SMALL TIME CENTRAL LIMIT THEOREMS FOR SEMIMARTINGALES WITH APPLICATIONS

STEFAN GERHOLD, MAX KLEINERT, PIET PORKERT,
AND MYKHAYLO SHKOLNIKOV

ABSTRACT. We give conditions under which the normalized marginal distribution of a semimartingale converges to a Gaussian limit law as time tends to zero. In particular, our result is applicable to solutions of stochastic differential equations with locally bounded and continuous coefficients. The limit theorems are subsequently extended to functional central limit theorems on the process level. We present two applications of the results in the field of mathematical finance: to the pricing of at-the-money digital options with short maturities and short time implied volatility skews.

1. INTRODUCTION

Limit theorems for finite-dimensional stochastic processes as time goes to infinity have been a classical object of study in probability theory and many results on the existence and uniqueness of invariant distributions, the convergence of the processes to the latter and the limiting behavior of the fluctuations around the limiting distributions have been obtained (see e.g. [20], [22], [28], [29], [30] and the references therein). More recently, small time asymptotics of finite-dimensional *continuous time* stochastic processes have attracted much attention. Apart from the theoretical interest, these have become of great importance in various applied fields such as mathematical finance, where the increasingly high frequency of trades in financial markets requires pricing models behaving reasonably both on very short and on long time horizons.

In the works [2], [3], [4], [6], [13], [14], [21] and the references therein the authors study the behavior of the random variables $\mathbb{E}[f(X_{t_0+\delta})|\mathcal{F}_{t_0}^X]$ for small values of $\delta > 0$, where X is a finite-dimensional (jump-)diffusion process, a Lévy process or more generally a semimartingale, $(\mathcal{F}_t^X)_{t\geq 0}$ is the filtration it generates and the function f is taken from a space of suitable real-valued test functions. In [4], this program is carried out for general finite-dimensional semimartingales

and under appropriate continuity assumptions on the characteristics of X as well as smoothness assumptions on the function f, the almost sure limit

$$(1.1) \qquad \lim_{\delta \searrow 0} \delta^{-1} \big(\mathbb{E}[f(X_{t_0+\delta})|\mathcal{F}_{t_0}^X] - f(X_{t_0}) \big)$$

is determined.

Here, we are interested in small time Central Limit Theorems for finite dimensional semimartingales; that is, instead of the *almost sure* limit (1.1) we are concerned with the limit

$$(1.2) \qquad \lim_{\delta \searrow 0} \delta^{-1/2} \big(f(X_\delta) - f(X_0) \big)$$

in distribution. More precisely, we give sufficient conditions on the semimartingale X under which, for every suitable test function f, the limit (1.2) exists and is given by a centered normal random variable (whose variance depends on the particular choice of the function f). The most closely related result in the literature seems to be Theorem 2.5 of Doney and Maller [11], which characterizes the *Lévy processes* that satisfy a small time Central Limit Theorem.

In addition to the just described Central Limit Theorems, we prove Functional Central Limit Theorems on the process level and give two applications of our results in the field of mathematical finance: to the pricing of digital options and the asymptotics of implied volatility skews. To outline the first of the two applications, we recall that the price of a digital option with strike K and maturity δ on an underlying security with price process X in the presence of a constant interest rate $r > 0$ is given by the formula

$$(1.3) \qquad \mathbb{E}[e^{-r\delta} \mathbf{1}_{\{X_\delta > K\}}] = e^{-r\delta} \, \mathbb{P}(X_\delta > K).$$

In the limit $\delta \searrow 0$, that is for short maturities, this price tends to 0 if $K > X_0$ (*out-of-the-money* options) and to 1 if $K < X_0$ (*in-the-money* options) as soon as X has right-continuous sample paths. The evaluation of the limit in the case $K = X_0$ (*at-the-money* options) is however much trickier and, in general, the limit can take all values in the interval $[0, 1]$ as we show below. However, if a Central Limit Theorem of the type described above holds for the semimartingale X, then the limit must be given by $\frac{1}{2}$. Moreover, in a special case we can bound the price in (1.3) for any fixed value of $\delta > 0$ from above and below by completely explicit functions tending to $\frac{1}{2}$ in the limit $\delta \searrow 0$. By a well known relation between digital prices and implied volatility skews, we deduce bounds on the latter in certain models with stochastic interest rates.

For the sake of a cleaner exposition, we first give the assumptions on the semimartingale X and state our main results in the case of continuous trajectories.

Assumption 1. *Let $T > 0$, $x_0 \in \mathbb{R}^m$. Let $X = (X_t^1, \dots, X_t^m)_{t \in [0,T]}^\top$ be an $\mathbb{R}^m$-valued continuous semimartingale with canonical decomposition (see e.g. page 337 in [24]) $X - x_0 = M + A$, where M is a continuous local martingale, and A has locally finite variation. Assume that*

(1) $X_0 = x_0$ a.s.;

(2) there exists an a.s. positive stopping time τ_A such that a.s.

$$A_t^j = \int_0^t b_s^j \, \mathrm{d}s, \quad t \in [0, \tau_A], \quad j \in \{1, \ldots, m\},$$

for an adapted process b;

(3) there exists a random variable C_b, such that $|b_t^j| \leq C_b < \infty$ for a.e. $t \in [0, \tau_A]$ a.s., $j \in \{1, \ldots, m\}$;

(4) there exists an a.s. positive stopping time τ_M such that the covariation is a.s.

$$\langle M^j, M^k \rangle_t = \int_0^t \sum_{l=1}^m \sigma_s^{jl} \sigma_s^{kl} \, \mathrm{d}s, \quad t \in [0, \tau_M], \quad j, k \in \{1, \ldots, m\},$$

for a progressive process σ;

(5) there exists a deterministic constant $C_\sigma < \infty$, such that $|\sigma_t^{jk}| \leq C_\sigma$ for a.e. $t \in [0, \tau_M]$ a.s., $j, k \in \{1, \ldots, m\}$;

(6) as $t \searrow 0$, $\sigma_t \to L$ a.s., where L is a deterministic $m \times m$-matrix;

With this notation the Central Limit Theorem and the Functional Central Limit Theorem for continuous semimartingales read as follows.

Theorem 2 (Central Limit Theorem). *Let X satisfy Assumption 1. Then for every $f : \mathbb{R}^m \to \mathbb{R}^n$ such that there exists an open neighborhood U of x_0 with $f \in C^2(U, \mathbb{R}^n)$, we have*

$$\frac{1}{\sqrt{t}}(f(X_t) - f(x_0)) \xrightarrow{d} N_f \text{ as } t \searrow 0,$$

where N_f is a normal random vector with mean 0 and covariance matrix

$$V = (Df)(x_0)L(Df(x_0)L)^\top.$$

Here, $(Df)(x_0)$ stands for the Jacobian of f at x_0.

Theorem 3 (Functional Central Limit Theorem). *Let X satisfy Assumption 1. Then for every $f : \mathbb{R}^m \to \mathbb{R}^n$ such that there exists an open neighborhood U of x_0 with $f \in C^2(U, \mathbb{R}^n)$, the processes*

$$Y^{f,u} := \left(\frac{f(X_{ut}) - f(x_0)}{\sqrt{u}} \right)_{t \in [0,T]}, \quad u \in (0,1),$$

converge in law to a Brownian motion with variance-covariance matrix

$$V = (Df)(x_0)L(Df(x_0)L)^\top.$$

as $u \searrow 0$.

We remark at this point that Assumption 1 is satisfied for weak solutions of stochastic differential equations (SDEs) under minimal regularity assumptions on the coefficients.

Remark 4. *Let X be a weak solution of the m-dimensional SDE*

$$\mathrm{d}X_t^j = b_j(t, X_t)\,\mathrm{d}t + \sum_{k=1}^d \sigma_{jk}(t, X_t)\,\mathrm{d}B_t^k, \quad t \geq 0, \quad j \in \{0, \ldots, m\},$$

$$X_0 = x_0 \ a.s.,$$

where B is a standard d-dimensional Brownian motion, $x_0 \in \mathbb{R}^m$, $b : [0, T] \times \mathbb{R}^m \to \mathbb{R}^m$ is uniformly bounded in a neighborhood of $(0, x_0)$ and $\sigma : [0, T] \times \mathbb{R}^m \to \mathbb{R}^{m \times d}$ is continuous in $(0, x_0)$. Then, X satisfies Assumption 1 and, hence, Theorems 2 and 3 apply.

We also note that, if X satisfies Assumption 1 and the matrix L is non-singular, then the price of an at-the-money digital option in (1.3) (that is, when $K = x_0$) converges to $\frac{1}{2}$ in the limit $\delta \searrow 0$. This result can be significantly sharpened, when X is given by a weak solution of an SDE of the following type.

Theorem 5. *Suppose that the process X solves the stochastic differential equation*

$$(1.4) \qquad\qquad \mathrm{d}X_t = b(t, \cdot)\,\mathrm{d}t + \sigma(t)\,\mathrm{d}B_t,$$

$$X_0 = x_0,$$

where $b : [0, \infty) \times \Omega \to \mathbb{R}^m$ is a bounded predictable process, $\sigma : [0, \infty) \to \mathbb{R}^{m \times m}$ is a locally square integrable function taking values in the set of invertible matrices such that the smallest eigenvalue of $\sigma(\cdot)^\top \sigma(\cdot)$ is uniformly bounded away from 0 and B is a standard m-dimensional Brownian motion. Then, the bounds

$$(1.5) \qquad\qquad e^{f_1(t)} \leq \mathbb{P}\big(X_t^1 > X_0^1\big) \leq e^{f_2(t)}, \quad t > 0$$

apply. Here, the functions f_1, f_2 are given by

$$f_1(t) = -\left(1 + \sqrt{\frac{\|\sigma^{-1}b\|_{2,\infty}^2\, t}{2\log 2}}\right) \cdot \left(\log 2 + \sqrt{\frac{\|\sigma^{-1}b\|_{2,\infty}^2\, t \log 2}{2}}\right),$$

$$f_2(t) = -\left(\sqrt{\frac{2\log 2}{\|\sigma^{-1}b\|_{2,\infty}^2\, t}} - 1\right) \cdot \left(\sqrt{\frac{\|\sigma^{-1}b\|_{2,\infty}^2\, t \log 2}{2}} - \frac{1}{2}\|\sigma^{-1}b\|_{2,\infty}^2\, t\right),$$

where $\|\sigma^{-1}b\|_{2,\infty} = \sup_{t,\omega} |\sigma^{-1}(t)b(t,\omega)|_2$. Moreover, in the limit $t \searrow 0$, the functions e^{f_1}, e^{f_2} admit the series expansions

$$(1.6) \qquad e^{f_1(t)} = \frac{1}{2} - \sqrt{\frac{\log 2}{2}}\|\sigma^{-1}b\|_{2,\infty}\, t^{1/2} + O(t),$$

$$(1.7) \qquad e^{f_2(t)} = \frac{1}{2} + \sqrt{\frac{\log 2}{2}}\|\sigma^{-1}b\|_{2,\infty}\, t^{1/2} + O(t).$$

The rest of the paper is structured as follows. In section 2, we give the proofs of Theorems 2, 5 and 3 in this order. In addition, we provide examples of continuous semimartingales, for which the limit in (1.3) is not $\frac{1}{2}$ and, therefore, the Central Limit Theorem (Theorem 2) cannot hold with a non-degenerate Gaussian law in the limit. In section 3, we state and prove extensions of Theorems 2, 5 and 3 to

semimartingales with jumps. Finally, in section 4, we explain the consequences of these results for the prices of at-the-money digital options with short maturities and the small time asymptotics of implied volatility skews.

2. Continuous Semimartingales

We start with the proof of Theorem 2.

Proof of Theorem 2. Let f be as in the statement of the theorem and let N_f be an $\mathcal{N}(0, V)$ random vector on some probability space $(\widetilde{\Omega}, \widetilde{\mathcal{A}}, \widetilde{\mathbb{P}})$. We need to show

$$(2.1) \qquad \lim_{t \searrow 0} \mathbb{E}\Big[g\Big(\frac{f(X_t) - f(x_0)}{\sqrt{t}}\Big)\Big] = \mathbb{E}_{\tilde{\mathbb{P}}}\big[g(N_f)\big], \quad g \in C_{\mathrm{b}}(\mathbb{R}^n, \mathbb{R}).$$

To this end, we fix a function $g \in C_{\mathrm{b}}(\mathbb{R}^n, \mathbb{R})$, choose an open ball $\boldsymbol{B}$ such that $\overline{\boldsymbol{B}} \subset U$, and define the hitting time $\overline{\tau} := \tau_{\overline{\boldsymbol{B}}^c}$. Then with

$$(2.2) \qquad\qquad\qquad\qquad \tau := \overline{\tau} \wedge \tau_A \wedge \tau_M,$$

we have

$$\Big|\mathbb{E}\Big[g\Big(\frac{f(X_t) - f(x_0)}{\sqrt{t}}\Big)\Big] - \mathbb{E}_{\tilde{\mathbb{P}}}[g(N_f)]\Big|$$

$$\leq \Big|\mathbb{E}\Big[g\Big(\frac{f(X_t) - f(x_0)}{\sqrt{t}}\Big) - g\Big(\frac{f(X_{t \wedge \tau}) - f(x_0)}{\sqrt{t}}\Big)\Big]\Big|$$

$$+ \Big|\mathbb{E}\Big[g\Big(\frac{f(X_{t \wedge \tau}) - f(x_0)}{\sqrt{t}}\Big)\Big] - \mathbb{E}_{\tilde{\mathbb{P}}}[g(N_f)]\Big|.$$

Hence in order to show (2.1), it is sufficent to prove that the two summands in the latter upper bound tend to zero as $t \searrow 0$. Since the event $\{\tau = 0\}$ has probability zero, the first summand converges to zero by the Dominated Convergence Theorem. Moreover, the convergence of the second summand to zero will follow, if we can show

$$(2.3) \qquad\qquad \frac{f(X_{t \wedge \tau}) - f(x_0)}{\sqrt{t}} \xrightarrow{d} N_f, \qquad t \searrow 0.$$

In order to prove (2.3), we first note that Doob's Integral Representation Theorem (see e.g. Theorem 18.12 on page 358 of [24]) in combination with part (4) of Assumption 1 implies the existence of an m-dimensional Brownian motion B (possibly on an extension of the primary probability space) such that a.s.

$$(2.4) \qquad M_{t \wedge \tau}^j = \sum_{k=1}^m \int_0^{t \wedge \tau} \sigma_s^{jk} \, \mathrm{d}B_s^k, \quad t \in [0, T], \quad j \in \{1, \dots, m\}.$$

By part (2) of Assumption 1 and (2.4) we therefore have a.s.

$$(2.5) \quad X_{t \wedge \tau}^j = x_0 + \int_0^{t \wedge \tau} b_s^j \, \mathrm{d}s + \sum_{k=1}^m \int_0^{t \wedge \tau} \sigma_s^{jk} \, \mathrm{d}B_s^k, \quad t \in [0, T], \, j \in \{1 \dots, m\}.$$

In addition, we recall that, by the Cramér–Wold Theorem, (2.3) holds iff for every $s = (s_1, \ldots, s_n)^\top \in \mathbb{R}^n$

$$(2.6) \qquad \sum_{j=1}^{n} s_j \frac{f_j(X_{t \wedge \tau}) - f_j(x_0)}{\sqrt{t}} \xrightarrow{d} \sum_{j=1}^{n} s_j N_f^j$$

as $t \searrow 0$. To show this, we fix $s = (s_1, \ldots, s_n)^\top \in \mathbb{R}^n$. Applying the local Itô formula (see e.g. Corollary 17.19 on page 341 of [24]) in combination with (2.5), we have with $\Psi_t = (\psi_t^{jk})_{1 \le j,k \le m} := \sigma_t \sigma_t^\top$ for all $j \in \{1, \ldots, n\}$:

$$f_j(X_{t \wedge \tau}) - f_j(x_0) = \int_0^{t \wedge \tau} (\mathcal{L}_s f_j)(X_s) \, \mathrm{d}s + \sum_{k,l=1}^{m} \int_0^{t \wedge \tau} \frac{\partial f_j}{\partial x_l}(X_s) \, \sigma_s^{lk} \, \mathrm{d}B_s^k, \quad t \ge 0,$$

where

$$(\mathcal{L}_s f_j)(u) = \frac{1}{2} \sum_{k,l=1}^{m} \psi_s^{kl} \frac{\partial^2 f_j}{\partial x_k \partial x_l}(u) + \sum_{k=1}^{m} b_s^k \frac{\partial f_j}{\partial x_k}(u), \quad u \in U, \quad s \in [0, \tau].$$

Thus, we have for $t > 0$:

$$(2.7) \qquad \begin{aligned} \sum_{j=1}^{n} s_j \frac{f_j(X_{t \wedge \tau}) - f_j(x_0)}{\sqrt{t}} &= \frac{1}{\sqrt{t}} \sum_{j=1}^{n} s_j \int_0^{t \wedge \tau} (\mathcal{L}_s f_j)(X_s) \, \mathrm{d}s \\ &+ \frac{1}{\sqrt{t}} \sum_{j=1}^{n} s_j \sum_{k,l=1}^{m} \int_0^{t \wedge \tau} \frac{\partial f_j}{\partial x_l}(X_s) \, \sigma_s^{lk} \, \mathrm{d}B_s^k. \end{aligned}$$

By parts (3) and (5) of Assumption 1 and the choice of $\boldsymbol{B}$ there exists a random variable $C < \infty$ a.s. such that $\sup_{u \in \boldsymbol{B}} |(\mathcal{L}_s f_j)(u)| \le C$ a.s. for $s \in [0, \tau]$, $j \in \{1, \ldots, n\}$. Thus, we have

$$(2.8) \qquad \left| \frac{1}{\sqrt{t}} \sum_{j=1}^{m} s_j \int_0^{t \wedge \tau} (\mathcal{L}_s f_j)(X_s) \, \mathrm{d}s \right| \le C \sqrt{t} \sum_{j=1}^{m} s_j \to 0, \quad t \searrow 0 \quad \text{a.s.}$$

Before examining the second term on the right-hand side of (2.7) we observe that, for every $t \in [0, T]$, the random vector

$$N_t := \begin{pmatrix} N_t^1 \\ \vdots \\ N_t^n \end{pmatrix} := \begin{pmatrix} \frac{1}{\sqrt{t}} \sum_{k,l=1}^{m} \frac{\partial f_1}{\partial x_l}(x_0) L_{lk} B_t^k \\ \vdots \\ \frac{1}{\sqrt{t}} \sum_{k,l=1}^{m} \frac{\partial f_n}{\partial x_l}(x_0) L_{lk} B_t^k \end{pmatrix}$$

is $\mathcal{N}(0, V)$ distributed. In particular, the distribution of N_t is independent of t, and $N_t \stackrel{d}{=} N_f$ for every $t > 0$. With $\xi_{jk} := \sum_{l=1}^{m} \frac{\partial f_j}{\partial x_l}(x_0) L_{lk}$ we have for all

$h \in C_b(\mathbb{R}, \mathbb{R})$:

$$\left| \mathbb{E}\Big[h\Big(\frac{1}{\sqrt{t}} \sum_{j=1}^{n} s_j \sum_{k=1}^{m} \xi_{jk} B_{t \wedge \tau}^k \Big) \Big] - \mathbb{E}_{\tilde{\mathbb{P}}}\Big[h\Big(\sum_{j=1}^{n} s_j N_f^j \Big) \Big] \right|$$

$$= \left| \mathbb{E}\Big[h\Big(\frac{1}{\sqrt{t}} \sum_{j=1}^{n} s_j \sum_{k=1}^{m} \xi_{jk} B_{t \wedge \tau}^k \Big) \Big] - \mathbb{E}\Big[h\Big(\frac{1}{\sqrt{t}} \sum_{j=1}^{n} s_j \sum_{k=1}^{m} \xi_{jk} B_t^k \Big) \Big] \right|$$

$$\leq \left| \mathbb{E}\Big[\Big(h\Big(\frac{1}{\sqrt{t}} \sum_{j=1}^{n} s_j \sum_{k=1}^{m} \xi_{jk} B_{t \wedge \tau}^k \Big) - h\Big(\frac{1}{\sqrt{t}} \sum_{j=1}^{n} s_j \sum_{k=1}^{m} \xi_{jk} B_t^k \Big) \Big) \mathbf{1}_{\{\tau > t\}} \Big] \right|$$

$$+ \left| \mathbb{E}\Big[\Big(h\Big(\frac{1}{\sqrt{t}} \sum_{j=1}^{n} s_j \sum_{k=1}^{m} \xi_{jk} B_{t \wedge \tau}^k \Big) - h\Big(\frac{1}{\sqrt{t}} \sum_{j=1}^{n} s_j \sum_{k=1}^{m} \xi_{jk} B_t^k \Big) \Big) \mathbf{1}_{\{\tau \leq t\}} \Big] \right|$$

$$\leq 2\|h\|_\infty \, \mathbb{P}(\tau \leq t) \to 0$$

as $t \searrow 0$. Therefore, the random variables

$$(2.9) \qquad \frac{1}{\sqrt{t}} \sum_{j=1}^{n} s_j \sum_{k,l=1}^{m} \int_0^{t \wedge \tau} \frac{\partial f_j}{\partial x_l}(x_0) L_{lk} \, \mathrm{d}B_s^k = \frac{1}{\sqrt{t}} \sum_{j=1}^{n} s_j \sum_{k,l=1}^{m} \frac{\partial f_j}{\partial x_l}(x_0) L_{lk} \, B_{t \wedge \tau}^k$$

converge in distribution to $\sum_{j=1}^{n} s_j N_f^j$ as $t \searrow 0$. Next, we show that the difference between (2.9) and the second term on the right-hand side of (2.7) converges to zero in L^2. By the Cauchy–Schwarz inequality and Itô's isometry we have

$$\mathbb{E}\left[\Big(\sum_{j=1}^{n} s_j \sum_{k,l=1}^{m} \frac{1}{\sqrt{t}} \int_0^{t \wedge \tau} \Big(\frac{\partial f_j}{\partial x_l}(X_s)\sigma_s^{lk} - \frac{\partial f_j}{\partial x_l}(x_0)L_{lk} \Big) \mathrm{d}B_s^k \Big)^2 \right]$$

$$\leq n\, m^2 \sum_{j=1}^{n} s_j^2 \sum_{k,l=1}^{m} \mathbb{E}\left[\frac{1}{t} \Big(\int_0^{t \wedge \tau} \Big(\frac{\partial f_j}{\partial x_l}(X_s)\sigma_s^{lk} - \frac{\partial f_j}{\partial x_l}(x_0)L_{lk} \Big) \mathrm{d}B_s^k \Big)^2 \right]$$

$$= n\, m^2 \sum_{j=1}^{n} s_j^2 \sum_{k,l=1}^{m} \mathbb{E}\left[\frac{1}{t} \int_0^{t \wedge \tau} \Big(\frac{\partial f_j}{\partial x_l}(X_s)\sigma_s^{lk} - \frac{\partial f_j}{\partial x_l}(x_0)L_{lk} \Big)^2 \mathrm{d}s \right]$$

$$\leq n\, m^2 \sum_{j=1}^{n} s_j^2 \sum_{k,l=1}^{m} \mathbb{E}\left[\frac{t \wedge \tau}{t} \max_{s \in [0, t \wedge \tau]} \Big(\frac{\partial f_j}{\partial x_l}(X_s)\sigma_s^{lk} - \frac{\partial f_j}{\partial x_l}(x_0)L_{lk} \Big)^2 \right],$$

which indeed converges to zero as $t \searrow 0$ by the Dominated Convergence Theorem. The just established L^2 convergence implies convergence in distribution. Summarizing, we have in the limit $t \searrow 0$:

$$\left| \frac{1}{\sqrt{t}} \sum_{j=1}^{n} s_j \int_0^{t \wedge \tau} (\mathcal{L}_s f_j)(X_s) \, \mathrm{d}s \right| \to 0 \quad \text{a.s.,}$$

$$\sum_{j=1}^{n} s_j \sum_{k,l=1}^{m} \frac{1}{\sqrt{t}} \int_0^{t \wedge \tau} \frac{\partial}{\partial x_l} f_j(X_s)\sigma_s^{lk} \, \mathrm{d}B_s^k \xrightarrow{d} \sum_{j=1}^{n} s_j N_f^j,$$

so that by Slutsky's theorem (2.6) readily follows. $\qquad \square$

Remark 6. *Recall that, if a family of probability measures satisfies a large deviations principle (LDP) with a rate function I, then the validity of a CLT is related to the second derivative of I (see section 1.4 in [10] for a discussion in the case of Cramér's theorem). We now (heuristically) outline this connection in a very simple instance of our setup. Suppose that X satisfies a one-dimensional SDE (with zero drift for simplicity)*

$$X_t = x_0 + \int_0^t \sigma(X_s)\, \mathrm{d}B_s, \quad t \geq 0,$$

where σ is bounded, bounded away from zero and Lipschitz continuous. Then, due to the time-change formalism for one-dimensional diffusions (see e.g. Theorem 8.5.1 on page 148 in [32]), for each $\delta > 0$, we can view the random variable X_δ as the value of the diffusion

$$X_t^{(\delta)} = x_0 + \sqrt{\delta} \int_0^t \sigma(X_s^{(\delta)})\, \mathrm{d}W_s$$

at time 1, where W is the appropriate standard Brownian motion. Now, using the remark following Theorem 5.6.7 in [9] on page 214, and the contraction principle (see Theorem 4.2.1 on page 126 of [9]), we conclude that the random variables X_t satisfy an LDP as $t \searrow 0$ with rate function

$$I(x_0 + \varepsilon) = \frac{1}{2} \inf_{\substack{f \in H^1([0,1]): \\ f(0)=x_0, \\ f(1)=x_0+\varepsilon}} \int_0^1 \frac{\dot{f}(s)^2}{\sigma(f(s))^2}\, \mathrm{d}s.$$

Next, let Σ be an antiderivative of the function $1/\sigma$. By the assumptions on σ, the function $\Sigma(f(\cdot))$ belongs to $H^1([0,1])$ if and only if the function f belongs to $H^1([0,1])$. Hence, the latter infimum can be rewritten as

$$\inf_{\substack{v \in H^1([0,1]): \\ v(0)=\Sigma(x_0), \\ v(1)=\Sigma(x_0+\varepsilon)}} \int_0^1 \dot{v}(s)^2\, \mathrm{d}s.$$

Due to Jensen's inequality, the infimum is reached when v is the affine function connecting $\Sigma(x_0)$ and $\Sigma(x_0 + \varepsilon)$. Plugging it in, we end up with

$$(2.10) \qquad I(x_0 + \varepsilon) = \frac{1}{2}[\Sigma(x_0 + \varepsilon) - \Sigma(x_0)]^2 = \frac{1}{2}\left(\int_{x_0}^{x_0+\varepsilon} \frac{\mathrm{d}u}{\sigma(u)}\right)^2.$$

That is, for $\varepsilon > 0$ small and fixed, we have the asymptotics

$$(2.11) \qquad \mathbb{P}(X_t \geq x_0 + \varepsilon) \simeq \exp(-I(x_0 + \varepsilon)/t),$$

where $\simeq$ stands for exponential equivalence. Now, pretend that we can apply the LDP (2.11) with a time-dependent ε defined by $\varepsilon = z\sqrt{t}$, where $z > 0$. Since $I(x_0) = I'(x_0) = 0$, we have

$$I(x_0 + z\sqrt{t}) = \tfrac{1}{2}I''(x_0)z^2 t + o(t),$$

and so

$$\mathbb{P}\left(\frac{X_t - x_0}{\sqrt{t}} \geq z\right) \simeq \exp\left(-\frac{z^2}{2\sigma(x_0)^2} + o(1)\right).$$

The Gaussian limit law is thus correctly identified by this heuristic argument (the case $z < 0$ is similar).

If a semimartingale X satisfies Assumption 1, and the limit law in Theorem 2 is non-degenerate, we clearly have

(2.12) $$\lim_{t \searrow 0} \mathbb{P}(X_t > x_0) = \frac{1}{2}.$$

We now give some examples where the value of this limit is not $1/2$.

Example 7. *Let us consider the squared Brownian motion B^2 in one dimension (no confusion with our superindex convention should arise). Then clearly $\lim_{t \searrow 0} \mathbb{P}(B_t^2 > 0) = 1$, which does not contradict Theorem 2. Indeed, the martingale part in the canonical decomposition of B^2 is $B_t^2 - t = 2\int_0^t B_s \, dB_s$, which leads to*

(2.13) $$\langle B_t^2 - t \rangle = 4\int_0^t B_s^2 \, ds \to 0, \quad t \searrow 0 \quad a.s.$$

Since all items of Assumption 1 are satisfied, Theorem 2 tells us that $\frac{1}{\sqrt{t}} B_t^2$ converges in distribution to a degenerate normal random variable.

Example 8. *Denoting by Φ the standard normal cumulative distribution function, we see that for any $p \in (0, 1)$ and a standard Brownian motion B, the continuous process $X_t = B_t + \Phi^{-1}(p)\sqrt{t}$ satisfies $\mathbb{P}(X_t > 0) = p$ for all $t \geq 0$. (Although not related to the present topic, we recall that the process $B_t = W_t + \sqrt{t}$ occurs in Example 3.4 of Delbaen and Schachermayer [8]. They show that, when used as the price process of a financial security, X_t (and also $\exp(X_t)$) allows for immediate arbitrage; the arbitrage disappears if proportional transaction costs are introduced [19, Example 4.1].)*

The following example shows that each probability $p \in [0, 1)$ can even be realized by a continuous *martingale*. (Note also that that the *non-continuous* martingale $t - P_t$, where P_t is a Poisson process with parameter 1, satisfies $\lim_{t \searrow 0} \mathbb{P}(t - P_t > 0) = 1$.)

Example 9. *Consider the squared Bessel process of dimension $\delta \geq 0$, that is, the strong solution of the SDE*

$$dR_t^\delta = 2\sqrt{R_t^\delta} \, dB_t + \delta \, dt$$

with initial value $R_0^\delta = 0$. Then, the process $R_t^\delta - \delta t$, $t \geq 0$ is a martingale. We claim that

for all $p \in [0, 1)$, there is a $\delta \in [0, \infty)$ such that $\lim_{t \searrow 0} \mathbb{P}(R_t^\delta - \delta t > 0) = p$.

The scaling property of squared Bessel processes (see section 1 in chapter XI of [33]) shows

$$\lim_{t \searrow 0} \mathbb{P}(R_t^\delta - \delta t > 0) = \mathbb{P}(R_1^\delta > \delta).$$

We show now that when one varies δ in $[0, \infty)$, the right-hand side achieves all values $p \in [0, 1)$. For $\delta > 0$, the random variable R_1^δ has the gamma distribution with shape parameter $\delta/2$ and scale parameter 2 (see Corollary 1.4 in section 1 of chapter XI in [33]). In particular, it has mean δ and variance 2δ. We claim that

(2.14) $$\lim_{\delta \searrow 0} \mathbb{P}(R_1^\delta > \delta) = 0.$$

Let $\varepsilon > 0$ be arbitrary. By Chebyshev's inequality, we have

$$\mathbb{P}(R_1^\delta > \delta + \varepsilon) \leq \frac{2\delta}{\varepsilon^2},$$

and so $\lim_{\delta \searrow 0} \mathbb{P}(R_1^\delta > \delta + \varepsilon) = 0$. Therefore, recalling that, for any fixed $0 < \delta \leq 2$, the density function of R_1^δ is strictly decreasing, we have the estimates

$$\lim_{\delta \searrow 0} \mathbb{P}(R_1^\delta > \delta) = \lim_{\delta \searrow 0} \mathbb{P}(\delta + \varepsilon > R_1^\delta > \delta)$$

$$\leq \varepsilon \lim_{\delta \searrow 0} 2^{-\delta/2} \Gamma(\delta/2)^{-1} \delta^{\delta/2 - 1} e^{-\delta/2}$$

$$= \varepsilon \lim_{\delta \searrow 0} \frac{\delta^{\delta/2 - 1}}{\Gamma(1 + \delta/2)/(\delta/2)} = \frac{\varepsilon}{2}.$$

Thus, taking the limit $\varepsilon \searrow 0$, we end up with (2.14). For $\delta \to \infty$, the random variables $(R_1^\delta - \delta)/(2\delta)^{1/2}$ converge in distribution to a standard normal random variable [23, p. 340]. This implies

$$\lim_{\delta \to \infty} \mathbb{P}(R_1^\delta > \delta) = \frac{1}{2}.$$

It now follows from the Intermediate Value Theorem that, for every $p \in [0, \frac{1}{2})$, we can find a $\delta \geq 0$ such that

$$\lim_{t \searrow 0} \mathbb{P}(R_t^\delta - \delta t > 0) = p.$$

(Note that for $\delta = 0$, we have $R_t^\delta \equiv 0$, and so $\lim_{t \searrow 0} \mathbb{P}(R_t^\delta - \delta t > 0) = 0$.) Finally, by considering the martingales $\delta t - R_t^\delta$, we see that all values $p \in [0, 1)$ can be achieved.

We now take a look at higher order terms beyond the limit in (2.12). If $X_t = B_t + bt$ is a one-dimensional Brownian motion with drift $b \in \mathbb{R}$, we have $\mathbb{P}(X_t > x_0) = \frac{1}{2} + O(t^{1/2})$. Theorem 5, which we prove now, shows that this estimate persists for a larger class of Itô processes.

Proof of Theorem 5. Fix a $t > 0$ and make a change of probability measure according to the Girsanov Theorem, with the corresponding density being given

by

$$\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}} := Z_t^{-1} := e^{-N_t - \frac{1}{2}\langle N \rangle_t}$$

$$:= \exp\Big(-\int_0^t \sigma(s)^{-1} b(s,\cdot)\, \mathrm{d}B_s - \frac{1}{2}\int_0^t |\sigma(s)^{-1} b(s,\cdot)|_2^2 \,\mathrm{d}s \Big).$$

Under $\mathbb{Q}$, the process X solves the equation

$$(2.15) \qquad\qquad \mathrm{d}X_s = \sigma(s)\, \mathrm{d}B_s^{\mathbb{Q}}$$

on $[0,t]$ with initial condition $X_0 = x_0$ and where $B^{\mathbb{Q}}$ is a standard Brownian motion under $\mathbb{Q}$. Thus, $\mathbb{Q}(X_s^1 > X_0^1) = \frac{1}{2}$ for all $s \in (0,t]$. Moreover,

$$(2.16) \qquad\qquad \mathbb{P}(X_t^1 > X_0^1) = \mathbb{E}^{\mathbb{Q}}\big[Z_t \, \mathbf{1}_{\{X_t^1 > X_0^1\}} \big].$$

To obtain upper and lower bounds on the latter expression, we fix numbers $p,q > 1$ such that $p^{-1} + q^{-1} = 1$ and apply Hölder's inequality to deduce

$$\mathbb{Q}\big(X_t^1 > X_0^1\big) = \mathbb{E}^{\mathbb{Q}}\big[\mathbf{1}_{\{X_t^1 > X_0^1\}} Z_t^{1/p} Z_t^{-1/p} \big]$$

$$\leq \mathbb{E}^{\mathbb{Q}}\big[\mathbf{1}_{\{X_t^1 > X_0^1\}} Z_t \big]^{1/p} \mathbb{E}^{\mathbb{Q}}\big[Z_t^{-q/p} \big]^{1/q}.$$

Taking the p-th power and rearranging, we get

$$\mathbb{Q}\big(X_t^1 > X_0^1\big)^p \mathbb{E}^{\mathbb{Q}}\big[Z_t^{-q/p} \big]^{-p/q} \leq \mathbb{E}^{\mathbb{Q}}\big[\mathbf{1}_{\{X_t^1 > X_0^1\}} Z_t \big]$$

$$\leq \mathbb{Q}\big(X_t^1 > X_0^1\big)^{1/q} \mathbb{E}^{\mathbb{Q}}\big[Z_t^p \big]^{1/p},$$

where the last upper bound follows again by Hölder's inequality. This can be simplified to

$$\Big(\frac{1}{2}\Big)^p \mathbb{E}^{\mathbb{Q}}\big[Z_t^{-q/p} \big]^{-p/q} \leq \mathbb{P}(X_t^1 > X_0^1) \leq \Big(\frac{1}{2}\Big)^{1/q} \mathbb{E}^{\mathbb{Q}}\big[Z_t^p \big]^{1/p},$$

or

$$(2.17) \qquad \Big(\frac{1}{2}\Big)^p \mathbb{E}^{\mathbb{P}}\big[Z_t^{-q/p-1} \big]^{-p/q} \leq \mathbb{P}(X_t^1 > X_0^1) \leq \Big(\frac{1}{2}\Big)^{1/q} \mathbb{E}^{\mathbb{P}}\big[Z_t^{p-1} \big]^{1/p}.$$

To estimate the bounds further, we note that

$$Z_t^{-q/p-1} = e^{-(q/p+1)N_t - \frac{1}{2}(q/p+1)^2 \langle N \rangle_t} \cdot \exp\Big(\frac{1}{2}\Big(\frac{q}{p}+1\Big)\frac{q}{p} \langle N \rangle_t \Big)$$

$$(2.18) \qquad \leq e^{-(q/p+1)N_t - \frac{1}{2}(q/p+1)^2 \langle N \rangle_t} \cdot \exp\Big(\frac{1}{2}\Big(\frac{q}{p}+1\Big)\frac{q}{p} t \, \|\sigma^{-1} b\|_{2,\infty}^2 \Big)$$

and

$$Z_t^{p-1} = e^{(p-1)N_t - \frac{1}{2}(p-1)^2 \langle N \rangle_t} \cdot \exp\Big(\frac{1}{2}(p-1)p \langle N \rangle_t \Big)$$

$$(2.19) \qquad \leq e^{(p-1)N_t - \frac{1}{2}(p-1)^2 \langle N \rangle_t} \cdot \exp\Big(\frac{1}{2}(p-1)p\, t \, \|\sigma^{-1} b\|_{2,\infty}^2 \Big).$$

(Recall that we write $\|\sigma^{-1}b\|_{2,\infty}$ for $\sup_{t,x}|\sigma^{-1}b(t,x)|_2$.) The first factors in (2.18) resp. (2.19) are $\mathbb{P}$-martingales, since Novikov's condition is satisfied by our assumptions on b and σ. Therefore, inserting these estimates into (2.17), we obtain

$$
\sup_{\substack{p^{-1}+q^{-1}=1 \\ p>1}} \left(\frac{1}{2}\right)^p \exp\left(-\frac{1}{2}\left(\frac{q}{p}+1\right)t\,\|\sigma^{-1}b\|_{2,\infty}^2\right)
$$

$$
\leq \mathbb{P}(X_t^1 > X_0^1)
$$

$$
\leq \inf_{\substack{p^{-1}+q^{-1}=1 \\ p>1}} \left(\frac{1}{2}\right)^{1/q} \exp\left(\tfrac{1}{2}(p-1)\,t\,\|\sigma^{-1}b\|_{2,\infty}^2\right).
$$

It is easy to see that the lower bound is maximized by

$$
p = 1 + \sqrt{\frac{\|\sigma^{-1}b\|_{2,\infty}^2\,t}{2\log 2}}\,,
$$

whereas the upper bound is minimized by

$$
p = \sqrt{\frac{2\log 2}{\|\sigma^{-1}b\|_{2,\infty}^2\,t}}\,,
$$

which together give (1.5). Finally, the expansions given in the statement of the theorem can be computed by Taylor expansions of the explicit functions in the lower and upper bounds. $\qquad\square$

We conclude this section with the proof of Theorem 3.

Proof of Theorem 3. Let $\widetilde{B}$ be a Brownian motion with variance-covariance matrix V and let $(u_l)_{l\in\mathbb{N}}$ be a sequence with elements in $(0,1)$ such that $u_l \searrow 0$ as $l \to \infty$. It is sufficient to verify the convergence of the finite-dimensional distributions

$$
(2.20)\qquad (Y_{t_1}^{f,u_l},\ldots,Y_{t_w}^{f,u_l}) \xrightarrow{d} (\widetilde{B}_{t_1},\ldots,\widetilde{B}_{t_w}), \quad t_1,\ldots,t_w \in [0,T], \quad w \in \mathbb{N},
$$

and the tightness condition

$$
(2.21)\qquad \lim_{\delta\searrow 0}\,\overline{\lim_{l\to\infty}}\,\mathbb{P}\Big(\sup_{|s-t|\leq\delta}|Y_s^{f,u_l}-Y_t^{f,u_l}|>\varepsilon\Big)=0, \quad \varepsilon>0.
$$

Indeed, by Theorem 1.3.2 in [35], condition (2.21) implies the tightness of the laws of Y^{f,u_l}, $l \in \mathbb{N}$. Moreover, the convergence (2.20) allows to the identify the limit points with the law of $\widetilde{B}$.

First, we focus on (2.20). Fix $t_1,\ldots,t_w \in [0,T]$ for some $w \in \mathbb{N}$; then by the Cramér–Wold theorem it suffices to show

$$
\sum_{d=1}^{w}\sum_{j=1}^{n} s_{dj}(Y_{t_d}^{f,u_l})^j \xrightarrow{d} \sum_{d=1}^{w}\sum_{j=1}^{n} s_{dj}\widetilde{B}_{t_d}^j
$$

for all $s \in \mathbb{R}^{w \times n}$ as $l \to \infty$. Let τ be defined as in (2.2). Arguing as in the proof of Theorem 2, we see that it is enough to show

$$\sum_{d=1}^{w} \sum_{j=1}^{n} s_{dj} (Y_{t_d \wedge \tau}^{f,u_l})^j \xrightarrow{d} \sum_{d=1}^{w} \sum_{j=1}^{n} s_{dj} \widetilde{B}_{t_d}^j$$

as $l \to \infty$. However, this can be proven analogously to (2.6).

To show (2.21), note that we may work with the stopped processes $Y_{t \wedge \tau}^{f,u_l}$, $l \in \mathbb{N}$. Indeed, since τ, as defined in (2.2), is a.s. positive, we have

$$\varlimsup_{l \to \infty} \mathbb{P}\Big(\sup_{t \in [0,T]} |Y_t^{f,u_l} - Y_{t \wedge \tau}^{f,u_l}| > \varepsilon \Big) = 0, \quad \varepsilon > 0.$$

The triangle inequality thus shows that (2.21) is implied by

$$(2.22) \quad \lim_{\delta \searrow 0} \varlimsup_{l \to \infty} \mathbb{P}\Big(\sup_{|s-t| \leq \delta} |(Y_{s \wedge \tau}^{f,u_l})^j - (Y_{t \wedge \tau}^{f,u_l})^j| > \varepsilon \Big) = 0, \quad \varepsilon > 0, j \in \{1, \ldots, n\}.$$

By (2.7) we get

(2.23)
$$\mathbb{P}\Big(\sup_{|s-t| \leq \delta} |(Y_{s \wedge \tau}^{f,u_l})^j - (Y_{t \wedge \tau}^{f,u_l})^j| > \varepsilon \Big) \leq \mathbb{P}\Big(\sup_{|s-t| \leq \delta} \frac{1}{\sqrt{u_l}} \int_{u_l(s \wedge \tau)}^{u_l(t \wedge \tau)} |(\mathcal{L}_r f_j)(X_r)| \, dr > \frac{\varepsilon}{2} \Big)$$
$$+ \mathbb{P}\Big(\sup_{|s-t| \leq \delta} \Big| \sum_{k,v=1}^{m} \int_{u_l(s \wedge \tau)}^{u_l(t \wedge \tau)} \frac{\partial f_j}{\partial x_v}(X_r) \sigma_r^{vk} \, dB_r^k \Big| > \frac{\varepsilon \sqrt{u_l}}{2} \Big).$$

According to (2.8), we have

$$\mathbb{P}\Big(\sup_{|s-t| < \delta} \frac{1}{\sqrt{u_l}} \int_{u_l(s \wedge \tau)}^{u_l(t \wedge \tau)} |(\mathcal{L}_r f_j)(X_r)| \, dr > \frac{\varepsilon}{2} \Big) \leq \mathbb{P}\Big(C\sqrt{u_l}\, \delta > \frac{\varepsilon}{2} \Big) \xrightarrow{l \to \infty} 0, \quad \varepsilon > 0.$$

We now investigate the second term on the right-hand side of (2.23). After fixing δ, j and l, we define the process

$$F_t := \sum_{k,v=1}^{m} \int_0^{u_l(t \wedge \tau)} \frac{\partial f_j}{\partial x_v}(X_r) \sigma_r^{vk} \, dB_r^k, \quad t \in [0,T].$$

In addition, we introduce the processes

$$G_t^i := F_{i\delta+t} - F_{i\delta}, \quad t \in I_i := [0, \delta], \quad i \in \{0, \ldots, \lfloor T/\delta \rfloor - 1\},$$

and for $i = \lfloor T/\delta \rfloor$,

$$G_t^{\lfloor T/\delta \rfloor} := F_{\lfloor T/\delta \rfloor \delta + t} - F_{\lfloor T/\delta \rfloor \delta}, \quad t \in I_{\lfloor T/\delta \rfloor} := [0, T - \lfloor T/\delta \rfloor].$$

These are continuous local martingales and, thus, each of them can be represented as a time changed Brownian motion (see e.g. Theorem 18.4 on page 352 of [24]): $G_t^i = W_{\langle G^i \rangle_t}^i$. Moreover, the quadratic variation of G^i can be bounded according to

$$\langle G^i \rangle_t \leq \gamma C_\sigma^2 u_l \delta, \quad t \in I_i, \quad i \in \{0, \ldots, \lfloor T/\delta \rfloor\},$$

where $0 < \gamma < \infty$ only depends on m and the Jacobian of f on the ball $\overline{\boldsymbol{B}}$ (see the paragraph preceding (2.2) for the definition of the latter). Now, consider the event $\{\sup_{|t-s|<\delta} |F_t - F_s| > \frac{\varepsilon\sqrt{u_l}}{2}\}$. Clearly, on this event there exist s_0, $t_0 \in [0, T]$ such that $|s_0 - t_0| \le \delta$ and $|F_{t_0} - F_{s_0}| > \frac{\varepsilon\sqrt{u_l}}{2}$. Without loss of generality we may assume that $0 \le s_0 < \delta \le t_0 < 2\delta$ (the other cases can be dealt with in the same manner). Then, either $|F_\delta - F_{s_0}| > \frac{\varepsilon\sqrt{u_l}}{4}$, or $|F_{t_0} - F_\delta| > \frac{\varepsilon\sqrt{u_l}}{4}$. In the first case we get

$$(2.24) \qquad \frac{\varepsilon\sqrt{u_l}}{4} < |F_\delta - F_{s_0}| \le |F_{s_0} - F_0| + |F_\delta - F_0| \le 2 \sup_{r \in [0,\delta]} |F_r - F_0|.$$

In the second case we have

$$(2.25) \qquad \frac{\varepsilon\sqrt{u_l}}{4} < |F_{t_0} - F_\delta| \le \sup_{r \in [0,\delta]} |F_{\delta+r} - F_\delta|.$$

These considerations show that on the event $\{\sup_{|t-s|<\delta} |F_t - F_s| > \frac{\varepsilon\sqrt{u_l}}{2}\}$ there exists an index $i \in \{0, \dots \lfloor T/\delta \rfloor\}$ such that $\sup_{t \in I_i} |G_t^i| > \frac{\varepsilon\sqrt{u_l}}{8}$. Putting everything together we obtain

$$\mathbb{P}\left(\sup_{|s-t|<\delta} \left| \sum_{k,v=1}^m \int_{u_l(s \wedge \tau)}^{u_l(t \wedge \tau)} \frac{\partial f_j}{\partial x_v}(X_r) \sigma_r^{vk} \, dB_r^k \right| > \frac{\varepsilon\sqrt{u_l}}{2} \right)$$

$$\le \mathbb{P}\left(\sup_{t \in I_i} |G_t^i| > \frac{\varepsilon\sqrt{u_l}}{8} \text{ for at least one } i \right)$$

$$\le \sum_{i=0}^{\lfloor T/\delta \rfloor} \mathbb{P}\left(\sup_{t \in I_i} |G_t^i| > \frac{\varepsilon\sqrt{u_l}}{8} \right)$$

$$\le \left(\frac{T}{\delta} + 1 \right) \mathbb{P}\left(\sup_{0 \le r \le \gamma C_\sigma^2 u_l \delta} |W_r^i| > \frac{\varepsilon\sqrt{u_l}}{8} \right)$$

$$\le \left(\frac{T}{\delta} + 1 \right) \exp\left(-\frac{\varepsilon^2}{128 \, \gamma \, C_\sigma^2 \delta} \right) \xrightarrow{\delta \to 0} 0.$$

Here, the last estimate follows from Bernstein's inequality (see e.g. Exercise 3.16 on page 153 of [33]). We have established (2.22) and, thus, the proof is finished. □

3. Semimartingales with Jumps

This section is devoted to the extensions of Theorems 2, 3 and 5 to semimartingales with jumps. We start by stating the assumptions on the semimartingale X, which will replace Assumption 1 when jumps are present.

Assumption 10. *For a $T > 0$, let $X = (X_t^1, \dots, X_t^m)_{t \in [0,T]}^\top$ be an $\mathbb{R}^m$-valued semimartingale with decomposition $X = X^c + J$, such that*

(1) X^c is a continuous semimartingale satisfying Assumption 1;

(2) the process J is given by

$$J_t = \int_0^t \int_{\boldsymbol{B}_1} \psi(s,z) \left(\Pi(\mathrm{d}s, \mathrm{d}z) - \mu(\mathrm{d}s, \mathrm{d}z) \right) + \int_0^t \int_{\mathbb{R}^m \setminus \boldsymbol{B}_1} \varphi(s,z) \, \Pi(\mathrm{d}s, \mathrm{d}z),$$

where $\boldsymbol{B}_1$ denotes the unit ball in $\mathbb{R}^m$, Π is a Poisson random measure on $[0,T] \times \mathbb{R}^m$ with compensator μ; the $\mathbb{R}^m$-valued processes ψ, φ are predictable with respect to the filtration generated by Π and

$$\mathbb{E}\left[\int_0^T \int_{\boldsymbol{B}_1} |\psi(s,z)|^2 \, \mu(\mathrm{d}s, \mathrm{d}z) \right] < \infty;$$

(3) There exists an a.s. positive stopping time τ_J such that

$$\mathbb{E}\left[\left| \Pi - \mu \right| \left([0, t \wedge \tau_J] \times \boldsymbol{B}_1 \right) \right] = o(t^{1/2}) \quad \text{as} \quad t \searrow 0.$$

We can now formulate the analogue of Theorem 2 in the case of semimartingales with jumps.

Theorem 11 (Central Limit Theorem with jumps)**.** *Let X satisfy Assumption 10. Then for every $f : \mathbb{R}^m \to \mathbb{R}^n$ such that there exists an open neighborhood U of x_0 with $f \in C^2(U, \mathbb{R}^n)$, we have*

$$\frac{1}{\sqrt{t}}(f(X_t) - f(x_0)) \xrightarrow{d} N_f \quad \text{as } t \searrow 0,$$

where N_f is a normal random vector with mean 0 and covariance matrix

$$V = (Df)(x_0) L (Df(x_0) L)^\top.$$

Proof. Let $r > 0$ be such that the closed ball $\overline{\boldsymbol{B}}_r(x_0)$ with radius r around x_0 is contained in U. Further, we denote by $\overline{\boldsymbol{B}}_{r/2}(x_0)$ the closed ball with radius $r/2$ around x_0 and define the hitting time $\overline{\tau} := \tau_{\overline{\boldsymbol{B}}_{r/2}(x_0)^c}$. Finally, we introduce the stopping time

(3.1) $$\tau := \overline{\tau} \wedge \tau_A \wedge \tau_M \wedge \tau_J$$

and notice that τ a.s. positive. Therefore, by the same argument as in the proof of Theorem 2, it suffices to show

$$\frac{1}{\sqrt{t}}(f(X_{t \wedge \tau}) - f(x_0)) \xrightarrow{d} N_f \quad \text{as} \quad t \searrow 0.$$

By Itô's formula in the form of Proposition 8.19 in [7], we have for all $j \in \{1, \dots, n\}$ and $t \in [0, T]$:

(3.2)

$$f_j(X_{t \wedge \tau}) - f_j(x_0) = \int_0^{t \wedge \tau} (\mathcal{L}_s f_j)(X_s) \, \mathrm{d}s + \sum_{k,l=1}^m \int_0^{t \wedge \tau} \frac{\partial f_j}{\partial x_l}(X_s) \sigma_s^{lk} \, \mathrm{d}B_s^k$$

(3.3) $$+ \int_0^{t \wedge \tau} \int_{\boldsymbol{B}_1} \left(f_j(X_{s-} + \psi(s,z)) - f_j(X_{s-}) \right) \left(\Pi(\mathrm{d}s, \mathrm{d}z) - \mu(\mathrm{d}s, \mathrm{d}z) \right)$$

(3.4) $$+ \int_0^{t \wedge \tau} \int_{\mathbb{R}^m \setminus \boldsymbol{B}_1} \left(f_j(X_{s-} + \varphi(s,z)) - f_j(X_{s-}) \right) \Pi(\mathrm{d}s, \mathrm{d}z).$$

Arguing as in the proof of Theorem 2, we see that the vector of terms on the right-hand side of (3.2), rescaled by $\frac{1}{\sqrt{t}}$, converges in distribution to N_f as $t \searrow 0$. Thus, the theorem will follow if we can show that the terms (3.3) and (3.4), rescaled by $\frac{1}{\sqrt{t}}$, converge to zero in probability as $t \searrow 0$.

The term (3.3), rescaled by $\frac{1}{\sqrt{t}}$, can be decomposed into a sum $T_t^1 + T_t^2$ of the following two terms:

$$\frac{1}{\sqrt{t}} \int_0^{t \wedge \tau} \int_{\boldsymbol{B}_1} \big(f_j(X_{s-} + \psi(s,z)) - f_j(X_{s-})\big) \mathbf{1}_{\{|\psi(s,z)| < r/2\}} \big(\Pi(\mathrm{d}s, \mathrm{d}z) - \mu(\mathrm{d}s, \mathrm{d}z)\big),$$

$$\frac{1}{\sqrt{t}} \int_0^{t \wedge \tau} \int_{\boldsymbol{B}_1} \big(f_j(X_{s-} + \psi(s,z)) - f_j(X_{s-})\big) \mathbf{1}_{\{|\psi(s,z)| \geq r/2\}} \big(\Pi(\mathrm{d}s, \mathrm{d}z) - \mu(\mathrm{d}s, \mathrm{d}z)\big).$$

Then:

$$\mathbb{E}\big[|T_t^1|\big] \leq \frac{2\|f|_{\overline{\boldsymbol{B}}_r(x_0)}\|_\infty}{\sqrt{t}} \mathbb{E}\Big[\big|\Pi - \mu\big|\big([0, t \wedge \tau_J] \times \boldsymbol{B}_1\big)\Big],$$

which converges to zero as $t \searrow 0$ by part (3) of Assumption 10. Moreover, since J a.s. has only finitely many jumps of absolute size greater than $r/2$ on every finite time interval, T_t^2 converges to 0 a.s. as $t \searrow 0$.

Lastly, the term (3.4), rescaled by $\frac{1}{\sqrt{t}}$, converges to zero a.s. as $t \searrow 0$, since J a.s. has only finitely many jumps of absolute size greater than 1 on every finite time interval. □

As in the case of continuous semimartingales, the Central Limit Theorem can be strengthened to a Functional Central Limit Theorem, which in the presence of jumps reads as follows.

Theorem 12 (Functional Central Limit Theorem with jumps). *Let X satisfy Assumption 10. Then for every $f : \mathbb{R}^m \to \mathbb{R}^n$ such that there exists an open neighborhood U of x_0 with $f \in C^2(U, \mathbb{R}^n)$, the processes*

$$Y^{f,u} := \left(\frac{f(X_{ut}) - f(x_0)}{\sqrt{u}}\right)_{t \in [0,T]}, \quad u \in (0,1),$$

converge in law to a Brownian motion with variance-covariance matrix given by

$$V = (Df)(x_0)L(Df(x_0)L)^\top$$

as $u \searrow 0$.

Proof. For each f and u as in the statement of the theorem, we write $Q^{f,u}$ for the law of the process $Y^{f,u}$ on $D([0,T], \mathbb{R}^n)$, the space of right-continuous functions on $[0,T]$ having left limits; moreover, we denote by $Q_c^{f,u}$ the law of the continuous part of $Y^{f,u}$ on $C([0,T], \mathbb{R}^n)$. We claim first that the family $(Q^{f,u})_{u \in (0,1)}$ is tight on $D([0,T], \mathbb{R}^n)$ if and only if the family $(Q_c^{f,u})_{u \in (0,1)}$ is tight on $C([0,T], \mathbb{R}^n)$ and, moreover, that the limit points of the two families are the same.

To prove the claim, it suffices to show that for every $\varepsilon > 0$ and $j \in \{1, \ldots, n\}$:

$$(3.5) \qquad \mathbb{P}\Big(\sup_{t \in [0,T]} |(J_t^{f,u})^j| > \varepsilon \Big) \to 0 \quad \text{as} \quad u \searrow 0,$$

where $J^{f,u}$ denotes the jump part of $Y^{f,u}$. Indeed, if this is the case, then every converging subsequence of $(Q^{f,u})_{u \in (0,1)}$ in $D([0,T], \mathbb{R}^n)$ corresponds to a converging subsequence of $(Q_c^{f,u})_{u \in (0,1)}$ in $C([0,T], \mathbb{R}^n)$ and the limits of the two subsequences have to coincide. Now, since the stopping time defined in (3.1) is a.s. positive, (3.5) is implied by

$$(3.6) \qquad \mathbb{P}\Big(\sup_{t \in [0,T]} |(J_t^{f,u})^j| > \varepsilon, \ \tau > uT \Big) \to 0 \quad \text{as} \quad u \searrow 0.$$

Furthermore, by Itô's formula in the form of Proposition 8.19 in [7], we have on the event $\{\tau > uT\}$:

(3.7)

$$(J_t^{f,u})^j = \frac{1}{\sqrt{u}} \int_0^{ut} \int_{\boldsymbol{B}_1} \big(f_j(X_{s-} + \psi(s,z)) - f_j(X_{s-}) \big) \big(\Pi(\mathrm{d}s, \mathrm{d}z) - \mu(\mathrm{d}s, \mathrm{d}z) \big)$$

$$(3.8) \qquad + \frac{1}{\sqrt{u}} \int_0^{ut} \int_{\mathbb{R}^m \setminus \boldsymbol{B}_1} \big(f_j(X_{s-} + \varphi(s,z)) - f_j(X_{s-}) \big) \Pi(\mathrm{d}s, \mathrm{d}z).$$

As in the proof of Theorem 11, we decompose the integral on the right-hand side of (3.7) according to whether $|\psi(s,z)| < r/2$, or $|\psi(s,z)| \geq r/2$, and call the two resulting processes $(J^{f,u,1})^j$ and $(J^{f,u,2})^j$, respectively. Since the process $(J^{f,u,1})^j$ is obtained by integrating a predictable process with respect to a compensated Poisson random measure, it is a square-integrable martingale. Thus, by Doob's maximal inequality, we have

$$\mathbb{P}\Big(\sup_{t \in [0,T]} |(J_t^{f,u,1})^j| > \varepsilon/2, \ \tau > uT \Big)$$

$$\leq \frac{2}{\varepsilon \sqrt{u}} \mathbb{E}\Big[\Big| \int_0^{(uT) \wedge \tau} \int_{\boldsymbol{B}_1} \big(f_j(X_{s-} + \psi(s,z)) - f_j(X_{s-}) \big) \mathbf{1}_{\{|\psi(s,z)| < r/2\}} \overline{\Pi}(\mathrm{d}s, \mathrm{d}z) \Big| \Big],$$

where we wrote $\overline{\Pi}$ for $\Pi - \mu$. Moreover, the same argument as in the proof of Theorem 11 shows that the latter upper bound tends to zero as $u \searrow 0$ (by virtue of part (3) of Assumption 10). Finally, since a.s. the process $J^{f,u}$ has finitely many jumps of size greater than $r/2$ on every finite time interval, the random variables $\sup_{t \in [0,T]} |(J_t^{f,u,2})^j|$ converge to zero a.s. as $u \searrow 0$. In addition, by the same reasoning, the supremum over $t \in [0,T]$ of (3.8) tends to zero a.s. as $u \searrow 0$ as well. Putting everything together, we end up with (3.6), finishing the proof of the claim.

Lastly, one can proceed as in the proof of Theorem 3 to first show the tightness of the family $(Q_c^{f,u})_{u \in (0,1)}$ on $C([0,T], \mathbb{R}^n)$ and to subsequently identify each of its limit points with the law of a Brownian motion with variance-covariance matrix V. In view of the claim above, this finishes the proof. $\qquad \square$

We conclude this section by stating and proving the analogue of Theorem 5 in the presence of jumps.

Theorem 13. *Suppose that the process X solves the stochastic differential equation*

$$(3.9) \qquad \mathrm{d}X_t = b(t, \cdot) \, \mathrm{d}t + \sigma(t) \, \mathrm{d}B_t + \int_{\mathbb{R}^m} \psi(t, y) \, \Pi(\mathrm{d}t, \mathrm{d}y),$$
$$X_0 = x_0,$$

where $b : [0, \infty) \times \Omega \to \mathbb{R}^m$ is a bounded predictable process with respect to the filtration of the standard m-dimensional Brownian motion B, $\sigma : [0, \infty) \to \mathbb{R}^{m \times m}$ is a locally square integrable function taking values in the set of invertible matrices such that the smallest eigenvalue of $\sigma(\cdot)^\top \sigma(\cdot)$ is uniformly bounded away from 0 and ψ is a predictable process with respect to the filtration of the Poisson random measure Π.

Suppose further that Π is symmetric with respect to y (so that, in particular, its compensator vanishes) and that $\psi_1(t, y) = -\psi_1(t, y)$ for all $t \geq 0$ and $y \in \mathbb{R}^m$ with probability 1. Then, the bounds

$$(3.10) \qquad e^{f_1(t)} \leq \mathbb{P}\big(X_t^1 > X_0^1\big) \leq e^{f_2(t)}, \quad t > 0$$

of Theorem 5 apply with the same functions f_1, f_2 as there.

Proof. We start by fixing a $t > 0$ and changing the underlying probability measure $\mathbb{P}$ to an equivalent probability measure $\mathbb{Q}$ according to

$$(3.11) \qquad \frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}} = \exp\Big(-\int_0^t \sigma(s)^{-1} b(s, \cdot) \, \mathrm{d}B_s - \frac{1}{2} \int_0^t |\sigma(s)^{-1} b(s, \cdot)|_2^2 \, \mathrm{d}s \Big).$$

Then, in view of the independence of the continuous and the jump parts of X under $\mathbb{P}$ and the Girsanov Theorem, the process X solves the stochastic differential equation

$$(3.12) \qquad \mathrm{d}X_s = \sigma(s) \, \mathrm{d}B_s^{\mathbb{Q}} + \int_{\mathbb{R}^m} \psi(s, y) \, \Pi(\mathrm{d}s, \mathrm{d}y), \quad s \in [0, t]$$

with a standard Brownian motion $B^{\mathbb{Q}}$ under $\mathbb{Q}$ and initial condition $X_0 = x_0$. Moreover, the random variables

$$U_t^{(1)} := \Big(\int_0^t \sigma(s) \, \mathrm{d}B_s^{\mathbb{Q}} \Big)_1 \quad \text{and} \quad U_t^{(2)} := \int_0^t \int_{\mathbb{R}^m} \psi_1(s, y) \, \Pi(\mathrm{d}s, \mathrm{d}y)$$

are independent under $\mathbb{Q}$ and their distributions $\eta^{(1)}$ and $\eta^{(2)}$ are symmetric. Hence,

$$\begin{aligned}
\mathbb{Q}(X_t^1 > X_0^1) &= \mathbb{Q}\big(U_t^{(1)} + U_t^{(2)} > 0\big) \\
&= \int_0^\infty \mathbb{Q}\big(U_t^{(1)} > -c\big)\,\eta^{(2)}(\mathrm{d}c) + \int_{-\infty}^0 \mathbb{Q}\big(U_t^{(1)} > -c\big)\,\eta^{(2)}(\mathrm{d}c) \\
&= \int_0^\infty \mathbb{Q}\big(U_t^{(1)} > -c\big)\,\eta^{(2)}(\mathrm{d}c) + \int_{-\infty}^0 1 - \mathbb{Q}\big(U_t^{(1)} > c\big)\,\eta^{(2)}(\mathrm{d}c) \\
&= \frac{1}{2}.
\end{aligned}$$

From now on, one can follow the lines of the proof of Theorem 5 to finish the proof. $\qquad\square$

4. DIGITAL OPTIONS AND THE IMPLIED VOLATILITY SLOPE

Suppose that the one-dimensional, positive process S models the price of a financial asset, and that $\mathbb{P}$ is the pricing measure. The riskless rate is $r > 0$. The holder of a digital call option with maturity T and strike K receives the payoff $\mathbf{1}_{\{S_T > K\}}$ at maturity. Digital options are peculiar in that the owner receives the full payoff as soon as they are only slightly in the money, as opposed to call options, say, which kick in gradually. By the risk-neutral pricing formula, the value of the digital call at time zero is

$$D(K,T) := e^{-rT}\,\mathbb{E}[\mathbf{1}_{\{S_T > K\}}] = e^{-rT}\,\mathbb{P}(S_T > K).$$

There is a considerable literature on short-maturity approximations for option prices. For OTM (out-of-the-money; $S_0 < K$) or ITM (in-the-money; $S_0 > K$) digitals, the first order approximation is clear: As soon as the underlying S is a.s. right-continuous at $t = 0$, the Dominated Convergence Theorem yields

$$\lim_{T \to 0} D(K,T) = \begin{cases} 0 & \text{if } S_0 < K \qquad \text{(OTM)} \\ 1 & \text{if } S_0 > K \qquad \text{(ITM)}. \end{cases}$$

Finer information on the OTM decay (which trivially also covers the ITM behavior) comes from small time large deviations principles for the underlying. E.g., see Forde and Jacquier [15] for the case of the Heston model and references about other diffusion processes. Our CLT-type results are useful in the *ATM* case $S_0 = K$. As an immediate consequence of our limit theorems, we enunciate:

Theorem 14. *If the process S satisfies the assumptions of Theorem 11 (in particular, if it satisfies those of Theorem 2 or Remark 4), and the limit law is non-degenerate, then the limiting price of an at-the-money digital call is $1/2$:*

$$(4.1) \qquad\qquad \lim_{T \to 0} D(S_0, T) = \frac{1}{2}.$$

This (intuitive) result captures virtually all diffusion-based models that have been considered (Black-Scholes, CEV, Heston, Stein-Stein, etc.). Although it seems to be new in its generality, in particular for jump processes, some special

cases can be inferred from the literature (see the comment at the end of this section).

The jump processes used in financial modelling are often Lévy processes. It is clear that a compensated compound Poisson process will yield an (unrealistic) ATM digital price limit of either zero or one (see the remark before Example 9). As for the infinite activity case, limit laws are not the appropriate way to get a result like (4.1). Doney and Maller [11] have determined all Lévy processes that admit a short-time CLT, with a criterion involving the tail of the Lévy measure. While there do exist infinity activity Lévy processes that satisfy a CLT [11, Remark 9], the Lévy processes that have been considered in mathematical finance are typically *not* of this kind. For instance, it is easy to see from the characteristic function that the variance gamma process [27] does not admit *any* non-degenerate limit law for $t \to 0$, for any normalization. We will discuss these issues further in the more application-oriented companion paper [18].

Finally, we discuss the implied volatility skew. Suppose that the underlying S generates the call price surface $C(K,T)$:

$$C(K,T) = e^{-rT}\,\mathbb{E}[(S_T - K)^+], \quad K > 0, \quad T > 0.$$

Then the implied volatility (see e.g. [25]) for strike K and maturity T is the volatility $\sigma_{\mathrm{imp}}(K,T)$ that makes the Black-Scholes call price equal to $C(K,T)$:

$$C_{\mathrm{BS}}(K, \sigma_{\mathrm{imp}}, T) = C(K,T).$$

The map $K \to \sigma_{\mathrm{imp}}(K,T)$ is called the volatility smile for maturity T. It is also called the volatility skew, because it is often monotone instead of smile-shaped, but we will reserve the term skew for the *derivative* $\partial_K \sigma_{\mathrm{imp}}(K,T)$. If $C(K,T)$ is smooth in K, it equals (we omit arguments)

$$\partial_K \sigma_{\mathrm{imp}} = -\frac{\partial_K C_{\mathrm{BS}} - \partial_K C}{\partial_\sigma C_{\mathrm{BS}}}.$$

Under mild assumptions (e.g., if the law of S_T is absolutely continuous), we have

$$(4.2) \qquad \partial_K C = -e^{-rT}\,\mathbb{P}(S_T \geq K) = -D(K,T),$$

from which we deduce the (well-known) connection between the volatility skew and the price of a digital call (see e.g. [17]):

$$\partial_K \sigma_{\mathrm{imp}} = -\frac{D(K,T) + \partial_K C_{\mathrm{BS}}}{\partial_\sigma C_{\mathrm{BS}}}.$$

Inserting the explicit Black-Scholes vega and digital price (see e.g. [31]), we obtain

$$\partial_K \sigma_{\mathrm{imp}} = \frac{-D(K,T) + \Phi(-\sigma_{\mathrm{imp}}\sqrt{T}/2)}{K\sqrt{T}\,n(\sigma_{\mathrm{imp}}\sqrt{T}/2)},$$

with Φ and n denoting the standard normal cdf and density, respectively. For $T \to 0$, we have $\sigma_{\mathrm{imp}}\sqrt{T} = o(1)$ under the following mild assumptions [34,

Proposition 4.1]:

(4.3) $\qquad (S_0 - K)^+ \leq C(K, T) \leq S_0$ (no arbitrage bounds),

(4.4) $\qquad \lim_{T \to 0} C(K, T) = (S_0 - K)^+,$

(4.5) $\qquad T \mapsto C(K, T)$ is non-decreasing.

Therefore,

$$(4.6)\quad \partial_K \sigma_{\text{imp}} \sim \frac{\sqrt{2\pi}}{K\sqrt{T}}\left(\frac{1}{2} - D(K, T) - \frac{\sigma_{\text{imp}}\sqrt{T}}{2\sqrt{2\pi}} + O((\sigma_{\text{imp}}\sqrt{T})^3)\right), \quad T \to 0.$$

We see that the small time behavior of the skew is related to that of the digital price. At the money, the latter will typically tend to $1/2$ (see Theorem 14), and so higher order estimates are needed to get the first order asymptotics of the ATM skew $\partial_K \sigma_{\text{imp}}|_{K=S_0}$. To this end, we apply our Theorem 5, and compare our findings with the standard model free slope bounds [16, page 36]

$$(4.7)\qquad -\frac{\sqrt{2\pi}}{S_0\sqrt{T}}(1 - \Phi(d_2))e^{-rT + d_1^2/2} \leq \frac{\partial\sigma_{\text{imp}}}{\partial K} \leq \frac{\sqrt{2\pi}}{S_0\sqrt{T}}\Phi(d_2)e^{-rT + d_1^2/2},$$

where

$$d_1 = \frac{\log(S_0/K) + (r + \frac{1}{2}\sigma_{\text{imp}}^2)T}{\sigma_{\text{imp}}\sqrt{T}},$$

$$d_2 = d_1 - \sigma_{\text{imp}}\sqrt{T}.$$

Such bounds can give guidance on model choice; recall that the market slope seems to grow like $T^{-1/2}$ for short maturities [1]. Note that the following result accomodates stochastic interest rates, and recall that we assume in this section that the dimension is $m = 1$. Under stochastic interest rates, the digital call price is

$$(4.8)\qquad D(K, T) = \mathbb{E}\left[e^{-\int_0^T r(s)\,\mathrm{d}s}\mathbf{1}_{\{S_T > K\}}\right].$$

To calculate the implied volatility, a *deterministic* rate r has to be chosen (e.g., by $e^{-rT} = \mathbb{E}[\exp(-\int_0^T r(s)\,\mathrm{d}s)]$). This choice is irrelevant for Theorem 15, though.

Theorem 15. *Assume that the price process satisfies the SDE*

$$\mathrm{d}S_t/S_t = r(t)\,\mathrm{d}t + \sigma(t)\,\mathrm{d}B_t$$

with the stochastic short rate process $(r(t))_{t\geq 0}$, and that the log-price $X = \log S$, whose drift is $b(t) = r(t) - \frac{1}{2}\sigma^2(t)$, satisfies the assumptions of Theorem 5. Assume further that $\partial_K C(K, T) = -D(K, T)$ holds (cf. (4.2)) and that (4.4) and (4.5) are satisfied. Then we have the ATM slope bounds

$$\partial_K \sigma_{\text{imp}}|_{K=S_0} \geq \frac{\sqrt{2\pi}}{K\sqrt{T}}\left(-C\sqrt{T} - \frac{\sigma_{\text{imp}}\sqrt{T}}{2\sqrt{2\pi}} + O(T) + O((\sigma_{\text{imp}}\sqrt{T})^3)\right),$$

$$\partial_K \sigma_{\text{imp}}|_{K=S_0} \leq \frac{\sqrt{2\pi}}{K\sqrt{T}}\left(C\sqrt{T} - \frac{\sigma_{\text{imp}}\sqrt{T}}{2\sqrt{2\pi}} + O(T) + O((\sigma_{\text{imp}}\sqrt{T})^3)\right),$$

where

$$C = \sqrt{\frac{\log 2}{2}} \|\sigma^{-1}b\|_{2,\infty}.$$

Proof. According to (4.8), the ATM digital price equals

$$D(S_0, T) = \mathbb{E}\big[e^{-\int_0^T r(s)ds} \mathbf{1}_{\{X_T > x_0\}}\big].$$

The discount factor is $1 + O(T)$, so we can apply Theorem 5 to conclude

$$\frac{1}{2} - C\sqrt{T} + O(T) \le D(S_0, T) \le \frac{1}{2} + C\sqrt{T} + O(T).$$

Now, the result follows from (4.6). Note that $\sigma_{\mathrm{imp}}\sqrt{T} = o(1)$ by [34, Proposition 4.1], since we assume (4.4) and (4.5), and (4.3) is satisfied in our setup. $\square$

The bounds in Theorem 15 are asymptotically stronger than the general estimate (4.7), which is of order $O(T^{-1/2})$, since $\sigma_{\mathrm{imp}}\sqrt{T} = o(1)$. If the Berestycki-Busca-Florent formula [5] holds, then implied volatility tends to a constant. Therefore, our bounds are *considerably* stronger than (4.7) in this case, namely of order $O(1)$. The models covered by Theorem 15 thus do *not* match the empirical slope behavior $T^{-1/2}$, similarly to stochastic volatility models [26], whose slope also behaves like $O(1)$.

To conclude our discussion of ATM digitals and the implied volatility skew, note that, for some diffusion processes, the result in Theorem 14 is implicitly in the literature. To wit, by (4.6), a non-exploding ATM slope requires a limit price of 1/2 of the digital. See Durrleman [12, page 59] for a general expression for the implied volatility slope that shows that it does not explode, e.g., in the Heston model.

REFERENCES

[1] E. ALÒS, J. A. LEÓN, M. PONTIER, AND J. VIVES, *A Hull and White formula for a general stochastic volatility jump-diffusion model with applications to the study of the short-time behavior of the implied volatility*, J. Appl. Math. Stoch. Anal., (2008), pp. Art. ID 359142, 17.

[2] E. ALÒS, J. A. LEÓN, AND J. VIVES, *On the short-time behavior of the implied volatility for jump-diffusion models with stochastic volatility*, Finance Stoch., 11 (2007), pp. 571–589.

[3] E. BENHAMOU, E. GOBET, AND M. MIRI, *Smart expansion and fast calibration for jump diffusions*, Finance Stoch., 13 (2009), pp. 563–589.

[4] A. BENTATA AND R. CONT, *Short-time asymptotics for marginal distributions of semi-martingales*, arXiv:1202.1302, (2012).

[5] H. BERESTYCKI, J. BUSCA, AND I. FLORENT, *Asymptotics and calibration of local volatility models*, Quant. Finance, 2 (2002), pp. 61–69. Special issue on volatility modelling.

[6] H. BERESTYCKI, J. BUSCA, AND I. FLORENT, *Computing the implied volatility in stochastic volatility models*, Comm. Pure Appl. Math., 57 (2004), pp. 1352–1373.

[7] R. CONT AND P. TANKOV, *Financial modelling with jump processes*, Chapman & Hall/CRC Financial Mathematics Series, Chapman & Hall/CRC, Boca Raton, FL, 2004.

[8] F. DELBAEN AND W. SCHACHERMAYER, *The existence of absolutely continuous local martingale measures*, Ann. Appl. Probab., 5 (1995), pp. 926–945.

[9] A. DEMBO AND O. ZEITOUNI, *Large deviations techniques and applications*, vol. 38 of Stochastic Modelling and Applied Probability, Springer-Verlag, Berlin, 2010. Corrected reprint of the second (1998) edition.

[10] F. DEN HOLLANDER, *Large deviations*, vol. 14 of Fields Institute Monographs, American Mathematical Society, Providence, RI, 2000.

[11] R. A. DONEY AND R. A. MALLER, *Stability and attraction to normality for Lévy processes at zero and at infinity*, J. Theoret. Probab., 15 (2002), pp. 751–792.

[12] V. DURRLEMAN, *From implied to spot volatilities*, PhD thesis, Princeton University, 2004.

[13] J. FENG, M. FORDE, AND J.-P. FOUQUE, *Short-maturity asymptotics for a fast mean-reverting Heston stochastic volatility model*, SIAM J. Financial Math., 1 (2010), pp. 126–141.

[14] J. E. FIGUEROA-LÓPEZ AND C. HOUDRÉ, *Small-time expansions for the transition distributions of Lévy processes*, Stochastic Process. Appl., 119 (2009), pp. 3862–3889.

[15] M. FORDE AND A. JACQUIER, *Small-time asymptotics for implied volatility under the Heston model*, Int. J. Theor. Appl. Finance, 12 (2009), pp. 861–876.

[16] J.-P. FOUQUE, G. PAPANICOLAOU, AND K. R. SIRCAR, *Derivatives in financial markets with stochastic volatility*, Cambridge University Press, Cambridge, 2000.

[17] J. GATHERAL, *The Volatility Surface, A Practitioner's Guide*, Wiley, 2006.

[18] S. GERHOLD AND M. SHKOLNIKOV, *Short maturity implied volatility skew and digital option prices.* Working paper, 2012.

[19] P. GUASONI, *No arbitrage under transaction costs, with fractional Brownian motion and beyond*, Math. Finance, 16 (2006), pp. 569–582.

[20] R. Z. HASMINSKII, *Stochastic stability of differential equations*, vol. 7 of Monographs and Textbooks on Mechanics of Solids and Fluids: Mechanics and Analysis, Sijthoff & Noordhoff, Alphen aan den Rijn, 1980. Translated from the Russian by D. Louvish.

[21] J. JACOD, *Asymptotic properties of power variations of Lévy processes*, ESAIM Probab. Stat., 11 (2007), pp. 173–196.

[22] J. JACOD AND A. N. SHIRYAEV, *Limit theorems for stochastic processes*, vol. 288 of Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], Springer-Verlag, Berlin, second ed., 2003.

[23] N. L. JOHNSON, S. KOTZ, AND N. BALAKRISHNAN, *Continuous univariate distributions. Vol. 1*, Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics, John Wiley & Sons Inc., New York, second ed., 1994. A Wiley-Interscience Publication.

[24] O. KALLENBERG, *Foundations of modern probability*, Probability and its Applications (New York), Springer-Verlag, New York, second ed., 2002.

[25] R. W. LEE, *Implied volatility: statics, dynamics, and probabilistic interpretation*, in Recent advances in applied probability, Springer, New York, 2005, pp. 241–268.

[26] A. L. LEWIS, *Option valuation under stochastic volatility*, Finance Press, Newport Beach, CA, 2000. With Mathematica code.

[27] D. MADAN, P. CARR, AND E. CHANG, *The variance gamma process and option pricing*, European Finance Review, 2 (1998), pp. 79–105.

[28] S. P. MEYN AND R. L. TWEEDIE, *Stability of Markovian processes. I. Criteria for discrete-time chains*, Adv. in Appl. Probab., 24 (1992), pp. 542–574.

[29] ——, *Stability of Markovian processes. II. Continuous-time processes and sampled chains*, Adv. in Appl. Probab., 25 (1993), pp. 487–517.

[30] ——, *Stability of Markovian processes. III. Foster-Lyapunov criteria for continuous-time processes*, Adv. in Appl. Probab., 25 (1993), pp. 518–548.

[31] M. MUSIELA AND M. RUTKOWSKI, *Martingale methods in financial modelling*, vol. 36 of Stochastic Modelling and Applied Probability, Springer-Verlag, Berlin, second ed., 2005.

[32] B. OKSENDAL, *Stochastic differential equations: An introduction with applications*, Springer-Verlag, Berlin, sixth ed., 2010.

[33] D. Revuz and M. Yor, *Continuous martingales and Brownian motion*, vol. 293 of Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], Springer-Verlag, Berlin, third ed., 1999.

[34] M. Roper and M. Rutkowski, *On the relationship between the call price surface and the implied volatility surface close to expiry*, Int. J. Theor. Appl. Finance, 12 (2009), pp. 427–441.

[35] D. W. Stroock and S. R. S. Varadhan, *Multidimensional diffusion processes*, Classics in Mathematics, Springer-Verlag, Berlin, 2006. Reprint of the 1997 edition.

Vienna University of Technology, Wiedner Hauptstr. 8/105-1, A-1040 Austria
E-mail address: sgerhold@fam.tuwien.ac.at

Vienna University of Technology and arithmetica versicherungs- und finanz-mathematische Beratungs-GmbH
E-mail address: kleinert.max@gmail.com

Christian Doppler Laboratory for Portfolio Risk Management, Vienna University of Technology, Wiedner Hauptstr. 8/105-1, A-1040 Austria
E-mail address: piet.porkert@fam.tuwien.ac.at

Department of Statistics, University of California, Berkeley, CA 94720-3860
E-mail address: mshkolni@gmail.com