On the Markovian projection in the Brunick-Shreve mimicking result

Martin Forde*

Dept. Mathematics
King's College London,
London WC2R 2LS

9th May 2019

Abstract: For a one-dimensional Itô process $X_t = \int_0^t \sigma_s dW_s$ and a general $\mathcal{F}_t^X$-adapted non-decreasing path-dependent functional Y_t, we derive a number of forward equations for the characteristic function of (X_t, Y_t) for absolutely and non absolutely continuous functionals Y_t. The functional Y_t can be the maximum, the minimum, the local time, the quadratic variation, the occupation time or a general additive functional of X. Inverting the forward equation, we obtain a new Fourier-based method for computing the Markovian projection $\mathbb{E}(\sigma_t^2|X_t, Y_t)$ explicitly from the marginals of (X_t, Y_t), which can be viewed as a natural extension of the Dupire formula for local volatility models; $\mathbb{E}(\sigma_t^2|X_t, Y_t)$ is a fundamental quantity in the important mimicking theorems in Brunick&Shreve[BS12]. We also establish mimicking theorems for the case when Y is the local time or the quadratic variation of X (which is not covered by [BS12]), and we derive similar results for trivariate Markovian projections. †

1. Introduction

There has been a growing literature on the problem of constructing a process that mimics certain properties of a given Itô process, but is simpler in the sense that the mimicking process solves a stochastic differential equation, or more generally a stochastic functional differential equation, while the original Itô process may have drift and diffusion terms that are themselves adapted stochastic processes. The classical paper of Gyöngy[Gyö86] considers a multi-dimensional Itô process, and constructs a weak solution to a stochastic differential equation which mimics the marginals of the original Itô process at each fixed time. The drift and covariance coefficient for the mimicking process can be interpreted as the expected value of the instantaneous drift and covariance of original Itô process, conditioned on its terminal level.

Brunick&Shreve[BS12] relax the conditions of non-degeneracy and boundedness on the covariance of the Itô process imposed in [Gyö86], and they also significantly extend the Gyöngy result. More specifically, the main result Theorem 3.5 in [BS12] proves that we can match the *joint* distribution at each fixed time of various functionals of the Itō process, including the maximum-to-date or the running average of one component of the Itô process. The mimicking process now takes the form of a stochastic functional differential equation (SFDE) and the diffusion coefficient for the SFDE is given by the so-called *Markovian projection*; in the case when we are mimicking the law of the terminal value of the process X_t and another path-dependent functional Y_t, the Markovian projection is given by the conditional expectation $\hat{\sigma}(x, y, t)^2 = \mathbb{E}(\sigma_t^2 \,|\, X_t = x, Y_t = y)$.

[BS12] do not provide a constructive method for computing $\hat{\sigma}(x, y, t)^2$; however, for the standard problem of just mimicking the law of the terminal value of the process, this can be computed from the well known Dupire forward equation for continuous semimartingales, in terms of infinitesimal calendar and butterfly spreads of put or call options. This equation was derived heuristically in [Dup96] and can be proved rigorously using the Tanaka-Meyer formula for continuous semimartingales, see Klebaner[Kle02]. Bentata&Cont[BC09II] extend this analysis to derive a forward partial integro-differential equation for the call option price (in the sense of distributions) when the underlying asset follows a (possibly) discontinuous semimartingale. In another article, [BC09] have also extended the Gyöngy mimicking result to jump diffusion processes, but they assume a priori that the Markovian projection is continuous; it is not clear if/when this holds if we do not also assume a priori that the original Itô process admits a positive density at the point (x, y) of interest.

The other main technical obstacle in establishing fitting and mimicking results of this nature is establishing *uniqueness* for the associated forward Kolmogorov equation (or associated partial integro-differential equation when there is a jump component), in the sense of distributions. This can be done when Y_t is an a.s. absolutely continuous functional using standard existence and uniqueness theorems for the forward Kolmogorov equation associated with the mimicking

*martin.forde@kcl.ac.uk

†The authors would like to thank Gerard Brunick, Jin Feng, Chris Rogers and Hongzhong Zhang for useful discussions.

diffusion process, which is degenerate when we are just mimicking the marginals of the two quantities (X_t, Y_t), because there is only one driving Brownian motion. It is less clear how to proceed for a.s. non-absolutely continuous functionals likes the running maximum or local time, because the mimicking process now takes the form of a non-standard stochastic functional differential equation for which the theory is less developed.

In this article, we consider an $\mathbb{R}$-valued square integrable Itô semimartingale of the form $dX_t = \sigma_t dW_t$, and a general $\mathcal{F}_t^X$-adapted non-decreasing process Y_t (this is our path-dependent functional of interest). We first consider the case when Y_t is a.s. non absolutely continuous and $X_t = g(Y_t)$ for some continuous function $g(.)$, on the *growth set* of Y_t; this condition is satisfied when for example when Y is the running maximum of X with $g(y) = y$, or if $Y_t = L_t^a$ the local time of X at a with $g(y) = a$. In this setup, we derive a general forward equation for the Fourier-Laplace transform of the law of (X_t, Y_t), and the forward equation can be inverted to compute the Markovian projection $\hat{\sigma}(x, y, t)^2$ explicitly via a Fourier-Laplace inversion, without the a priori assumption that (X_t, Y_t) has a density at (x, y) or that $\hat{\sigma}(t, ., .)$ is continuous at (x, y). In section 5, we consider the case when Y_t is an a.s. absolutely continuous functional; we first derive a mimicking result for the case when $Y_t = \langle X \rangle_t$ which is not considered in [BS12], and we then derive a forward equation for the Fourier-Laplace transform of the law of (X_t, Y_t) and a similar equation when Y is an additive functional of X. In both cases, we can use the forward equation to compute the appropriate Markovian projection, and we conclude the article with a similar forward equation for a trivariate Markovian projection.

2. The modelling set up

We let $X : [0, T] \times \Omega \mapsto \mathbb{R}$ denote an Itô process, i.e. a continuous martingale defined on a filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ satisfying the usual conditions, with stochastic integral representation

$$X_t \quad = \quad \int_0^t \sigma_s dW_s \,, \tag{2.1}$$

where W is a standard one-dimensional Brownian motion adapted to $\mathcal{F}_t$, and σ_t is a adapted process with $\mathbb{E}(\int_0^t \sigma_s^2 ds) < \infty$ for all $t \le T$. Let $\mathcal{F}_t^X$ denote the natural filtration of X. Throughout, we let $(Y_t)_{t \ge 0}$ denote an a.s. continuous non-decreasing $\mathcal{F}_t^X$-adapted process with $Y_0 = 0$ - this is our path-dependent functional of interest. We assume that Y_t has full support on $\mathbb{R}^+$ and that $\mathbb{E}(Y_t) < \infty$ for all $t \le T$.

We begin with a short technical lemma.

Lemma 2.1. *There exists a function $\hat{\sigma}^2 : \mathbb{R} \times \mathbb{R}^+ \times (0, T] \mapsto \mathbb{R}^+$ such that $\hat{\sigma}^2(., ., t)$ is (Borel) measurable and*

$$\mathbb{E}(\sigma_t^2 \,|\, X_t, Y_t) \quad = \quad \hat{\sigma}^2(X_t, Y_t, t) \qquad\qquad a.s.$$

Proof. See Appendix C. □

Remark 2.2. *We refer to $\hat{\sigma}^2(x, y, t)$ as the Markovian projection of σ_t^2 on (X_t, Y_t).*

2.1. The Brunick-Shreve mimicking result

We now briefly summarize the main result in Brunick&Shreve[BS12] for the special case when the dimension $n = 1$ and the process under consideration is driftless.

For an Itô process of the form in (2.1), [BS12] consider a certain class of path-dependent functionals Y of X (which they refer to as *updating functions*), which can include $Y_t = \bar{X}_t$ (the running maximum of X), $Y_t = \underline{X}_t$ (the running minimum of X) or an additive functional of the form $Y_t = \int_0^t g(X_s) ds$, but cannot include $\langle X \rangle_t$ the quadratic variation of X or L_t^a the local time of X at $x = a$ because these functionals are not continuous in the sup norm topology. For continuous functionals Y in the class of updating functions, the main Theorem 3.6 in [BS12] proves that there exists a filtered probability space $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathcal{F}}_t, \hat{\mathbb{P}})$ that supports a continuous adapted process $\hat{X}$ on $\mathbb{R}$ and a one-dimensional Brownian motion $\hat{W}$ satisfying

$$\hat{X}_t \quad = \quad \int_0^t \hat{\sigma}(\hat{X}_s, \hat{Y}_s, s) d\hat{W}_s \,,$$

where

$$\hat{\sigma}(\hat{X}_t, \hat{Y}_t, t)^2 \quad = \quad \mathbb{E}(\sigma_t^2 | \, X_t, Y_t)$$

Lebesgue a.e. on $[0,T]$, such that the distribution of Y_t under $\mathbb{P}$ agrees with the distribution of $\hat{Y}_t$ under $\hat{\mathbb{P}}$ for all $t \in [0,T]$.

From here on, we make the following assumption throughout:

Assumption 2.3. *We assume that $\mathbb{E}(\sigma_t^2) < K < \infty$ for all $t \in [0,T]$ and that σ_t is positive and stochastically continuous.*

2.2. Mimicking the joint marginals of the terminal level and the local time

As mentioned above, Theorem 3.6 in [BS12] does not cover the case when the path-dependent functional is the local time of X at zero. However, by adapting their argument, we will now prove a similar mimicking result for local time[1]. Let L_t denote the local time of X at zero and let $Y_t = |X_t| - L_t = \int_0^t \operatorname{sgn}(X_s)dX_s$ and $Z_t = (X_t, Y_t)$. Then Z_t is an Itô process, and we have

$$\mathbb{E}\big(\sigma_t^2 \begin{bmatrix} 1 & \operatorname{sgn}(X_t) \\ \operatorname{sgn}(X_t) & 1 \end{bmatrix} \mid X_t, Y_t\big) \quad = \quad \begin{bmatrix} 1 & \operatorname{sgn}(X_t) \\ \operatorname{sgn}(X_t) & 1 \end{bmatrix} \hat{\sigma}^2(X_t, Y_t)\,,$$

t-a.e., where $\hat{\sigma}^2(X_t, Y_t) = \mathbb{E}(\sigma_t^2 | X_t, Y_t)$. Applying Theorem 3.6 in [BS12] to Z_t, we can then mimic the marginals of Z_t for each $t > 0$ with a diffusion-type process of the form

$$\begin{aligned} d\hat{X}_t &= \hat{\sigma}(\hat{X}_t, \hat{Y}_t, t) d\hat{W}_t\,, \\ d\hat{Y}_t &= \hat{\sigma}(\hat{X}_t, \hat{Y}_t, t) \operatorname{sgn}(\hat{X}_t)\, d\hat{W}_t \end{aligned}$$

on some $(\hat{\Omega}, \hat{\mathcal{F}}, \{\hat{\mathcal{F}}_t\}, \hat{\mathbb{P}})$. Then $(\hat{X}_t, |\hat{X}|_t - \hat{Y}_t)$ has the same distribution under $\hat{\mathbb{P}}$ as $(X_t, |X_t| - Y_t) = (X_t, L_t)$ under $\mathbb{P}$, and $|\hat{X}|_t - \hat{Y}_t$ is a version of the local time process $\hat{L}_t$ of $\hat{X}_t$.

3. Computing $\mathbb{E}(\sigma_t^2 \mid X_t, Y_t)$ explicitly - the non-absolutely continuous case

In the following subsections, we show how to compute the Markovian projection $\hat{\sigma}^2(.,.,t)$ explicitly for non absolutely continuous functionals, by deriving a forward equation for the Fourier-Laplace transform of the law of (X_t, Y_t).

Theorem 3.1. *Assume the process $(Y_t)_{t\geq 0}$ is not absolutely continuous a.s. and there exists a continuous function $g : \mathbb{R}^+ \mapsto \mathbb{R}$ such that*

$$X_t \quad = \quad g(Y_t) \tag{3.1}$$

a.s. on the growth set of Y_t [2]. Set $\phi(k,\lambda,t) = \mathbb{E}(e^{ikX_t - \lambda Y_t})$ for $k \in \mathbb{R}, \lambda \in \mathbb{C}$, $\mathrm{Re}(\lambda) \geq 0$ and assume that at least one of k or λ is not zero. Then we have the following forward equation for ϕ:

$$\partial_t[\phi + \lambda\, \mathbb{E}(G(k,\lambda,Y_t))] \quad = \quad -\frac{1}{2}k^2 U\,, \tag{3.2}$$

where $G(k,\lambda,y) = \int_0^y e^{ikg(u) - \lambda u} du$ and

$$U(k,\lambda,t) \quad = \quad \mathbb{E}(\sigma_t^2\, e^{ikX_t - \lambda Y_t}) \quad = \quad \mathbb{E}(\hat{\sigma}(X_t, Y_t, t)^2\, e^{ikX_t - \lambda Y_t})$$

is the Fourier-Laplace transform of the bounded measure $q(dx,dy,t) = \hat{\sigma}(x,y,t)^2 p(dx,dy,t)$, where $p(dx,dy,t) = \mathbb{P}(X_t \in dx, Y_t \in dy)$. [3].

Proof. Applying the Itô formula to $Z_t = e^{ikX_t - \lambda Y_t}$ we obtain

$$dZ_t \quad = \quad ikZ_t dX_t \,-\, \frac{1}{2}k^2 Z_t \sigma_t^2 \,-\, \lambda\, e^{ikg(Y_t) - \lambda Y_t} dY_t\,,$$

[1]We are grateful to Gerard Brunick for pointing this out.

[2]By growth set, we mean the support of the random measure induced by the process Y on $[0,T]$, i.e. the complement of the largest open set of zero measure.

[3]We know q_t is bounded because we imposed that $\mathbb{E}(\sigma_t^2) = \mathbb{E}(\hat{\sigma}^2(X_t, Y_t, t)) < \infty$.

where we have used that $X_t = g(Y_t)$ on the growth set of Y_t in the final term. Integrating, we obtain

$$Z_t - 1 \;=\; ik\int_0^t Z_s dX_s \,-\, \frac{1}{2}k^2\int_0^t Z_s\sigma_s^2 ds \,-\, \lambda G(k,\lambda,Y_t)\,.$$

Taking expectations, and using that $|Z_t| \leq 1$ and $\mathbb{E}(\int_0^t \sigma_s^2 ds) < \infty$, we can apply Fubini's theorem to obtain

$$\mathbb{E}[Z_t \,+\, \lambda G(k,\lambda,Y_t)] \,-\, 1 \;=\; -\frac{1}{2}k^2\int_0^t \mathbb{E}(Z_s\sigma_s^2)ds\,. \tag{3.3}$$

But σ_t is stochastically continuous and Z_t is continuous a.s., so $Z_u\sigma_u^2 \to Z_s\sigma_s^2$ a.s. as $u \to s$. Moreover, $|Z_t\sigma_t^2| \leq \sigma_t^2 \leq K$ for all $t \in [0,T]$. Thus by the dominated convergence theorem $\lim_{u\to s}\mathbb{E}(Z_u\sigma_u^2) = \mathbb{E}(Z_s\sigma_s^2)$ so the integrand $\mathbb{E}(Z_s\sigma_s^2)$ in (3.3) is continuous in s. Thus we can differentiate (3.3) everywhere with respect to t to obtain (3.2). Finally, using iterated expectations, we have that

$$\mathbb{E}(Z_t\sigma_t^2) \;=\; \mathbb{E}(Z_t\hat{\sigma}(X_t,Y_t)^2).$$

□

Remark 3.2. *Clearly $\phi_t(k,\lambda)$ and $\mathbb{E}(G(k,\lambda,Y_t))$ can both be computed just from the marginals of (X_t,Y_t). Thus, using (3.2) we can back out the unknown $U(k,\lambda,t)$ for $k \neq 0$ [4], from $\phi(k,\lambda,t)$ for all $t \in (0,T]$, and thus $q(dx,dy,t)$ (via a Laplace-Fourier inversion). $q(dx,dy,t)$ is absolutely continuous with respect to $p(dx,dy,t) = \mathbb{P}(X_t \in dx, Y_t \in dy)$ because $\mathbb{E}(\sigma_t^2) < \infty$. If $p(dx,dy,t) = \rho_t(x,y)dxdy + \alpha_t(dx,dy)$ where α_t is singular with respect to Lebesgue measure, then $q(dx,dy,t) = \hat{\sigma}(x,y,t)^2(\rho_t(x,y)dxdy + \alpha_t(dx,dy))$ and we can then compute $\hat{\sigma}(x,y,t)^2$ from $q(dx,dy,t)$ and $p(dx,dy,t)$ at atoms and non atomic points of $p(dx,dy,t)$.*

3.1. The local volatility case and the Dupire formula

In this subsection, we derive a simpler forward equation for the characteristic function of X alone, which yields an alternative methodology to the Dupire formula for computing the usual one-dimensional Markovian projection $\mathbb{E}(\sigma_t^2|X_t = x)$.

Proposition 3.3. *Let $\phi(k,t) = \mathbb{E}(e^{ikX_t})$ for $k \in \mathbb{R}$. Then we have the following forward equation for ϕ:*

$$\partial_t\phi \;=\; -\frac{1}{2}k^2 U \tag{3.4}$$

where

$$U(k,t) \;=\; \mathbb{E}(\sigma_t^2\, e^{ikX_t}) \;=\; \mathbb{E}(\hat{\sigma}(X_t,t)^2\, e^{ikX_t})$$

is the Fourier transform of the bounded measure $q(dx,t) = \hat{\sigma}(x,t)^2 p(dx,t)$, where $p(dx,t) = \mathbb{P}(X_t \in dx)$ and $\hat{\sigma}(x,t)^2 = \mathbb{E}(\sigma_t^2|X_t = x)$.

Remark 3.4. *Using (3.4) and performing a Fourier inversion, we can back out $q(dx,t)$ (and thus $\hat{\sigma}(x,t)^2$) from $\phi(k,t)$ for all $t \in (0,T]$. If $p(dx,t)$ has a $C^{2,1}$ density $p(x,t)$, then (3.4) is just the Fourier transform of the forward Kolmogorov equation*

$$\partial_t p \;=\; \partial_{xx}^2(\frac{1}{2}\hat{\sigma}(x,t)^2 p)\,,$$

and integrating this twice in x we obtain the celebrated Dupire formula

$$\partial_T C \;=\; \frac{1}{2}\hat{\sigma}(K,T)^2\partial_{KK}^2 C$$

for the price of a call option $C(K,T) = \mathbb{E}(X_T - K)^+$. Proposition 3.3 provides a way of computing $\hat{\sigma}(x,t)$ without having to calculate (or estimate) $\partial_T C$ or $\partial_{KK}^2 C$, which is notoriously difficult in practice with noisy and incomplete option price data.

[4] For $k = 0$ we just use the continuity of ϕ.

4. Examples

4.1. The running maximum

If $Y_t = \max_{0\le s\le t} X_s$ the running maximum of X_t, then from Doob's maximal inequality (see page 14 in [KS91]) we have that $\mathbb{E}(Y_t^2) \le 4\mathbb{E}(X_t^2) < \infty$ because X_t is a square integrable martingale. In this case, $g(y) = y$ because Y can only increase if $X_t = Y_t$ and (3.2) becomes

$$\partial_t[\phi + \frac{\lambda}{ik-\lambda}\,\mathbb{E}(e^{(ik-\lambda)Y_t})] \;=\; -\frac{1}{2}k^2 U\,.$$

4.2. The running minimum

If $Y_t = -\underline{X}_t = -\min_{0\le s\le t} X_s$, i.e. minus the running minimum of X_t, then $g(y) = -y$ and (3.2) becomes

$$\partial_t[\phi - \frac{\lambda}{ik+\lambda}\,\mathbb{E}(e^{-(ik+\lambda)Y_t})] \;=\; -\frac{1}{2}k^2 U\,.$$

Remark 4.1. *If we have an Itô process of the form in* (2.1) *and observed barrier option prices given by* $\mathbb{E}((X_t - K)^+ 1_{\underline{X}_t > b})$ *for all* K, b, t *with* $b \le K$, $t \in [0,T]$, *then formally differentiating twice with respect to* K *and once in* b *we can recover the marginals of* $(X_t, \underline{X}_t)$ *for all* $t \in [0,T]$. *Using the Brunick-Shreve result, we can then mimic these down-and-out call option prices with a diffusion-type process of the form* $dX_t = \sigma(X_t, \underline{X}_t)dW_t$.

4.3. Local time

If $Y_t = L_t^a$ is the local time for X at $x = a$, then $g(y) = a$ and (3.2) becomes

$$\partial_t[\phi - e^{ika}\,\mathbb{E}(e^{-\lambda L_t^a})] \;=\; -\frac{1}{2}k^2 U\,.$$

5. Absolutely continuous functionals

5.1. The forward Kolmogorov equation for absolutely continuous functionals

In this section, we assume $(Y_t)_{t\ge 0}$ is absolutely continuous a.s. with $Y_t = \int_0^t b_s ds$ for some adapted process b_t such that $\int_0^t |b_s| ds < \infty$ a.s. This includes the case when Y is an additive functional: $Y_t = \int_0^t b(X_s, s)ds$, or when $Y_t = \langle X\rangle_t$ the quadratic variation of X or more generally a weighted variance swap-type functional $Y_t = \int_0^t b(X_s)\sigma_s^2 ds$ (note that these are all $\mathcal{F}_t^X$-adapted processes).

Lemma 5.1. *There exists a measurable function* $\hat{b} : \mathbb{R}\times\mathbb{R}^+\times(0,T] \mapsto \mathbb{R}$ *such that* $\hat{b}(.,.,t)$ *is measurable and*

$$\mathbb{E}(b_t \,|\, X_t, Y_t) \;=\; \hat{b}(X_t, Y_t, t) \qquad a.s.$$

Proof. Follows by a similar argument to Lemma 2.1. □

Proposition 5.2. $p(dx,dy,t) = \mathbb{P}(X_t \in dx, Y_t \in dy)$ *satisfies a degenerate forward Kolmogorov equation in the following weak sense:*

$$\forall t\in[0,T], \qquad \frac{d}{dt}\int \psi\, p(dx,dy,t) \;=\; \int (\psi_t + \hat{b}(x,y,t)\psi_y + \frac{1}{2}\hat{\sigma}(x,y,t)^2\psi_{xx})\, p(dx,dy,t) \tag{5.1}$$

[5] *for all test functions* $\psi \in C_b^{2,1,1}(\mathbb{R}\times\mathbb{R}^+\times[0,T])$. *This implies that* $p(dx,dy,t) = \mathbb{P}(X_t\in dx, Y_t\in dy)$ *is a solution to the degenerate forward Kolmogorov equation*

$$\partial_t p \;=\; -\partial_y(\hat{b}(x,y,t)p) + \partial_{xx}^2(\frac{1}{2}\hat{\sigma}(x,y,t)^2 p) \tag{5.2}$$

in the sense of distributions.

Proof. See Appendix A. □

[5] where integration is over $\mathbb{R}\times\mathbb{R}^+$.

5.2. A mimicking result for quadratic variation

Now let $\Omega = C_{x_0,y_0}([0,t], \mathbb{R}\times\mathbb{R}^+)$, $(X_t, Y_t) = (\omega_1(t), \omega_2(t))$ the canonical process on Ω, and $\mathcal{B}_t$ its natural filtration. Recall that a probability measure $\mathbb{Q}$ on $(\Omega, \mathcal{B}_T)$ is a solution to the *martingale problem* associated with a second order differential operator $\mathcal{A}$ (in the Stroock-Varadhan sense) if the process

$$f(X_t, Y_t) - f(x_0, y_0) - \int_0^t \mathcal{A}f(X_s, Y_s)ds \tag{5.3}$$

is a $(\mathbb{Q}, \mathcal{B}_t)$-martingale for all $f \in C_c^\infty(\mathbb{R}\times\mathbb{R}^+)$. The martingale problem is said to be well posed if there exists a unique $\mathbb{Q}$ to the martingale problem.

Proposition 5.3. *For the case when $(Y_t)_{t\geq 0}$ is absolutely continuous a.s., if $\hat{b}, \hat{\sigma}^2$ are bounded and continuous and have two bounded continuous spatial derivatives, then there exists a weak solution $(\hat{X}, \hat{Y}, \hat{W}), (\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathcal{F}}_t, \hat{\mathbb{P}})$ to*

$$\begin{cases} d\hat{X}_t &= \hat{\sigma}(\hat{X}_t, \hat{Y}_t, t)d\hat{W}_t\,, \\ d\hat{Y}_t &= \hat{\sigma}(\hat{X}_t, \hat{Y}_t, t)^2 dt \end{cases} \tag{5.4}$$

with $X_0 = Y_0 = 0$ which is unique in law, and if $\hat{Y}$ has full support under $\hat{\mathbb{P}}$, then $\hat{\mathbb{P}}(\hat{X}_t \in dx, \hat{Y}_t \in dy) = \mathbb{P}(X_t \in dx, Y_t \in dy)$ for all $t \in [0,T]$, i.e. $(\hat{X}_t, \hat{Y}_t)$ mimics the marginals of (X_t, Y_t) for all $t \in [0,T]$.

Proof. From Proposition 4.1 in Figalli[Fig08] (see also Theorems 3.2.6 and Corollary 6.3.3 in Stroock&Varadhan[SV79], and Proposition 5.4.11 in [KS91]), from the conditions on $\hat{b}, \hat{\sigma}^2$, then there exists a unique solution $\mathbb{Q}$ to the martingale problem for

$$\mathcal{A} \;=\; \frac{1}{2}\hat{\sigma}(x,y,t)^2\partial^2_{xx} + \hat{b}(x,y,t)\partial_y$$

and thus there exists a weak solution $(\hat{X}, \hat{Y}, \hat{W}), (\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathcal{F}}_t, \hat{\mathbb{P}})$ to (5.4) which is unique in law. Moreover, from Proposition 5.2 we know that $p(dx, dy, t)$ and $\hat{p}(dx, dy, t) = \hat{\mathbb{P}}(X_t \in dx, Y_t \in dy)$ satisfy the same weak forward Kolmogorov equation (5.1). But from the second part of the proof of the same Proposition in [Fig08], we also know that the solution to (5.1) is unique. $\square$

Remark 5.4. *Proposition 5.3 is not covered by the main result in [Gyö86] because Gyöngy assumes that the mimicking process is non-degenerate, which is not the case here because the Y process has no diffusion coefficient. Proposition 5.3 is not covered in entirety by Theorem 3.6 in [BS12] because the latter requires that the functional Y be continuous in the sup norm topology, which is not the case when e.g. $Y_t = \langle X\rangle_t$.*

Remark 5.5. *Using Hörmander's theorem for a diffusion with time-dependent coefficients (see Eq 1.5 in [CM02]), we can find conditions in terms of Lie brackets under which $p(dx, dy, t)$ admits a smooth density $p(x, y, t)$, and is thus a classical solution to the forward Kolmogorov equation.*

The mimicking result in Proposition 5.4 can be applied to the following absolutely continuous functionals:

- An additive functional: $Y_t = \int_0^t b(X_s, s)ds$ - in this case $\hat{b}(x,y,t) = \mathbb{E}(b(X_t,t)|X_t = x, Y_t = y) = b(x,y,t)$ and Proposition 5.4 holds if $\hat{b}, \hat{\sigma}^2$ satisfy the regularity conditions in the proposition.
- Quadratic variation: $Y_t = \langle X\rangle_t = \int_0^t \sigma_s^2 ds$ - in this case $\hat{b}(x,y,t) = \mathbb{E}(b(X_t, Y_t, t)|X_t = x, Y_t = y) = \hat{\sigma}(x,y,t)^2$, and Proposition 5.4 holds if $\hat{\sigma}^2$ satisfies the regularity conditions in the proposition.

6. Computing $\mathbb{E}(\sigma_t^2 \,|\, X_t, Y_t)$ explicitly - absolutely continuous cases

6.1. Quadratic variation

Proposition 6.1. *Set $\phi(k,\lambda,t) = \mathbb{E}(e^{ikX_t - \lambda\langle X\rangle_t})$ for $k \in \mathbb{R}, \lambda \in \mathbb{C}$, $\mathrm{Re}(\lambda) \geq 0$. Then we have the following forward equation for ϕ*

$$\partial_t\phi \;=\; -(\frac{1}{2}k^2 + \lambda)U\,, \tag{6.1}$$

where

$$U(k,\lambda,t) \;=\; \mathbb{E}(\sigma_t^2\, e^{ikX_t - \lambda\langle X\rangle_t}) \;=\; \mathbb{E}(\hat{\sigma}(X_t, \langle X\rangle_t, t)^2\, e^{ikX_t - \lambda\langle X\rangle_t})$$

is the Fourier-Laplace transform of the bounded measure $u(dx, dy, t) = \hat{\sigma}^2(x,y,t)\,\mathbb{P}(X_t \in dx, \langle X\rangle_t \in dy)$, where $\hat{\sigma}^2(x,y,t) = \mathbb{E}(\sigma_t^2 \,|\, X_t = x, \langle X\rangle_t = y)$.

Proof. Applying the Itô formula to $Z_t = e^{ikX_t - \lambda\langle X\rangle_t}$ we obtain

$$Z_t \;=\; 1 + \int_0^t ikZ_s\,dX_s \;-\; (\frac{1}{2}k^2 + \lambda)\int_0^t \sigma_s^2 ds\,.$$

Taking expectations, applying Fubini and using the stochastic continuity of σ_t to differentiate wrt t as before and then using iterated expectations, we obtain (6.1). □

6.2. Additive functionals and the occupation time

Proposition 6.2. *Let $Y_t = \int_0^t g(X_s, s)ds$, where g is bounded and Borel measurable and set $\phi(k,\lambda,t) = \mathbb{E}(e^{ikX_t - \lambda Y_t})$ for $k \in \mathbb{R}, \lambda \in \mathbb{C}$, $\mathrm{Re}(\lambda) \geq 0$. Then we have the following forward equation for ϕ*

$$\partial_t\phi \;\;=\;\; -\frac{1}{2}k^2 U \;-\; \lambda\,\mathbb{E}(g(X_t,t))\,, \tag{6.2}$$

where

$$U(k,\lambda,t) \;\;=\;\; \mathbb{E}(\sigma_t^2 e^{ikX_t-\lambda Y_t}) \;\;=\;\; \mathbb{E}(\hat{\sigma}(X_t,Y_t,t)^2\, e^{ikX_t-\lambda Y_t})$$

is the Fourier-Laplace transform of the bounded measure $u(dx,dy,t) = \hat{\sigma}^2(x,y,t)\,\mathbb{P}(X_t \in dx, Y_t \in dy)$, where $\hat{\sigma}^2(x,y,t) = \mathbb{E}(\sigma_t^2\,|\,X_t = x, Y_t = y)$ a.s.

Proof. Applying the Itô formula to $Z_t = e^{ikX_t - \lambda Y_t}$ we obtain

$$Z_t \;\;=\;\; 1 + \int_0^t ikZ_s\,dX_s \;-\; \frac{1}{2}k^2\int_0^t Z_s\sigma_s^2 ds \;-\; \lambda\int_0^t Z_s g(X_s,s)ds$$

and we then proceed as before. □

Remark 6.3. *Using* (6.1) *and* (6.2) *and performing a Laplace-Fourier inversion, we can back out $u(dx,dy,t)$ and thus $\hat{\sigma}(x,y,t)^2$ from $\phi(k,\lambda,t)$ at all $t \in (0,T]$.*

Remark 6.4. *The special case $g(x,t) = 1_{x>b}$ corresponds to the occupation time of X_t above the level b. The case $g(x,t) = x$ corresponds to t times the running average of X_t.*

7. Tri-variate Markovian projections

7.1. Computing $\mathbb{E}(\sigma_t^2\,|\,(X_t, Y_t, \int_0^t b(X_s)ds)$

Proposition 7.1. *Let (X_t, Y_t) satisfy the same conditions as Theorem 3.1 and assume that $g'(y) = c$[6] is constant and let $\Gamma_t = \int_0^t b(X_s,s)ds$, where b is bounded and Borel measurable and set $\phi(k,\lambda,\gamma,t) = \mathbb{E}(e^{ikX_t-\lambda Y_t-\gamma\Gamma_t})$ for $k \in \mathbb{R}, \lambda, \gamma \in \mathbb{C}$, $\mathrm{Re}(\lambda), \mathrm{Re}(\gamma) \geq 0$. Then we have the following forward equation*

$$\partial_t\mathbb{E}(Z_t + \frac{\lambda}{ikc-\lambda}\bar{Z}_t) \;\;=\;\; -\frac{1}{2}k^2\mathbb{E}(Z_t\sigma_t^2) - \;\gamma\mathbb{E}((Z_t - \frac{\lambda}{ikc-\lambda}\bar{Z}_t)b(X_t,t)) \tag{7.1}$$

where $Z_t = e^{ikX_t-\lambda L_t-\gamma\Gamma_t}$, $\bar{Z}_t = e^{-\lambda L_t-\gamma\Gamma_t}$ and $\mathbb{E}(Z_t\sigma_t^2)$ is the Fourier-Laplace transform of the bounded measure $u(dx,dy,dz,t) = \hat{\sigma}^2(x,y,z,t)\,\mathbb{P}(X_t \in dx, Y_t \in dy, \Gamma_t \in dz)$, where $\hat{\sigma}^2(X_t,Y_t,\Gamma_t) = \mathbb{E}(\sigma_t^2\,|\,X_t,Y_t,\Gamma_t)$ a.s.

Proof. See Appendix. □

Remark 7.2. *Using* (7.1) *and performing a Laplace-Fourier inversion, we can back out $u(dx,dy,dz,t)$ and thus $\hat{\sigma}(x,y,z,t)^2$ from $\phi(k,\lambda,\gamma,t)$ at all $t \in (0,T]$.*

Remark 7.3. *We can derive a similar forward equation for $(X_t, Y_t, \langle X\rangle_t)$ (we defer the details for the sake of brevity).*

[6] The local time and the maximum process both satisfy this condition.

References

[BC09] A.Bentata and R.Cont, "Forward equations for option prices in semimartingale models", 2009, to appear in *Finance and Stochastics.*

[BC09II] A.Bentata and R.Cont, "Mimicking the marginals of a semimartingale", 2009.

[BS12] Brunick, G. and S.Shreve, "Matching an Itô Process by a Solution of a Stochastic Differential Equation", 2011, to appear in *Ann. Appl. Probab.*

[CM02] Cattiaux, P and L.Mesnager, "Hypoelliptic non-homogeneous diffusions", *Probab. Theory Related Fields*, Vol. 123, 4, 453-483, 2002.

[Dup96] Dupire, B., "A unified theory of volatility", Paribas working paper, 1996. Also appears in "Derivatives Pricing: The Classic Collection", pages 185-196. Risk Books, 2004.

[Fig08] Figalli, A., "Existence and uniqueness of martingale solutions for SDEs with rough or degenerate coefficients", *J. Funct. Anal.*, Vol. 254, 1, 109-153, 2008.

[Gyö86] Gyöngy, I., "Mimicking the one-dimensional marginal distributions of processes having an Itô differential", *Prob. Th. Rel. Fields*, Vol. 71, 4, 501-516, 1986.

[KS91] Karatzas, I. and S.Shreve, "Brownian motion and Stochastic Calculus", Springer-Verlag, 1991.

[Kle02] Klebaner, F., (2002), "Option pricing when the stock is a semimartingale", *Elect. Comm. in Probab.*, 7, 79-83.

[SV79] Stroock, D.W. and S.R.S.Varadhan, "Multidimensional diffusion processes", Springer-Verlag, Berlin-New York, 1979.

Appendix A: Proof of Proposition 5.2

X_t and Y_t are continuous semimartingales, so we can apply Itô's formula to the test function ψ:

$$d\psi(X_t,Y_t,t) \;=\; \psi_t(X_t,Y_t,t)dt + \psi_x(X_t,Y_t,t)dX_t + \psi_y(X_t,Y_t,t)dY_t + \frac{1}{2}\psi_{xx}(X_t,Y_t,t)\sigma_t^2 dt$$

Integrating and using that $\psi(X_0,Y_0,0)=\psi(0,0,0)=0$, we obtain

$$\begin{aligned}\psi(X_t,Y_t,t) \;&=\; \int_0^t \psi_t(X_s,Y_s,s)ds + \int_0^t \psi_x(X_s,Y_s,s)dX_s + \int_0^t \psi_y(X_s,Y_s,s)dY_s \\ &+\; \int_0^t \frac{1}{2}\psi_{xx}(X_s,Y_s,s)\sigma_s^2 ds\,.\end{aligned}$$

ψ_x, ψ_{xx}, ψ_y and ψ_t are bounded because $\psi \in C_b^{2,1,1}$ and $\mathbb{E}(\int_0^t \sigma_s^2 ds) < \infty$ by assumption so we can take expectations, apply Fubini's theorem and re-arrange to obtain

$$\mathbb{E}(\psi(X_t,Y_t,t) \;=\; \int_0^t \mathbb{E}((\psi_t(X_s,Y_s,s) + \psi_y(X_s,Y_s,s)b_s + \frac{1}{2}\psi_{xx}(X_s,Y_s,s)\sigma_s^2)ds\,.$$

Finally, using iterated expectations we obtain the result.

Appendix B: Proof of Proposition 7.1

Applying the Itô formula to $Z_t = e^{ikX_t-\lambda Y_t-\gamma\Gamma_t}$ we obtain

$$dZ_t \;=\; ikZ_t dX_t - \frac{1}{2}k^2 Z_t\sigma_t^2 - \lambda\bar{Z}_t dY_t - \gamma Z_t b(X_t,t)dt\,,$$

where $\bar{Z}_t = e^{ikg(Y_t)-\lambda Y_t-\gamma\Gamma_t}$, and we have used that $X_t = g(Y_t)$ on the growth set of Y_t in the penultimate term. But from Ito's lemma we also have

$$d\bar{Z}_t \;=\; \bar{Z}_t((ikc-\lambda)\,dY_t - \gamma b(X_t,t)dt).$$

Combining both expressions, we obtain

$$dZ_t \;=\; ikZ_t dX_t - \frac{1}{2}k^2 Z_t\sigma_t^2 - \frac{\lambda}{ikc-\lambda}\,(d\bar{Z}_t + \gamma b(X_t,t)dt) - \gamma Z_t b(X_t,t)dt$$

and we then proceed as before.

Appendix C: Proof of Lemma 2.1

Similar to the proof of Proposition 4.4. in [Gyö86], we now recall the definition of $\hat{\sigma}^2(x,y,t) = \mathbb{E}(\sigma_t^2|X_t = x, Y_t = y)$ via the Radon-Nikodým theorem: we consider the measure q_t defined by the formula

$$q_t(A) \quad = \quad \mathbb{E}(1_{(X_t,Y_t)\in A}\,\sigma_t^2)\,. \tag{C-1}$$

for every $B \in \mathcal{B}(\mathbb{R}\times\mathbb{R}^+)$. q_t is absolutely continuous with respect to $p(dx, dy, t)$ (the distribution of (X_t, Y_t) on $\mathbb{R}\times\mathbb{R}^+$). Thus, by the Radon-Nikodým theorem, there exists a measurable function $\hat{\sigma}^2(t, ., .)$ such that

$$q_t(A) \quad = \quad \int_{\mathbb{R}\times\mathbb{R}^+} 1_A\,\hat{\sigma}^2(x,y,t)\,p(dx,dy,t)\,.$$

For every t, $\hat{\sigma}^2(t, ., .)$ is unique up to a set of p_t-measure 0. We then define $\mathbb{E}(\sigma_t^2|X_t = x, Y_t = y) = \hat{\sigma}^2(x,y,t)$ and from the standard Kolmogorov definition of conditional expectation, $\mathbb{E}(\sigma_t^2|X_t, Y_t) = \hat{\sigma}^2(X_t, Y_t, t)$ a.s.