{\mathrm{TV}}\big(\widetilde{\nu}_u, \widetilde{\nu}_u^*\big)\, du\right], \tag{3.2}$$

$$d_{\mathrm{W}_3}(M_s + M_s^*, M_t + M_t^*) \;\le\; \frac{1}{6}\,\mathbb{E}\left[\int_s^t d_{\mathrm{W}}\big(\widetilde{\nu}_u, \widetilde{\nu}_u^*\big)\, du\right], \tag{3.3}$$

$$d_{\mathrm{W}_3}(M_s + M_s^*, M_t + M_t^*) \;\le\; \frac{2}{3}\,\mathbb{E}\left[\int_s^t d_{\mathrm{FM}}\big(\widetilde{\nu}_u, \widetilde{\nu}_u^*\big)\, du\right]. \tag{3.4}$$

Remark 3.2 The above bounds are obtained by arguments similar to each other. Although (3.4) is stronger than (3.3) since $d_{\mathrm{FM}}(\cdot,\cdot) \le d_{\mathrm{W}}(\cdot,\cdot)$, we note that (3.3) has a smaller constant.

Proof. We start by bounding the absolute difference $\big|\mathbb{E}[h(M_t+M_t^*)]-\mathbb{E}[h(M_s+M_s^*)]\big|$ for $h \in \mathcal{C}^2(\mathbb{R})$. By Itô's formula for forward-backward martingales (see Proposition 2.1) applied to $f(x_1, x_2) := h(x_1 + x_2)$ we have, for $0 \le s \le t$,

$$\begin{aligned}
&h(M_t + M_t^*) = h(M_s + M_s^*) \\
&\quad + \int_{s^+}^t h'(M_{u^-} + M_u^*) dM_u + \frac{1}{2}\int_s^t h''(M_u + M_u^*)\; d\langle M^c, M^c\rangle_u \\
&\quad + \sum_{s<u\le t} \big(h(M_{u^-} + M_u^* + \Delta M_u) - h(M_{u^-} + M_u^*) - \Delta M_u\; h'(M_{u^-} + M_u^*)\big) \\
&\quad - \int_s^{t^-} h'(M_u + M_{u^+}^*) d^* M_u^* - \frac{1}{2}\int_s^t h''(M_u + M_u^*)\; d\langle M^{*c}, M^{*c}\rangle_u \\
&\quad - \sum_{s\le u<t} \big(h(M_u + M_{u^+}^* + \Delta^* M_u^*) - h(M_u + M_{u^+}^*) - \Delta^* M_u^*\; h'(M_u + M_{u^+}^*)\big),
\end{aligned}$$

where d and d^* denote the forward and backward Itô differential as defined in (2.3) and (2.4) and $\Delta^* M_t^* = M_t^* - M_{t^+}^*$. Taking expectations and taking into account the vanishing of martingale terms, we find

$$\begin{aligned}
&\mathbb{E}\big[h(M_t + M_t^*)\big] = \mathbb{E}\big[h(M_s + M_s^*)\big] \\
&\quad + \frac{1}{2}\mathbb{E}\left[\int_s^t h''(M_u + M_u^*) d\big(\langle M^c, M^c\rangle_u - \langle M^{*c}, M^{*c}\rangle_u\big)\right] \\
&\quad + \mathbb{E}\left[\int_s^t \int_{-\infty}^{+\infty} \big(h(M_u + M_u^* + x) - h(M_u + M_u^*) - x h'(M_u + M_u^*)\big)\nu_u(dx)\, du\right]
\end{aligned} \tag{3.5}$$

$$-\mathbb{E}\left[\int_s^t \int_{-\infty}^{+\infty} \big(h(M_u+M_u^*+x)-h(M_u+M_u^*)-xh'(M_u+M_u^*)\big)\nu_u^*(dx)\,du\right].$$

For $h \in \mathcal{C}^2(\mathbb{R})$, the Taylor formula

$$h(y+x)=h(y)+xh'(y)+|x|^2\int_0^1(1-\tau)h''(y+\tau x)d\tau$$

allows us to rewrite (3.5) as

$$\begin{aligned}
&\mathbb{E}\big[h(M_t+M_t^*)\big]-\mathbb{E}\big[h(M_s+M_s^*)\big]\\
&= \frac{1}{2}\mathbb{E}\left[\int_s^t h''(M_u+M_u^*)\big(|\sigma_u|^2-|\sigma_u^*|^2\big)\,du\right] \qquad (3.6)\\
&\quad +\mathbb{E}\left[\int_0^1(1-\tau)\int_s^t\int_{-\infty}^{+\infty} h''(M_u+M_u^*+\tau x)|x|^2\big(\nu_u(dx)-\nu_u^*(dx)\big)\,dud\tau\right]\\
&= \mathbb{E}\left[\int_0^1(1-\tau)\int_s^t\int_{-\infty}^{+\infty} h''(M_u+M_u^*+\tau x)\big(\widetilde{\nu}_u(dx)-\widetilde{\nu}_u^*(dx)\big)\,dud\tau\right] \qquad (3.7)
\end{aligned}$$

where $\widetilde{\nu}_u(dx)$ and $\widetilde{\nu}_u^*(dx)$ are defined in (3.1). When $h \in \mathcal{C}_b^2(\mathbb{R})$, (3.7) entails

$$\big|\mathbb{E}\big[h\big(M_t+M_t^*\big)\big]-\mathbb{E}\big[h\big(M_s+M_s^*\big)\big]\big| \le \frac{\|h''\|_\infty}{2}\mathbb{E}\left[\int_s^t d_{\mathrm{TV}}\big(\widetilde{\nu}_u,\widetilde{\nu}_u^*\big)\,du\right]. \qquad (3.8)$$

When $h \in \mathcal{C}_b^3(\mathbb{R})$, then $h'' \in \mathrm{Lip}(\|h^{(3)}\|_\infty)$, and (3.7) implies

$$\begin{aligned}
&\big|\mathbb{E}\big[h\big(M_t+M_t^*\big)\big]-\mathbb{E}\big[h\big(M_s+M_s^*\big)\big]\big|\\
&\le \mathbb{E}\left[\int_0^1(1-\tau)\int_s^t\left|\int_{-\infty}^{+\infty} h''(M_u+M_u^*+\tau x)\big(\widetilde{\nu}_u(dx)-\widetilde{\nu}_u^*(dx)\big)\right|dud\tau\right] \qquad (3.9)\\
&\le \mathbb{E}\left[\int_0^1(1-\tau)\int_s^t(\tau\|h^{(3)}\|_\infty+\|h''\|_\infty)d_{\mathrm{FM}}\big(\widetilde{\nu}_u,\widetilde{\nu}_u^*\big)dud\tau\right]
\end{aligned}$$

since the function $x \mapsto h''(M_u+M_u^*+\tau x)$ is almost surely $(\tau\|h^{(3)}\|_\infty)$-Lipschitz and bounded by $\|h''\|_\infty$. As a consequence, we have

$$\big|\mathbb{E}\big[h\big(M_t+M_t^*\big)\big]-\mathbb{E}\big[h\big(M_s+M_s^*\big)\big]\big| \le \Big(\frac{1}{6}\|h^{(3)}\|_\infty+\frac{1}{2}\|h''\|_\infty\Big)\mathbb{E}\left[\int_s^t d_{\mathrm{FM}}\big(\widetilde{\nu}_u,\widetilde{\nu}_u^*\big)\,du\right]. \qquad (3.10)$$

Alternatively, (3.9) also entails the (weaker) bound

$$\big|\mathbb{E}\big[h\big(M_t+M_t^*\big)\big]-\mathbb{E}\big[h\big(M_s+M_s^*\big)\big]\big| \le \frac{1}{6}\|h^{(3)}\|_\infty\mathbb{E}\left[\int_s^t d_{\mathrm{W}}\big(\widetilde{\nu}_u,\widetilde{\nu}_u^*\big)\,du\right]. \qquad (3.11)$$

Due to (2.7), the inequalities (3.8), (3.10) and (3.11) immediately give the bounds for the smooth Wasserstein distances in (3.2), (3.4) and (3.3). □

The next proposition presents bounds for the Wasserstein distance.

Theorem 3.3 *In the same setting as in Theorem 3.1, the following bounds hold true for the Wasserstein distance:*

(1) For $s \leq t$ close enough (see (3.18)):

$$d_{\rm W}(M_s + M_s^*, M_t + M_t^*) \leq \sqrt{\big(4 + \mathbb{E}\big[|M_t + M_t^*|\big] + \mathbb{E}\big[|M_s + M_s^*|\big]\big)\mathbb{E}\left[\int_s^t d_{\rm TV}\big(\widetilde{\nu}_u, \widetilde{\nu}_u^*\big)\ du\right]}. \quad (3.12)$$

(2) For $s \leq t$ close enough (see (3.19)):

$$d_{\rm W}(M_s + M_s^*, M_t + M_t^*) \leq \frac{3}{\sqrt[3]{16}}\big(4 + \mathbb{E}\big[\big|M_s + M_s^*\big|\big] + \mathbb{E}\big[\big|M_t + M_t^*\big|\big]\big)^{2/3} \left(\mathbb{E}\left[\int_s^t d_{\rm W}\big(\widetilde{\nu}_u, \widetilde{\nu}_u^*\big)\, du\right)\right)^{1/3}. \quad (3.13)$$

(3) For $s \leq t$ close enough (see (3.20)):

$$\begin{aligned} d_{\rm W}(M_s + M_s^*, M_t + M_t^*) &\leq \frac{b\big(1 + \sqrt{\alpha_*(a,b)}\big)}{\alpha_*(a,b)} \sim_{b \to 0} \sqrt[3]{\frac{9a^2b}{4}} \\ &= O\left(\left(\mathbb{E}\left[\int_s^t d_{\rm FM}\big(\widetilde{\nu}_u, \widetilde{\nu}_u^*\big)\, du\right]\right)^{1/3}\right), \end{aligned} \quad (3.14)$$

where $f(t) = O(g(t))$ means $f(t)/g(t)$ is bounded as $t \searrow s$, and we set

$$a := \frac{1}{2}\big(4 + \mathbb{E}\big[\big|M_s + M_s^*\big|\big] + \mathbb{E}\big[\big|M_t + M_t^*\big|\big]\big), \quad b := \mathbb{E}\left[\int_s^t d_{\rm FM}\big(\widetilde{\nu}_u, \widetilde{\nu}_u^*\big)\, du\right]. \quad (3.15)$$

Proof. The argument is similar to the proof of Theorem 3.1, however since the definition of the (usual) Wasserstein distance $d_{\rm W}$ requires to use $h \in {\rm Lip}(1)$, we cannot directly apply (3.8), (3.10) or (3.11) which have been obtained for $h \in \mathcal{C}^3(\mathbb{R})$ in the proof of Theorem 3.1. For this reason, we consider the approximation h_α of $h \in {\rm Lip}(1)$ given by

$$h_\alpha(x) := \int_{-\infty}^{+\infty} h\big(y\sqrt{\alpha} + x\sqrt{1-\alpha}\big)\phi(y)\ dy, \qquad 0 < \alpha < 1, \quad (3.16)$$

where ϕ is the density of the standard $\mathcal{N}(0,1)$-distribution and we apply (3.8), (3.10) or (3.11) to h_α and combine with the approximation Lemma A.1 in order to recover a bound for $h \in {\rm Lip}(1)$.

(1) Using (3.8) and (A.1) in Lemma A.1, we have for all $\alpha \in (0,1)$:

$$\begin{aligned}
&\big|\mathbb{E}\big[h\big(M_t + M_t^*\big)\big] - \mathbb{E}\big[h\big(M_s + M_s^*\big)\big]\big| \\
&\quad \leq \ \big|\mathbb{E}\big[h\big(M_t + M_t^*\big)\big] - \mathbb{E}\big[h_\alpha\big(M_t + M_t^*\big)\big]\big| + \big|\mathbb{E}\big[h_\alpha\big(M_t + M_t^*\big)\big] - \mathbb{E}\big[h_\alpha\big(M_s + M_s^*\big)\big]\big| \\
&\qquad + \big|\mathbb{E}\big[h_\alpha\big(M_s + M_s^*\big)\big] - \mathbb{E}\big[h\big(M_s + M_s^*\big)\big]\big| \\
&\quad \leq \ \frac{\sqrt{\alpha}}{2}\big(4 + \mathbb{E}\big[\big|M_t + M_t^*\big|\big] + \mathbb{E}\big[\big|M_s + M_s^*\big|\big]\big) + \frac{1}{2\sqrt{\alpha}}\mathbb{E}\left[\int_s^t d_{\mathrm{TV}}\big(\widetilde{\nu}_u, \widetilde{\nu}_u^*\big)\ du\right]. \qquad (3.17)
\end{aligned}$$

Next, minimizing (3.17) in $\alpha \in (0,1)$ with

$$\alpha_* := \frac{\mathbb{E}\left[\int_s^t d_{\mathrm{TV}}\big(\widetilde{\nu}_u, \widetilde{\nu}_u^*\big)\ du\right]}{4 + \mathbb{E}[|M_s + M_s^*|] + \mathbb{E}[|M_t + M_t^*|]} < 1, \tag{3.18}$$

for instance when

$$\mathbb{E}\left[\int_s^t d_{\mathrm{TV}}\big(\widetilde{\nu}_u, \widetilde{\nu}_u^*\big)\ du\right] < 4,$$

and taking the maximum in $h \in \mathrm{Lip}(1)$, we obtain the bound (3.12).

(2) Similarly, using now (3.11), we have for all $\alpha \in (0,1)$:

$$\begin{aligned}
&\big|\mathbb{E}\big[h\big(M_t + M_t^*\big)\big] - \mathbb{E}\big[h\big(M_s + M_s^*\big)\big]\big| \\
&\quad \leq \ \frac{\sqrt{\alpha}}{2}\big(4 + \mathbb{E}\big[\big|M_s + M_s^*\big|\big] + \mathbb{E}\big[\big|M_t + M_t^*\big|\big]\big) + \frac{1}{3\alpha}\mathbb{E}\left[\int_s^t d_{\mathrm{W}}\big(\widetilde{\nu}_u, \widetilde{\nu}_u^*\big)\, du\right].
\end{aligned}$$

Optimizing the quantity $a\sqrt{\alpha} + b/\alpha$ in $\alpha > 0$ with

$$a = \frac{1}{2}\big(4 + \mathbb{E}\big[\big|M_s + M_s^*\big|\big] + \mathbb{E}\big[\big|M_t + M_t^*\big|\big]\big), \quad b = \frac{1}{3}\mathbb{E}\left[\int_s^t d_{\mathrm{W}}\big(\widetilde{\nu}_u, \widetilde{\nu}_u^*\big)\, du\right]$$

and

$$\alpha_* := \sqrt[3]{\frac{4b^2}{a^2}} < 1, \tag{3.19}$$

we obtain

$$\begin{aligned}
&d_{\mathrm{W}}(M_s + M_s^*, M_t + M_t^*) \\
&\quad \leq \ \frac{3}{\sqrt[3]{16}}\big(4 + \mathbb{E}\big[\big|M_s + M_s^*\big|\big] + \mathbb{E}\big[\big|M_t + M_t^*\big|\big]\big)^{2/3}\left(\mathbb{E}\left[\int_s^t d_{\mathrm{W}}\big(\widetilde{\nu}_u, \widetilde{\nu}_u^*\big)\, du\right]\right)^{1/3}
\end{aligned}$$

when $s \leq t$ are close enough, for instance when

$$\mathbb{E}\left[\int_s^t d_{\mathrm{W}}\big(\widetilde{\nu}_u, \widetilde{\nu}_u^*\big)\, du\right] < 3.$$

(3) Similarly, using (3.10) and still (A.1) in Lemma A.1, we have

$$\begin{aligned} & \left|\mathbb{E}\big[h\big(M_t+M_t^*\big)\big]-\mathbb{E}\big[h\big(M_s+M_s^*\big)\big]\right| \\ & \quad \leq \frac{\sqrt{\alpha}}{2}\big(4+\mathbb{E}\big[\big|M_s+M_s^*\big|\big]+\mathbb{E}\big[\big|M_t+M_t^*\big|\big]\big)+\Big(\frac{1}{3\alpha}+\frac{1}{2\sqrt{\alpha}}\Big)\mathbb{E}\left[\int_s^t d_{\mathrm{FM}}\big(\widetilde{\nu}_u,\widetilde{\nu}_u^*\big)\,du\right]. \end{aligned}$$

Next, with Lemma A.2, we optimize the above quantity of the form $a\sqrt{\alpha}+b/(2\sqrt{\alpha})+b/(3\alpha)$ in $\alpha>0$, with a,b given in (3.15), under appropriate conditions on $d_{\mathrm{FM}}\big(\widetilde{\nu}_u,\widetilde{\nu}_u^*\big)$, so that b is small enough to ensure that $\alpha_*(a,b)$ in (A.2) satisfies

$$\alpha_*(a,b)<1. \tag{3.20}$$

The bound (3.10) is then optimized into

$$\begin{aligned} d_{\mathrm{W}}(M_s+M_s^*,M_t+M_t^*) & \leq b\frac{1+\sqrt{\alpha_*(a,b)}}{\alpha_*(a,b)}\sim_{b\to 0}\sqrt[3]{\frac{9a^2b}{4}} \\ & = O\left(\left(\mathbb{E}\left[\int_s^t d_{\mathrm{FM}}\big(\widetilde{\nu}_u,\widetilde{\nu}_u^*\big)\,du\right]\right)^{1/3}\right). \end{aligned} \tag{3.21}$$

□

The above bounds (3.2)–(3.4) and (3.12)–(3.14) rely on the distance $d_{\mathcal{H}}\big(\widetilde{\nu}_t,\widetilde{\nu}_t^*\big)$ for $\mathcal{H}\in\{\mathrm{TV},\mathrm{FM},\mathrm{W}\}$ which satisfies

$$\begin{aligned} d_{\mathcal{H}}\big(\widetilde{\nu}_t,\widetilde{\nu}_t^*\big) & = \sup_{h\in\mathcal{H}}\left|\int_{-\infty}^{+\infty}h(x)\widetilde{\nu}_t(dx)-\int_{-\infty}^{+\infty}h(x)\widetilde{\nu}_t^*(dx)\right| \\ & = \sup_{h\in\mathcal{H}}\left|(|\sigma_t|^2-|\sigma_t^*|^2)h(0)+\int_{-\infty}^{+\infty}y^2h(y)\nu_t(dy)-\int_{-\infty}^{+\infty}y^2h(y)\nu_t^*(dy)\right| \end{aligned}$$

from the definition (3.1). When $\mathcal{H}=\mathrm{TV}$, this gives

$$d_{\mathrm{TV}}\big(\widetilde{\nu}_t,\widetilde{\nu}_t^*\big)=\big||\sigma_t|^2-|\sigma_t^*|^2\big|+\int_{-\infty}^{+\infty}|y|^2|\nu_t-\nu_t^*|_{\mathrm{TV}}(dy),$$

where $|\nu_t-\nu_t^*|_{\mathrm{TV}}(dy)$ denotes the total variation measure obtained from the Jordan decomposition of $\nu_u-\nu_u^*$, see Section 2, while when $\mathcal{H}=\mathrm{W}$ we have

$$d_{\mathrm{W}}\big(\widetilde{\nu}_t,\widetilde{\nu}_t^*\big)\leq\sup_{h\in\mathcal{H}}\left(\big||\sigma_t|^2-\big|\sigma_t^*\big|^2\big||h(0)|+\left|\int_{-\infty}^{+\infty}y^2h(y)\nu_t(dy)-\int_{-\infty}^{+\infty}y^2h(y)\nu_t^*(dy)\right|\right).$$

4 Wasserstein bounds for stochastic integrals with jumps

We now consider random variables F given by the sum

$$F = \int_0^T \sigma_t dB_t + Z_T \tag{4.1}$$

of a Wiener integral and the value at time T of a pure-jump martingale $(Z_t)_{t\in[0,T]}$ with compensator $\nu_t(dx)dt$, where $(B_t)_{t\in[0,T]}$ is a standard Brownian motion and $(\sigma_t)_{t\in[0,T]}$ is adapted with respect to the filtration generated by $(B_t)_{t\in[0,T]}$. In this section, the random variable G is given by the backward counterpart of (4.1), i.e. as the sum

$$G = \int_0^T \sigma_t^* d^* B_t^* + Z_0^* \tag{4.2}$$

of a backward Wiener integral and the value at time 0 of a backward pure-jump martingale $(Z_t^*)_{t\in[0,T]}$ with compensator $\nu_t^*(dx)dt$, where $(B_t^*)_{t\in[0,T]}$ is a backward Brownian motion and $(\sigma_t^*)_{t\in[0,T]}$ is adapted with respect to the (backward) filtration generated by $(B_t^*)_{t\in[0,T]}$.

In order to bound the smooth Wasserstein distances $d_{\mathrm{W}_r}(F,G)$ and the Wasserstein distance $d_{\mathrm{W}}(F,G)$, we will apply the results of Section 3 to suitable forward and backward martingales recovering respectively F and G as their initial and final values.

Corollary 4.1 *Let F and G be given by (4.1) and (4.2). Then, with ν_t and ν_t^* given in (3.1), the following bounds hold true:*

$$d_{\mathrm{W}_2}(F,G) \le \frac{1}{2}\,\mathbb{E}\left[\int_0^T d_{\mathrm{TV}}\big(\widetilde{\nu}_t, \widetilde{\nu}_t^*\big)\ dt\right], \tag{4.3}$$

$$d_{\mathrm{W}_3}(F,G) \le \frac{1}{6}\,\mathbb{E}\left[\int_0^T d_{\mathrm{W}}\big(\widetilde{\nu}_t, \widetilde{\nu}_t^*\big)\, dt\right], \tag{4.4}$$

$$d_{\mathrm{W}_3}(F,G) \le \frac{2}{3}\,\mathbb{E}\left[\int_0^T d_{\mathrm{FM}}\big(\widetilde{\nu}_t, \widetilde{\nu}_t^*\big)\, dt\right]. \tag{4.5}$$

The same comments as in Remark 3.2 apply to the bounds (4.4) and (4.5).

Proof. We assume without loss of generality that F and G are independent, and apply Theorem 3.1. To that purpose, denoting by $(\mu(dt,dy))_{(t,y)\in[0,T]\times\mathbb{R}}$ and $\big(\mu^*(dt,dy)\big)_{(t,y)\in[0,T]\times\mathbb{R}}$ the jump measures of Z in (4.1) and of Z^* in (4.2), we have

$$Z_t = \int_0^t \int_{-\infty}^{+\infty} y(\mu(ds,dy) - \nu_s(dy)ds) \quad \text{and} \quad Z_t^* = \int_t^T \int_{-\infty}^{+\infty} y(\mu^*(ds,dy) - \nu_s^*(dy)ds).$$

Due to the independence of F and G, we have that $(B_t)_{t\in[0,T]}$ and $\big(\mu(dt,dx)\big)_{t\in[0,T]}$ are independent from $(B_t^*)_{t\in[0,T]}$ and $\big(\mu^*(dt,dx)\big)_{t\in[0,T]}$. Next, letting

$$M_t = \int_0^t \sigma_s dB_s + \int_0^t \int_{-\infty}^{+\infty} y\big(\mu(ds,dy) - \nu_s(dy)ds\big),$$

and

$$M_t^* = \int_t^T \sigma_s^* d^* B_s^* + \int_t^T \int_{-\infty}^{+\infty} y(\mu^*(ds,dy) - \nu_s^*(dx)ds),$$

we note that $F = M_T$ and $G = M_0^*$. Let also $(\mathcal{F}_t^M)_{t\in[0,T]}$ be the (forward) filtration generated by $(B_t)_{t\in[0,T]}$ and $\big(\mu(dt,dx)\big)_{t\in[0,T]}$, given by

$$\mathcal{F}_t^M = \sigma\big(B_s,\ \mu([0,s]\times A)\ :\ 0\le s\le t, A\in\mathcal{B}_b(\mathbb{R})\big),$$

and let $(\mathcal{F}_t^{M^*})_{t\in[0,T]}$ be the (backward) filtration generated by $(B_t^*)_{t\in[0,T]}$ and $\big(\mu^*(dt,dx)\big)_{t\in[0,T]}$, given by

$$\mathcal{F}_t^{M^*} = \sigma(B_T^* - B_s^*,\ \mu^*([s,T]\times A)\ :\ t\le s\le T, A\in\mathcal{B}_b(\mathbb{R})), \quad t\in[0,T].$$

Note that $(M_t)_{t\in[0,T]}$ is a forward martingale with respect to $(\mathcal{F}_t^M)$ while $(M_t^*)_{t\in[0,T]}$ is a backward martingale with respect $(\mathcal{F}_t^{M^*})$. In order to apply Theorem 3.1, we consider the following forward and backward filtrations:

$$\mathcal{F}_t = \mathcal{F}_t^M \vee \mathcal{F}_0^{M^*}, \quad \mathcal{F}_t^* = \mathcal{F}_T^M \vee \mathcal{F}_t^{M^*}, \quad t\in[0,T].$$

Due to the independence of $\mathcal{F}_T^M$ and $\mathcal{F}_0^{M^*}$, the process $(M_t)_{t\in[0,T]}$ is a forward $(\mathcal{F}_t)$-martingale satisfying (2.1) since M_t is $\mathcal{F}_T^M$-measurable for all $t\in[0,T]$. Similarly, $(M_t^*)_{t\in[0,T]}$ is a backward $(\mathcal{F}_t^*)$-martingale satisfying (2.2) since M_t^* is $\mathcal{F}_T^{M^*}$-measurable for all $t\in[0,T]$.

Since the local characteristics of $(M_t)_{t\in[0,T]}$ are $|\sigma_t|^2$ and $\nu_t^{(M)} = \nu_t$ and those of $(M_t^*)_{t\in[0,T]}$ are $|\sigma_t^*|^2$ and $\nu_t^{(M^*)} = \nu_t^*$, Theorem 3.1 applies to $(M_t)_{t\in[0,T]}$ and $(M_t^*)_{t\in[0,T]}$, and since $M_T = F$, $M_0^* = G$, and $M_0 = M_T^* = 0$, the bound (3.2) yields

$$d_{\mathrm{W}_2}(F,G) \le \frac{1}{2}\,\mathbb{E}\left[\int_s^t d_{\mathrm{TV}}\big(\widetilde{\nu}_u, \widetilde{\nu}_u^*\big)\ du\right],$$

which is (4.3). Similarly, (3.3) yields (4.4), and (3.4) yields (4.5). □

Following the same strategy, we derive similarly from Theorem 3.3 the following bounds for the Wasserstein distance of F and G in (4.1) and (4.2). Observe that for $s = 0$ and $t = T$, the closeness requirements about s and t in Theorem 3.3, see (3.18)-(3.20), are also satisfied under the assumptions in (1)-(3) below, under suitable conditions on the martingale characteristics σ_t, σ_t^*, ν_t, ν_t^*, $t \in [0, T]$.

Corollary 4.2 *Let F and G be given by (4.1) and (4.2). Then, with ν_t and ν_t^* given in (3.1), the following bounds hold true:*

(1) Under the condition

$$\mathbb{E}\left[\int_0^T d_{\mathrm{TV}}\big(\widetilde{\nu}_t, \widetilde{\nu}_t^*\big)dt\right] < 4 + \mathbb{E}[|F|] + \mathbb{E}[|G|],$$

we have

$$d_{\mathrm{W}}(F, G) \leq \sqrt{(4 + \mathbb{E}[|F|] + \mathbb{E}[|G|])\mathbb{E}\left[\int_0^T d_{\mathrm{TV}}\big(\widetilde{\nu}_t, \widetilde{\nu}_t^*\big)dt\right]}.$$

(2) Under the condition

$$\mathbb{E}\left[\int_0^T d_{\mathrm{W}}\big(\widetilde{\nu}_t, \widetilde{\nu}_t^*\big)\ dt\right] < \frac{3}{4}(4 + \mathbb{E}[|F|] + \mathbb{E}[|G|]),$$

we have

$$d_{\mathrm{W}}(F, G) \leq \frac{\sqrt[3]{9}}{2}(4 + \mathbb{E}[|F|] + \mathbb{E}[|G|])^{2/3}\left(\mathbb{E}\left[\int_0^T d_{\mathrm{W}}\big(\widetilde{\nu}_t, \widetilde{\nu}_t^*\big)\ dt\right]\right)^{1/3}. \tag{4.6}$$

(3) Letting

$$a := \frac{1}{2}(4 + \mathbb{E}[|F|] + \mathbb{E}[|G|]) \quad \textit{and} \quad b := \mathbb{E}\left[\int_0^T d_{\mathrm{FM}}\big(\widetilde{\nu}_t, \widetilde{\nu}_t^*\big)\ dt\right],$$

and assuming that $\alpha_*(a, b)$ in (A.2) satisfies $\alpha_*(a, b) < 1$, e.g. for b small enough, we have

$$d_{\mathrm{W}}(F, G) \leq b\frac{1 + \sqrt{\alpha_*(a, b)}}{(\alpha_*(a, b))^{3/2}} \sim \sqrt[3]{\frac{9a^2 b}{4}}, \quad b \to 0. \tag{4.7}$$

5 Application to stochastic integrals

In this section, the results of the previous sections are specialized to the comparison of random variables F and G respectively given as the sum of standard Brownian Itô integral and a pure jump-martingale, and as the sum of a standard Wiener integral and a compound Poisson process. In the following result, the random variable G is expressed using a standard forward stochastic integral.

Corollary 5.1 *Consider*

$$F = \int_0^T \sigma_t dB_t + Z_T,$$

as in (4.1), and let

$$G := \int_0^T f(t) dB_t + N_T \tag{5.1}$$

where $(f(t))_{t\in[0,T]}$ is a deterministic function and N_T is a (compensated) compound Poisson process with compensator $\mu(t,dx)dt$, satisfying the condition

$$\int_0^T \int_{-\infty}^{\infty} \min(1,x^2)\mu(t,dx)dt < \infty.$$

Then the bounds of Corollary 4.1 and Corollary 4.2 apply respectively to the smooth Wasserstein distance $d_{W_r}(F,G)$, $r = 2,3$, and to the Wasserstein distance $d_W(F,G)$ by taking

$$\widetilde{\nu}_t(dx) := |\sigma_t|^2 \delta_0(dx) + |x|^2 \nu_t(dx) \quad \textit{and} \quad \widetilde{\mu}(t,dx) := |f(t)|^2 \delta_0(dx) + |x|^2 \mu(t,dx). \tag{5.2}$$

Proof. Since for a deterministic integrand $(f(t))_{t\in[0,T]}$ the forward and backward stochastic integrals coincide (see e.g. [Nua06], Relations (3.13)-(3.14) page 176), we have

$$\int_0^T f(t) dB_t = \int_0^T f(t) d^* B_t^*, \tag{5.3}$$

where $B_t^* = B_T - B_t$, $t \in [0,T]$, defines a backward Brownian motion. Next, we let

$$N_T := \int_0^T \int_{-\infty}^{+\infty} y(N(dt,dy) - \mu(t,dy)dt)$$

where $N(dt,dy)$ is a Poisson random measure with compensator $\mu(t,dx)dt$, and set

$$\mathcal{F}_t^{N^*} := \sigma\big(N(A\times[s,T]) : A \in \mathcal{B}(\mathbb{R}),\ t \le s \le T\big), \quad t \in [0,T],$$

for the backward filtration generated by $N(dt,dy)$. Observe that

$$Z_t^* := \int_t^T \int_{-\infty}^{+\infty} y\big(N(ds,dy) - \mu(s,dy)ds\big), \quad t \in [0,T],$$

defines an $(\mathcal{F}_t^{N^*})$-backward martingale. Indeed, we have

$$\begin{aligned}
\mathbb{E}\big[N_T \mid \mathcal{F}_t^{N^*}\big] &= \mathbb{E}\left[\int_0^T \int_{-\infty}^{+\infty} y\big(N(ds,dy) - \mu(s,dy)ds\big) \,\Big|\, \mathcal{F}_t^{N^*}\right] \\
&= \mathbb{E}\left[\int_0^t \int_{-\infty}^{+\infty} y\big(N(ds,dy) - \mu(s,dy)ds\big) \,\Big|\, \mathcal{F}_t^{N^*}\right] \\
&\quad + \int_t^T \int_{-\infty}^{+\infty} y\big(N(ds,dy) - \mu(s,dy)ds\big) && (5.4) \\
&= \mathbb{E}\left[\int_0^t \int_{-\infty}^{+\infty} y\big(N(ds,dy) - \mu(s,dy)ds\big)\right] + Z_t^* && (5.5) \\
&= Z_t^*,
\end{aligned}$$

where we used the facts that $\int_t^T \int_{-\infty}^{+\infty} y\big(N(ds,dy) - \mu(s,dy)ds\big)$ is $\mathcal{F}_t^{N^*}$-measurable in (5.4), and that $\int_0^t \int_{-\infty}^{+\infty} y\big(N(ds,dy) - \mu(s,dy)ds\big)$ is measurable with respect to

$$\sigma\big(N([0,s] \times A) : A \in \mathcal{B}(\mathbb{R}),\ 0 \le s \le t\big),$$

which is independent of $\mathcal{F}_t^{N^*}$, in (5.5). Using (5.3) and $N_T = Z_0^*$, the random variables G writes

$$G = \int_0^T f(t) d^* B_t^* + Z_0^*,$$

and both Corollaries 4.1 and 4.2 apply with $\nu_t^*(\cdot) := \mu(t,\cdot)$, $t \in [0,T]$. □

Next, we specify examples of pure-jump martingale components $(Z_t)_{t\in[0,T]}$ and $(N_t)_{t\in[0,T]}$ appearing in the definitions of F and G in (4.1) and (5.1), and we derive explicit bounds in case those components are given by a jump process, a compound Poisson process, or a Poisson stochastic integral in the framework of Corollary 5.1.

Corollary 5.2 *Consider $(Y_t)_{t\in[0,T]}$ a point process with jumps of size 1 and compensator $(\lambda_t)_{t\in[0,T]}$, and let*

$$F = \int_0^T \sigma_t dB_t + Z_T = \int_0^T \sigma_t dB_t + \int_0^T J_t (dY_t - \lambda_t dt),$$

where $(B_t)_{t\in[0,T]}$ *is a standard Brownian motion,* $(\sigma_t)_{t\in[0,T]}$ *is a process adapted with respect to the filtration generated by* $(B_t)_{t\in[0,T]}$*, and* $(J_t)_{t\in[0,T]}$ *is predictable with respect to the filtration* $(\mathcal{F}_t^Y)_{t\in[0,T]}$ *generated by* $(Y_t)_{t\in[0,T]}$*. Let also*

$$G = \int_0^T f(t)dB_t + \sum_{i=1}^{N_T} U_i - \mathbb{E}[U] \int_0^T \mu(t)dt,$$

where $(U_i)_{i\geq 1} \subset L^2(\Omega)$ *is an i.i.d. sequence distributed as* U*,* $(N_t)_{t\in[0,T]}$ *is a Poisson process with intensity* $(\mu(t))_{t\in[0,T]}$*, and* $(f(t))_{t\in[0,T]}$ *is a deterministic function. Then the distance bounds* $d_{\mathrm{W}_3}(F,G)$ *in* (4.5) *of Corollary 4.1 and* $d_{\mathrm{W}}(F,G)$ *in* (4.7) *of Corollary 4.2-*(3) *are controlled by*

$$\mathbb{E}\left[\int_0^T d_{\mathrm{FM}}\big(\widetilde{\nu}_t, \widetilde{\nu}_t^*\big)dt\right] \leq \mathbb{E}\left[\int_0^T \big||\sigma_t|^2 - |f(t)|^2\big|dt\right] + \mathbb{E}\left[\int_0^T \lambda_t |J_t|^2 |J_t - U| dt\right]$$
$$+\mathbb{E}\left[\int_0^T \lambda_t \big||J_t|^2 - U^2\big|dt\right] + \mathbb{E}\big[U^2\big]\mathbb{E}\left[\int_0^T \big|\lambda_t - \mu(t)\big|dt\right]. \tag{5.6}$$

Proof. In the present setting we have $\nu_t(dx) = \lambda_t \delta_{J_t}(dx)$ and $\mu(t,dx) = \mu(t)\mathbb{P}_U(dx)$ where $\mathbb{P}_U$ is the probability distribution of U, and the measures in (5.2) are given by

$$\widetilde{\nu}_t(dx) = |\sigma_t|^2 \delta_0(dx) + \lambda_t |J_t|^2 \delta_{J_t}(dx), \quad \widetilde{\nu}_t^*(dx) = |f(t)|^2 \delta_0(dx) + \mu(t)x^2 \mathbb{P}_U(dx).$$

Since

$$\int_{-\infty}^{\infty} h(x)\ \widetilde{\nu}_t(dx) = |\sigma_t|^2 h(0) + \lambda_t |J_t|^2 h(J_t)$$

and

$$\int_{-\infty}^{\infty} h(x) \widetilde{\nu}_t^*(dx) = |f(t)|^2 h(0) + \mathbb{E}[U^2 h(U)],$$

we have

$$d_{\mathrm{FM}}\big(\widetilde{\nu}_t, \widetilde{\nu}_t^*\big) = \sup_{h\in\mathcal{FM}} \big|\big(|\sigma_t|^2 - |f(t)|^2\big)h(0) + \lambda_t |J_t|^2 h(J_t) - \mu(t)\mathbb{E}[U^2 h(U)]\big| \tag{5.7}$$
$$\leq \sup_{h\in\mathcal{FM}} \big(\big||\sigma_t|^2 - |f(t)|^2\big||h(0)| + \lambda_t \big||J_t|^2 h(J_t) - \mathbb{E}\big[U^2 h(U)\big]\big| + \big|\lambda_t - \mu(t)\big|\big|\mathbb{E}\big[U^2 h(U)\big]\big|\big). \tag{5.8}$$

Next, regarding the second term in (5.8), we have

$$|J_t|^2 h(J_t) - \mathbb{E}[U^2 h(U)] = \mathbb{E}\big[|J_t|^2 h(J_t) - U^2 h(U)\big|J_t\big]$$

with

$$\big||J_t|^2 h(J_t) - U^2 h(U)\big| \quad = \quad \big||J_t|^2\big(h(J_t) - h(U)\big) + (|J_t|^2 - U^2)h(U)\big|$$

$$\leq \quad |J_t|^2|J_t - U| + \big||J_t|^2 - U^2\big|,$$

which yields (5.6) by plugging the above bound in (5.8). $\square$

By rearranging (5.8) as

$$\begin{aligned} d_{\mathrm{FM}}\big(\widetilde{\nu}_t, \widetilde{\nu}_t^*\big) &= \sup_{h\in\mathcal{FM}} \big|\big(|\sigma_t|^2 - |f(t)|^2\big)h(0) + \lambda_t|J_t|^2 h(J_t) - \mu(t)\mathbb{E}[U^2h(U)]\big| \\ &\leq \sup_{h\in\mathcal{FM}} \big(\big||\sigma_t|^2 - |f(t)|^2\big||h(0)| + \mu(t)\big||J_t|^2h(J_t) - \mathbb{E}\big[U^2h(U)\big]\big| + |\lambda_t - \mu(t)||J_t|^2|h(J_t)|\big), \end{aligned}$$

with

$$\begin{aligned} \big||J_t|^2h(J_t) - U^2h(U)\big| &= \big|U^2(h(J_t) - h(U)) + (|J_t|^2 - U^2)h(J_t)\big| \\ &\leq |U|^2|J_t - U| + \big||J_t|^2 - U^2\big|, \end{aligned}$$

we find that (5.6) can be rewritten as

$$\begin{aligned} \mathbb{E}\left[\int_0^T d_{\mathrm{FM}}\big(\widetilde{\nu}_t, \widetilde{\nu}_t^*\big)dt\right] &\leq \mathbb{E}\left[\int_0^T \big||\sigma_t|^2 - |f(t)|^2\big|dt\right] + \mathbb{E}\left[\int_0^T |\lambda_t - \mu(t)||J_t|^2 dt\right] \\ &\quad + \mathbb{E}\left[\int_0^T \big||J_t|^2 - U^2\big|\mu(t)dt\right] + \mathbb{E}\left[U^2\int_0^T |J_t - U|\mu(t)dt\right]. \end{aligned}$$

As a consequence, the bound (4.7) becomes small when $|\sigma_t|^2$, λ_t and J_t are respectively close to $|f(t)|^2$, $\mu(t)$ and U, uniformly in $t \in [0, T]$.

When G is given by a centered Gauss-Poisson mixture with N a centered Poissonian random variable parameterized by μT, i.e.

$$F = \int_0^T \sigma_t dB_t + \int_0^T J_t(dN_t - \lambda_t dt) \quad \text{and} \quad G = \int_0^T f(t)dB_t + N_T,$$

with $U_i = 1$, i.e. $\mathbb{P}_U(dx) = \delta_1(dx)$, $\mu(t) := \mu \geq 0$ and $(f(t))_{t\in[0,T]}$ in (5.1) deterministic with $\int_0^T f(t)dB_t \sim \mathcal{N}\left(0, \int_0^T |f(t)|^2 dt\right)$, then (5.6) becomes

$$\begin{aligned} \mathbb{E}\left[\int_0^T d_{\mathrm{FM}}\big(\widetilde{\nu}_t, \widetilde{\nu}_t^*\big)dt\right] &\leq \mathbb{E}\left[\int_0^T \big||\sigma_t|^2 - |f(t)|^2\big|dt\right] + \mathbb{E}\left[\int_0^T |J_t|^2|J_t - 1|\lambda_t dt\right] \\ &\quad + \mathbb{E}\left[\int_0^T \big||J_t|^2 - 1\big|\lambda_t dt\right] + \mathbb{E}\left[\int_0^T |\lambda_t - \mu| dt\right]. \end{aligned}$$

The next result deals with the distance of Poisson stochastic integrals with respect to a Gaussian distribution.

Corollary 5.3 *Consider the Poisson stochastic integral*

$$F = \int_0^T J_t(dN_t - \lambda dt)$$

of a $(\sigma(N_s : 0 \le s \le t))_{t\in[0,T]}$*-predictable process* $(J_t)_{t\in[0,T]}$*, where* $(N_t)_{t\in[0,T]}$ *is a Poisson point process with intensity* $(\lambda_t)_{t\in[0,T]}$*, and let* G *denote the Wiener integral*

$$G = \int_0^T f(t)dB_t,$$

where $(f(t))_{t\in[0,T]}$ *is a deterministic function. Then the distance bounds* $d_{\mathrm{W}_3}(F,G)$ *in* (4.5) *of Corollary 4.1 and* $d_{\mathrm{W}}(F,G)$ *in* (4.7) *of Corollary 4.2-(3) are controlled by*

$$\mathbb{E}\left[\int_0^T d_{\mathrm{FM}}\big(\widetilde{\nu}_t, \widetilde{\nu}_t^*\big)dt\right] \le \mathbb{E}\left[\int_0^T \big||f(t)|^2 - \lambda|J_t|^2\big|dt\right] + \lambda\mathbb{E}\left[\int_0^T |J_t|^3 dt\right].$$

Proof. We have $\nu_t(dx) = \lambda\delta_{J_t}(dx)$, $\sigma_t = 0$, $t \in [0,T]$, and $N_T = 0$ and $\mu(t,dx) = 0$. Hence (5.2) reads

$$\widetilde{\nu}_t(dx) = \lambda|J_t|^2\delta_{J_t}(dx), \qquad \widetilde{\nu}_t^*(dx) = |f(t)|^2\delta_0(dx),$$

and for any $h \in \mathcal{FM}$, i.e. $h \in \mathrm{Lip}(1)$ with $\|h\|_\infty \le 1$, from (5.7) we have

$$\begin{aligned} d_{\mathrm{FM}}\big(\widetilde{\nu}_t, \widetilde{\nu}_t^*\big) &= \sup_{h\in\mathcal{FM}} \big|\lambda|J_t|^2h(J_t) - |f(t)|^2h(0)\big| \\ &\le \sup_{h\in\mathcal{FM}} \big(\lambda|J_t|^2\big|h(J_t) - h(0)\big| + \big|\lambda|J_t|^2 - |f(t)|^2\big||h(0)|\big) \\ &\le \lambda|J_t|^3 + \big|\lambda|J_t|^2 - |f(t)|^2\big|. \end{aligned}$$

□

In particular, when $f(t) = 1$, $J_t = 1/\sqrt{n}$ and $\lambda = n$, we find the Poisson to Gaussian convergence bound

$$\mathbb{E}\left[\int_0^T d_{\mathrm{FM}}\big(\widetilde{\nu}_t, \widetilde{\nu}_t^*\big)dt\right] \le \frac{T^{1/3}}{n^{1/6}},$$

but we do not recover the standard Berry-Esseen rate of Corollary 3.4 in [PSTU10] due to the power 1/3 in (4.6), nor the faster rates of e.g. Corollary 5.3 in [Pri18] or [Pri19]. However, those latter results apply only to the case where $(J_t)_{t\in[0,T]}$ is a deterministic function, while the bounds of the present paper have a wider range of applications. In the case of Itô integrals we have the following result which involves only one bounding term in contrast with related results based on the Malliavin calculus, cf. the bound (4.2) in [Pri15].

Corollary 5.4 *Consider the Itô integrals*

$$F = \int_0^T \sigma_t dB_t \quad \text{and} \quad G = \int_0^T f(t) dB_t,$$

where $(\sigma_t)_{t\in[0,T]}$ is adapted with respect to the filtration generated by $(B_t)_{t\in[0,T]}$ and $(f(t))_{t\in[0,T]}$ is a deterministic function. Then the distance bounds $d_{\mathrm{W}_3}(F,G)$ in (4.5) of Corollary 4.1 and $d_{\mathrm{W}}(F,G)$ in (4.7) of Corollary 4.2-(3) are controlled by

$$\mathbb{E}\left[\int_0^T d_{\mathrm{FM}}\big(\widetilde{\nu}_t, \widetilde{\nu}_t^*\big) dt\right] \leq \mathbb{E}\left[\int_0^T \big||\sigma_t|^2 - |f(t)|^2\big| dt\right]. \tag{5.9}$$

Proof. We have $Z_T = N_T = 0$ in (4.1), (4.2), i.e. $\nu_t(dx) = \mu(t,dx) = 0$ and in this case, (5.2) reads

$$\widetilde{\nu}_t(dx) = |\sigma_t|^2 \delta_0(dx), \quad \widetilde{\nu}_t^*(dx) = |f(t)|^2 \delta_0(dx),$$

and we have

$$d_{\mathrm{FM}}\big(\widetilde{\nu}_t, \widetilde{\nu}_t^*\big) = \sup_{h\in\mathcal{FM}} \big||\sigma_t|^2 h(0) - |f(t)|^2 h(0)\big| = \big||\sigma_t|^2 - |f(t)|^2\big|,$$

which allows us to conclude. □

The bound (5.9) is of interest only when $(\sigma_t)_{t\in[0,T]}$ is random, since when $(\sigma(t))_{t\in[0,T]}$ is a deterministic function we can use the more natural inequality

$$d_{\mathrm{W}}(F,G) \leq \left| \int_0^T \sigma^2(t) dt - \int_0^T f^2(t) dt \right|,$$

which implies (5.9) and therefore (4.5). In the non-deterministic case we are not able to recover the above inequality due to the use of a triangle inequality in (3.9). See also (1.3) for a related inequality.

A Appendix

A.1 Approximation Lemma

The following lemma is a generalization of a by-product of Corollary 3.6 in [NPR10].

Lemma A.1 *Let $h \in Lip(1)$ and consider the function h_α defined in (3.16) for $0 < \alpha < 1$. Then we have $h_\alpha \in \mathcal{C}_b^\infty(\mathbb{R})$, with*

$$\|h_\alpha^{(n)}\|_\infty \leq \frac{(1-\alpha)^{n/2}}{\alpha^{(n-1)/2}} \sqrt{(n-1)!}, \qquad n \geq 1.$$

Moreover, for any integrable random variable X we have

$$|\mathbb{E}[h_\alpha(X)] - \mathbb{E}[h(X)]| \le \sqrt{\alpha}\Big(1 + \frac{\mathbb{E}[|X|]}{2}\Big). \tag{A.1}$$

Proof. We start by assuming that h is in $\mathcal{C}^n(\mathbb{R})$ with $\|h'\|_\infty = \|h\|_L \le 1$ and bounded derivatives of orders 1 to n. In this case, an iterated integration by parts with respect to the standard normal density ϕ yields

$$\begin{aligned}
h_\alpha^{(n)}(x) &= (1-\alpha)^{n/2} \int_{-\infty}^{+\infty} h^{(n)}\big(y\sqrt{\alpha} + x\sqrt{1-\alpha}\big)\phi(y)\ dy \\
&= \frac{(1-\alpha)^{n/2}}{\alpha^{(n-1)/2}} \int_{-\infty}^{+\infty} h'\big(y\sqrt{\alpha} + x\sqrt{1-\alpha}\big) H_{n-1}(y)\phi(y)\ dy \\
&\le \frac{(1-\alpha)^{n/2}}{\alpha^{(n-1)/2}} \left(\int_{-\infty}^{+\infty} \big(h'\big(y\sqrt{\alpha} + x\sqrt{1-\alpha}\big)\big)^2 \phi(y)\ dy\right)^{1/2} \left(\int_{-\infty}^{+\infty} (H_{n-1}(y))^2\phi(y)\ dy\right)^{1/2} \\
&\le \frac{(1-\alpha)^{n/2}}{\alpha^{(n-1)/2}} \sqrt{(n-1)!},
\end{aligned}$$

by an application of the Cauchy-Schwarz inequality, where H_{n-1} is the Hermite polynomial of order $n-1 \ge 0$. In case the function h is only Lipschitz, we conclude by approximating h with a sequence of $\mathcal{C}^n$ functions. Finally, the bound (A.1) is obtained as in [NPR10], as follows:

$$\begin{aligned}
|\mathbb{E}[h(X)] - \mathbb{E}[h_\alpha(X)]| &\le \left|\mathbb{E}\left[\int_{-\infty}^{+\infty} h\big(y\sqrt{\alpha} + X\sqrt{1-\alpha}\big) - h\big(X\sqrt{1-\alpha}\big)\phi(y)\ dy\right]\right| \\
&\quad + \big|\mathbb{E}\big[h\big(X\sqrt{1-\alpha}\big) - h(X)\big]\big| \\
&\le \sqrt{\alpha}\|h\|_L \int_{-\infty}^{+\infty} |y|\phi(y)\ dy + \|h\|_L\big|1 - \sqrt{1-\alpha}\big|\mathbb{E}\big[|X|\big] \\
&\le \sqrt{\alpha}\|h\|_L \left(\frac{2}{\sqrt{2\pi}} + \mathbb{E}[|X|]\right)
\end{aligned}$$

using the bound $\big|1 - \sqrt{1-\alpha}\big| \le \sqrt{\alpha}$ for $\alpha \in (0,1)$. □

A.2 Cardano type lemma

The following lemma is based on Cardano's formula for cubic equations.

Lemma A.2 *Let $a > 0$ and $b \in (0, 24a)$. The minimum of $\alpha \mapsto a\sqrt{\alpha} + b/(2\sqrt{\alpha}) + b/(3\alpha)$ is attained at*

$$\alpha_*(a,b) = \left(\sqrt[3]{\frac{b}{3a}\Big(1 + \sqrt{1 - b/(24a)}\Big)} + \sqrt[3]{\frac{b}{3a}\Big(1 - \sqrt{1 - b/(24a)}\Big)}\right)^2 \sim_{b\to 0} \left(\frac{2b}{3a}\right)^{2/3}, \tag{A.2}$$

and is equal to

$$b\frac{1+\sqrt{\alpha_*(a,b)}}{\alpha_*(a,b)} = b\frac{\sqrt[3]{3a/b}+\sqrt[3]{1+\sqrt{1-b/(24a)}}+\sqrt[3]{1-\sqrt{1-b/(24a)}}}{\sqrt[3]{b/(3a)}\Big(\sqrt[3]{1+\sqrt{1-b/(24a)}}+\sqrt[3]{1-\sqrt{1-b/(24a)}}\Big)^2} \tag{A.3}$$

$$\sim_{b\to 0} \sqrt[3]{\frac{9a^2b}{4}},$$

Proof. Letting $\beta = \sqrt{\alpha} > 0$, we have

$$\frac{\partial}{\partial\beta}\Big(a\beta + \frac{b}{2\beta} + \frac{b}{3\beta^2}\Big) = \frac{6a\beta^3 - 3b\beta - 4b}{6\beta^3},$$

where $6a\beta^3 - 3b\beta - 4b$ admits a unique zero $\beta^* \in \mathbb{R}_+$ given by the Cardano formula (when $b < 24a$) as

$$\beta^* = \sqrt[3]{\frac{b}{3a}\big(1+\sqrt{1-b/(24a)}\big)} + \sqrt[3]{\frac{b}{3a}\big(1-\sqrt{1-b/(24a)}\big)},$$

which yields (A.2). The value (A.3) of the maximum follows easily. □

References

[ABD+20] B. Arras, J.-C. Breton, A. Deshayes, O. Durieu, and R. Lachièze-Rey. Some recent advances for limit theorems. *ESAIM: Proceedings and Surveys*, 68:73–96, 2020.

[ABP08] M. Arnaudon, J.-C. Breton, and N. Privault. Convex ordering for random vectors using predictable representation. *Potential Anal.*, 29(4):327–349, 2008.

[AH19] B. Arras and C. Houdré. *On Stein's Method for Infinitely Divisible Laws with Finite First Moment*. SpringerBriefs in Probability and Mathematical Statistics. Springer, New York, 2019.

[BJ00] N. Bellamy and M. Jeanblanc. Incompleteness of markets driven by a mixed diffusion. *Finance and Stochastics*, 4(2):209–222, 2000.

[BLP13] J.-C. Breton, B. Laquerrière, and N. Privault. Convex comparison inequalities for non-Markovian stochastic integrals. *Stochastics*, 85(5):789–806, 2013.

[BP08] J.-C. Breton and N. Privault. Bounds on option prices in point process diffusion models. *Int. J. Theor. Appl. Finance*, 11(6):597–610, 2008.

[BP20a] J.-C. Breton and N. Privault. Integrability and regularity of the flow of stochastic differential equations with jumps. *Theory Probab. Appl.*, 65(1):82–101, 2020.

[BP20b] J.-C. Breton and N. Privault. Wasserstein distance estimates for jump-diffusion processes. Preprint, 22 pages, 2020.

[BR06] J. Bergenthum and L. Rüschendorf. Comparison of option prices in semimartingale models. *Finance and Stochastics*, 10(2):229–249, 2006.

[Dud02] R.M. Dudley. *Real analysis and probability*, volume 74 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2002. Revised reprint of the 1989 original.

[EKJS98] N. El Karoui, M. Jeanblanc, and S. Shreve. Robustness of the Black and Scholes formula. *Math. Finance*, 8(2):93–126, 1998.

[ET07] E. Ekström and J. Tysk. Properties of option prices in models with jumps. *Math. Finance*, 17(3):381–397, 2007.

[JM76] J. Jacod and J. Mémin. Caractéristiques locales et conditions de continuité absolue pour les semi-martingales. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 35(1):1–37, 1976.

[KMP06] Th. Klein, Y. Ma, and N. Privault. Convex concentration inequalities via forward-backward stochastic calculus. *Electron. J. Probab.*, 11:27 pp. (electronic), 2006.

[MP13] Y.T. Ma and N. Privault. Convex concentration for additive functionals of jump stochastic differential equations. *Acta Math. Sin. (Engl. Ser.)*, 29:1449–1458, 2013.

[NP12] I. Nourdin and G. Peccati. *Normal approximations with Malliavin calculus: from Stein's method to universality*, volume 192 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 2012.

[NPR10] I. Nourdin, G. Peccati, and G. Reinert. Stein's method and stochastic analysis of Rademacher functionals. *Electron. J. Probab.*, 15(55):1703–1742, 2010.

[Nua06] D. Nualart. *The Malliavin calculus and related topics*. Probability and its Applications. Springer-Verlag, Berlin, second edition, 2006.

[Pri15] N. Privault. Stein approximation for Itô and Skorohod integrals by Edgeworth type expansions. *Electron. Comm. Probab.*, 20:Article 35, 2015.

[Pri18] N. Privault. Stein approximation for multidimensional Poisson random measures by third cumulant expansions. *ALEA Lat. Am. J. Probab. Math. Stat.*, 15:1141–1161, 2018.

[Pri19] N. Privault. Third cumulant Stein approximation for Poisson stochastic integrals. *J. Theoret. Probab.*, 32:1461–1481, 2019.

[PSTU10] G. Peccati, J. L. Solé, M. S. Taqqu, and F. Utzet. Stein's method and normal approximation of Poisson functionals. *Ann. Probab.*, 38(2):443–478, 2010.