On the positivity of local mild solutions to stochastic evolution equations

Carlo Marinelli[1] and Luca Scarpa[2]

[1] Department of Mathematics, University College London,
Gower Street, London WC1E 6BT, United Kingdom
`c.marinelli@ucl.ac.uk`
WWW home page: `https://www.homepages.ucl.ac.uk/~ucahcm0/`
[2] Faculty of Mathematics, University of Vienna,
Oskar-Morgenstern-Platz 1, 1090 Wien, Austria
`luca.scarpa@univie.ac.at`
WWW home page: `http://www.mat.univie.ac.at/~scarpa`

Abstract. We provide sufficient conditions on the coefficients of a stochastic evolution equation on a Hilbert space of functions driven by a cylindrical Wiener process ensuring that its mild solution is positive if the initial datum is positive. As an application, we discuss the positivity of forward rates in the Heath-Jarrow-Morton model via Musiela's stochastic PDE.

Keywords: Positivity, mild solutions, stochastic evolution equations.

1 Introduction

Let us consider a stochastic evolution equation of the type

$$du + Au\,dt = F(u)\,dt + B(u)\,dW, \qquad u(0) = u_0, \tag{1}$$

where A is a linear maximal monotone operator on a Hilbert space of functions H, the coefficients F and B satisfy suitable integrability assumptions, and W is a cylindrical Wiener process. Precise assumptions on the data of the Cauchy problem 1 are given in §2 below. Our goal is to establish a maximum principle for (local) mild solutions to (1), i.e. to provide sufficient conditions on the operator A and on the coefficients F and B such that positivity of the initial datum u_0 implies positivity of the solution u (see Theorem 2 below).

A simpler problem was studied in [10], where coefficients F and B are assumed to be Lipschitz continuous. Here we simply assume that F and B satisfy rather minimal integrability conditions and that a local mild solution exists. On the other hand, in [10] the linear operator A need only generate a positivity preserving semigroup, while here we require that A generates a sub-Markovian semigroup.

We refer to [10] for a discussion about the relation of other positivity results for solutions to stochastic partial differential equations with ours. It is however probably worth pointing out that most existing results seem to deal with equations in the variational setting (see, e.g., [1, 7, 8, 13]).

As an application, we provide an alternative, more direct proof of the positivity of forward rates in the Heath-Jarrow-Morton [5] framework with respect to the one in [10]. This is obtained, as is now classical, viewing forward curves as solutions to the so-called Musiela stochastic PDE (see, e.g., [3, 11]).

2 Assumptions and main result

Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space endowed with a complete right-continuous filtration $(\mathscr{F}_t)_{t\in[0,T]}$, with $T > 0$ a fixed final time, on which all random elements will be defined. Identities and inequalities between random variables are meant to hold $\mathbb{P}$-almost surely, and two stochastic processes are declared equal, unless otherwise stated, if they are indistinguishable. The σ-algebra of progressively measurable subsets of $\Omega \times [0,T]$ will be denoted by $\mathscr{R}$. We shall denote a cylindrical Wiener process on a separable Hilbert space U by W. Standard notation and terminology of stochastic calculus for semimartingales will be used throughout (see, e.g., [12]). In particular, given an adapted process X and a stopping time τ, X^τ will denote the process X stopped at τ. Similarly, if X is also càdlàg, $X^{\tau-}$ stands for the process X pre-stopped at τ.

For any separable Hilbert spaces E_1 and E_2, we use the symbols $\mathscr{L}(E_1, E_2)$ $\mathscr{L}^2(E_1, E_2)$ for the space of linear continuous and Hilbert-Schmidt operators from E_1 to E_2, respectively. The space of continuous bilinear maps from $E_1 \times E_1$ to E_2 will be denoted by $\mathscr{L}_2(E_1; E_2)$. The n-th order Fréchet and Gâteaux derivatives of a function $\Phi : E_1 \to E_2$ at a point $x \in E_1$ are denoted by $D^n\Phi(x)$ and $D^n_{\mathcal{G}}\Phi(x)$, respectively, omitting the superscript if $n = 1$, as usual.

We shall work under the following standing assumptions.

(A1) There exists an open set $\mathcal{O}$ in $\mathbb{R}^d$, $d \geq 1$, and a Borel measure μ such that $H = L^2(\mathcal{O}, \mu)$.

The norm and scalar product on H will be denoted by $\|\cdot\|$ and $\langle \cdot, \cdot \rangle$, respectively.

(A2) A is a linear maximal monotone operator on H such that its resolvent is sub-Markovian and is a contraction with respect to the $L^1(\mathcal{O}, \mu)$-norm.

Recall that the resolvent of A, i.e. the family of linear continuous operators on H defined by

$$J_\lambda := (I + \lambda A)^{-1}, \qquad \lambda > 0,$$

is said to be sub-Markovian if, for every $\lambda > 0$ and every $\phi \in H$ such that $0 \leq \phi \leq 1$ a.e. in $\mathcal{O}$, one has $0 \leq J_\lambda \phi \leq 1$ a.e. in $\mathcal{O}$.

(A3) $F : \Omega \times [0,T] \times H \to H$ and $B : \Omega \times [0,T] \times H \to \mathscr{L}^2(U, H)$ are $\mathscr{R} \otimes \mathscr{B}(H)$-measurable, and there exists a constant $C > 0$ such that

$$-\langle F(\omega, t, h), h_- \rangle + \frac{1}{2} \left\| 1_{\{h<0\}} B(\omega, t, h) \right\|^2_{\mathscr{L}^2(U,H)} \leq C \|h_-\|^2_{L^2(\mathcal{O})}$$

for all $(\omega, t, h) \in \Omega \times [0,T] \times H$. In particular, note that choosing $h = 0$ yields $F(\cdot, 0) = 0$ and $B(\cdot, 0) = 0$.

(A4) $u_0 \in L^0(\Omega, \mathscr{F}_0; H)$

Definition 1 *A local mild solution to the Cauchy problem* (1) *is a pair* (u, τ), *where* τ *is a stopping time with* $\tau \leq T$, *and* $u : [\![0, \tau[\![\to H$ *is a measurable adapted process with continuous trajectories such that, for any stopping time* $\sigma < \tau$, *one has*

(i) $S(t-\cdot)F(u)\mathbb{1}_{[\![0,\sigma]\!]} \in L^0(\Omega; L^1(0,t;H))$ *for all* $t \in [0,T]$;
(ii) $S(t-\cdot)B(u)\mathbb{1}_{[\![0,\sigma]\!]} \in L^0(\Omega; L^2(0,t;\mathscr{L}^2(U,H)))$ *for all* $t \in [0,T]$,

and

$$u = S(\cdot)u_0 + \int_0^{\cdot} S(\cdot - s)F(s,u(s))\,ds + \int_0^{\cdot} S(\cdot - s)B(s,u(s))\,dW(s).$$

The last identity is to be understood in the sense of indistinguishability of processes defined on the stochastic interval $[\![0,\tau[\![$. Here the stochastic convolution is defined on $[\![0,\sigma]\!]$, for every stopping time $\sigma < \tau$, as

$$\left(\int_0^t S(t-s)B(s,u(s))\mathbb{1}_{[\![0,\sigma]\!]}(s)\,dW(s)\right)_{t\in[0,\sigma]}.$$

The main result is the following.

Theorem 2 *Let* (u,τ) *be a local mild solution to the Cauchy problem* (1) *such that, for every stopping time* $\sigma < \tau$, *one has*

(i) $F(u)\mathbb{1}_{[\![0,\sigma]\!]} \in L^0(\Omega; L^1(0,T;H))$;
(ii) $B(u)\mathbb{1}_{[\![0,\sigma]\!]} \in L^0(\Omega; L^2(0,T;\mathscr{L}^2(U,H)))$.

If $u_0 \geq 0$ *a.e. in* $\mathcal{O}$, *then* $u^{\tau-}(t) \geq 0$ *a.e. in* $\mathcal{O}$ *for all* $t \in [0,T]$.

3 Auxiliary results

The arguments used in the proof of Theorem 2 (see §4 below) rely on the following results, that we recall here for the reader's convenience. The first is a continuous dependence result for mild solutions to stochastic evolution equations in the form (1) with respect to the coefficients and the initial datum. This is a consequence of a more general statement proved in [9, Corollary 3.4]. Let

$$\begin{gathered}(u_{0n})_n \subset L^0(\Omega, \mathscr{F}_0; H),\\ (f_n)_n, f \subset L^0(\Omega; L^1(0,T;H)),\\ (G_n)_n, G \subset L^0(\Omega; L^2(0,T;\mathscr{L}^2(U,H)))\end{gathered}$$

be such that the H-valued processes f_n, f, $G_n v$, and Gv are strongly measurable and adapted for all $v \in U$ and $n \in \mathbb{N}$. Then the Cauchy problems

$$du_n + Au_n\,dt = f_n\,dt + G_n\,dW, \qquad u_n(0) = u_{0n},$$

and

$$du + Au\,dt = f\,dt + G\,dW, \qquad u(0) = u_0,$$

admit unique mild solutions u_n and u, respectively.

Proposition 3 *Assume that*

$$\begin{aligned} u_{0n} &\longrightarrow u_0 && \text{in } L^0(\Omega; H),\\ f_n &\longrightarrow f && \text{in } L^0(\Omega; L^1(0,T;H)),\\ G_n &\longrightarrow G && \text{in } L^0(\Omega; L^2(0,T;\mathscr{L}^2(U,H))). \end{aligned}$$

Then $u_n \to u$ in $L^0(\Omega; C([0,T];H))$.

The second result we shall need is a generalized Itô formula, the proof of which can be found in [10].

Proposition 4 *Let $G\colon H \to \mathbb{R}$ be continuously Fréchet differentiable and DG be Gâteaux differentiable, with $D^2_{\mathcal{G}}G\colon H \to \mathscr{L}_2(H;\mathbb{R})$ such that $(\varphi,\zeta_1,\zeta_2) \mapsto D^2_{\mathcal{G}}G(\varphi)[\zeta_1,\zeta_2]$ is continuous, and assume that G, DG, and $D^2_{\mathcal{G}}G$ are polynomially bounded. Moreover, let the processes $f \in L^0(\Omega; L^1(0,T;H))$ and $\Phi \in L^0(\Omega; L^2(0,T;\mathscr{L}^2(U,H)))$ be measurable and adapted, and $v_0 \in L^0(\Omega, \mathscr{F}_0; H)$. Setting*

$$v := v_0 + \int_0^\cdot f(s)\,ds + \int_0^\cdot \Phi(s)\,dW(s),$$

one has

$$\begin{aligned} G(v) = G(v_0) &+ \int_0^\cdot \Big(DG(v)f + \frac{1}{2}\operatorname{Tr}\big(\Phi^* D^2_{\mathcal{G}}G(v)\Phi\big)\Big)(s)\,ds\\ &+ \int_0^\cdot DG(v(s))\Phi(s)\,dW(s). \end{aligned}$$

Finally, we recall an inequality for maximal monotone linear operators with sub-Markovian resolvent, due to Brézis and Strauss (see [2, Lemma 2]).[3]

Lemma 5 *Let $\beta\colon \mathbb{R} \to 2^{\mathbb{R}}$ be a maximal monotone graph with $0 \in \beta(0)$. Let $\varphi \in L^p(\mathcal{O})$ with $A\varphi \in L^p(\mathcal{O})$, and $z \in L^q(\mathcal{O})$ with $z \in \beta(\varphi)$ a.e. in $\mathcal{O}$, where $p,q \in [1,+\infty]$ and $1/p + 1/q = 1$. Then*

$$\int_{\mathcal{O}} (A\varphi)z \geq 0.$$

We include a sketch of proof for the reader's convenience, assuming for simplicity that $\beta : \mathbb{R} \to \mathbb{R}$ is continuous and bounded. Let $j : \mathbb{R} \to \mathbb{R}_+$ a (differentiable, convex) primitive of β and

$$A_\lambda := \frac{1}{\lambda}\big(I - (I+\lambda A)^{-1}\big) = \frac{1}{\lambda}(I - J_\lambda), \qquad \lambda > 0,$$

the Yosida approximation of A. It is well known that A_λ is a linear maximal monotone bounded operator on H and that, for every $v \in \mathsf{D}(A)$, $A_\lambda v \to Av$ as

[3] For a related inequality cf. also [14, Lemma 5.1].

$\lambda \to 0$. Let $v \in \mathsf{D}(A)$. The convexity of j implies, for every $\lambda > 0$,

$$\begin{aligned}\bigl\langle A_\lambda v, \beta(v)\bigr\rangle_{L^2} &= \frac{1}{\lambda}\bigl\langle v - J_\lambda v, j'(v)\bigr\rangle_{L^2}\\ &\geq \frac{1}{\lambda}\biggl(\int_{\mathcal{O}} j(v) - \int_{\mathcal{O}} j(J_\lambda v)\biggr) = \frac{1}{\lambda}\bigl(\|j(v)\|_{L^1} - \|j(J_\lambda v)\|_{L^1}\bigr).\end{aligned}$$

Since J_λ is sub-Markovian and j is convex, the generalized Jensen inequality for positive operators (see [4]) and the contractivity of J_λ in L^1 imply that

$$\bigl\|j(J_\lambda v)\bigr\|_{L^1} \leq \bigl\|J_\lambda j(v)\bigr\|_{L^1} \leq \bigl\|j(v)\bigr\|_{L^1},$$

i.e. that

$$\bigl\langle A_\lambda v, \beta(v)\bigr\rangle_{L^2} \geq 0$$

for every $\lambda \to 0$. Passing to the limit as $\lambda \to 0$ yields $\langle Av, \beta(v)\rangle_{L^2} \geq 0$.

4 Proof of Theorem 2

The proof is divided into two parts. First we show that a local mild solution u to (1) can be approximated by strong solutions to regularized equations. As a second step, we show that such approximating processes are positive, thanks to a suitable version of Itô's formula.

4.1 Approximation of the solution

Let (u,τ) be a local mild solution to (1). Let σ be a stopping time with $\sigma < \tau$, so that $u : [\![0,\sigma]\!] \to H$ is well defined, and set

$$\begin{aligned}\bar{u} &:= u^\sigma \in L^0(\Omega; C([0,T];H)),\\ \bar{F} &:= F(\cdot,u)\mathbb{1}_{[\![0,\sigma]\!]} \in L^0(\Omega; L^1(0,T;H)),\\ \bar{B} &:= B(\cdot,u)\mathbb{1}_{[\![0,\sigma]\!]} \in L^0(\Omega; L^2(0,T;\mathscr{L}^2(U,H))).\end{aligned}$$

Note that, by assumption (A3), $F(\cdot,0) = 0$ and $B(\cdot,0) = 0$, hence

$$\begin{aligned}\bar{F} &= F(\cdot,u)\mathbb{1}_{[\![0,\sigma]\!]} = F(\cdot,u\mathbb{1}_{[\![0,\sigma]\!]}),\\ \bar{B} &= B(\cdot,u)\mathbb{1}_{[\![0,\sigma]\!]} = B(\cdot,u\mathbb{1}_{[\![0,\sigma]\!]}).\end{aligned}$$

In particular, one has

$$\bar{u}(t) := S(t)u_0 + \int_0^t S(t-s)\bar{F}(s)\,ds + \int_0^t S(t-s)\bar{B}(s)\,dW(s) \tag{2}$$

for all $t \in [0,T]$ $\mathbb{P}$-a.s., or, equivalently, $\bar{u}$ is the unique global mild solution to the Cauchy problem

$$d\bar{u} + A\bar{u}\,dt = \bar{F}\,dt + \bar{B}\,dW, \qquad \bar{u}(0) = u_0.$$

Recalling that $J_\lambda \in \mathscr{L}(H, \mathsf{D}(A))$ for all $\lambda > 0$, one has

$$\begin{aligned}
\bar{F}_\lambda &:= J_\lambda F(\cdot, u)\mathbb{1}_{[\![0,\sigma]\!]} = J_\lambda \bar{F} \in L^0(\Omega; L^1(0,T;\mathsf{D}(A))),\\
\bar{B}_\lambda &:= J_\lambda B(\cdot, u)\mathbb{1}_{[\![0,\sigma]\!]} = J_\lambda \bar{B} \in L^0(\Omega; L^2(0,T;\mathscr{L}^2(U,\mathsf{D}(A)))),\\
u_{0\lambda} &:= J_\lambda u_0 \in L^0(\Omega, \mathscr{F}_0; \mathsf{D}(A)),
\end{aligned}$$

where the second assertion is an immediate consequence of the ideal property of Hilbert-Schmidt operators. The process $u_\lambda : \Omega \times [0,T] \to H$ defined as

$$u_\lambda(t) := S(t)u_{0\lambda} + \int_0^t S(t-s)\bar{F}_\lambda(s)\,ds + \int_0^t S(t-s)\bar{B}_\lambda(s)\,dW(s), \qquad t \in [0,T], \tag{3}$$

therefore belongs to $L^0(\Omega; C([0,T];\mathsf{D}(A)))$ and is the unique global strong solution to the Cauchy problem

$$du_\lambda + Au_\lambda\,dt = \bar{F}_\lambda\,dt + \bar{B}_\lambda\,dW, \qquad u_\lambda(0) = u_{0\lambda},$$

i.e.

$$u_\lambda + \int_0^\cdot Au_\lambda(s)\,ds = u_{0\lambda} + \int_0^\cdot \bar{F}_\lambda(s)\,ds + \int_0^\cdot \bar{B}_\lambda(s)\,dW(s) \tag{4}$$

in the sense of indistinguishable H-valued processes. Furthermore, since J_λ is contractive and converges to the identity in the strong operator topology of $\mathscr{L}(H,H)$ as $\lambda \to 0$, i.e. $J_\lambda h \to h$ for every $h \in H$, one has

$$\begin{aligned}
u_{0\lambda} &\longrightarrow u_0 \quad \text{in } L^0(\Omega; H),\\
\bar{F}_\lambda &\longrightarrow \bar{F} \quad \text{in } L^0(\Omega; L^2(0,T;H)),\\
\bar{B}_\lambda &\longrightarrow \bar{B} \quad \text{in } L^0(\Omega; L^2(0,T;\mathscr{L}^2(U,H))),
\end{aligned}$$

where the second convergence follows immediately by the dominated convergence theorem, and the third one by a continuity property of Hilbert-Schmidt operators (see, e.g., [6, Theorem 9.1.14]). Finally, thanks to Proposition 3, we deduce that

$$u_\lambda \longrightarrow \bar{u} \quad \text{in } L^0(\Omega; C([0,T];H)). \tag{5}$$

4.2 Positivity

Let us introduce the functional

$$\begin{aligned}
G\colon H &\longrightarrow \mathbb{R}_+,\\
G\colon \varphi &\longmapsto \frac{1}{2}\int_{\mathcal{O}} |\varphi_-|^2,
\end{aligned}$$

as well as the family, indexed by $n \in \mathbb{N}$, of regularized functionals

$$\begin{aligned}
G_n\colon H &\longrightarrow \mathbb{R}_+,\\
G_n\colon \varphi &\longmapsto \frac{1}{2}\int_{\mathcal{O}} g_n(\varphi),
\end{aligned}$$

where $g_n : \mathbb{R} \to \mathbb{R}_+$ is convex, twice continuously differentiable, identically equal to zero on $\mathbb{R}_+$, strictly positive and decreasing on $\mathbb{R}_-$, such that (g_n'') is uniformly bounded, and $g_n'(r) \to -r^-$ as $n \to \infty$ for all $r \in \mathbb{R}$. The existence of such an approximating sequence is well known (see, e.g., [15, §3] for details). One can verify (see, e.g., [10]) that, for every $n \in \mathbb{N}$, G_n is everywhere continuously Fréchet differentiable with derivative

$$\begin{aligned} DG_n \colon H &\longrightarrow \mathscr{L}(H,\mathbb{R}) \simeq H,\\ DG_n \colon \varphi &\longmapsto g_n'(\varphi), \end{aligned}$$

and that $DG_n \colon H \to H$ is Gâteaux differentiable with Gâteaux derivative given by

$$\begin{aligned} D^2_{\mathcal{G}} G_n \colon H &\longrightarrow \mathscr{L}(H,H) \simeq \mathscr{L}_2(H;\mathbb{R}),\\ D^2_{\mathcal{G}} G_n \colon \varphi &\longmapsto \Bigl[(\zeta_1,\zeta_2) \mapsto \int_{\mathcal{O}} g_n''(\varphi)\zeta_1\zeta_2\Bigr]. \end{aligned}$$

Furthermore, the map $(\varphi,\zeta_1,\zeta_2) \mapsto D^2_{\mathcal{G}} G_n(\varphi)(\zeta_1,\zeta_2)$ is continuous. Proposition 4 applied to the process u_λ defined by (4) then yields

$$\begin{aligned} G_n(u_\lambda) &+ \int_0^\cdot \bigl\langle Au_\lambda, DG_n(u_\lambda)\bigr\rangle(s)\,ds\\ &= G_n(u_{0\lambda}) + \int_0^\cdot DG_n(u_\lambda(s))\bar{B}_\lambda(s)\,dW(s)\\ &\quad + \int_0^\cdot \Bigl(DG_n(u_\lambda)\bar{F}_\lambda + \frac{1}{2}\operatorname{Tr}\bigl(\bar{B}_\lambda^* D^2_{\mathcal{G}} G_n(u_\lambda)\bar{B}_\lambda\bigr)\Bigr)(s)\,ds \end{aligned} \tag{6}$$

Recalling that $g_n' : \mathbb{R} \to \mathbb{R}$ is increasing, Lemma 5 implies that

$$\langle Au_\lambda, DG_n(u_\lambda)\rangle = \langle Au_\lambda, g'(u_\lambda)\rangle \geq 0, \tag{7}$$

hence also, denoting a complete orthonormal system of U by (e_j),

$$\begin{aligned} \int_{\mathcal{O}} g_n(u_\lambda(t)) \leq \int_{\mathcal{O}} g_n(u_{0\lambda}) &+ \int_0^t g_n'(u_\lambda(s))\bar{B}_\lambda(s)\,dW(s)\\ &+ \int_0^t g_n'(u_\lambda(s))\bar{F}_\lambda(s)\,ds + \frac{1}{2}\int_0^t \sum_{j=0}^\infty \int_{\mathcal{O}} g_n''(u_\lambda(s))\bigl|\bar{B}_\lambda(s)e_j\bigr|^2\,ds \end{aligned}$$

for every $t \in [0,T]$ and $n \in \mathbb{N}$. We are now going to pass to the limit as $n \to \infty$ in this inequality. Recalling that (g_n'') is uniformly bounded and that the paths of u_λ belong to $C([0,T];H)$ $\mathbb{P}$-a.s., the dominated convergence theorem yields

$$\begin{aligned} \int_{\mathcal{O}} g_n(u_\lambda(t)) &\longrightarrow \frac{1}{2}\bigl\|u_\lambda^-(t)\bigr\|^2 \quad \forall t \in [0,T],\\ \int_{\mathcal{O}} g_n(u_{0\lambda}) &\longrightarrow \frac{1}{2}\bigl\|u_{0\lambda}^-\bigr\|^2. \end{aligned}$$

Note that u_0 is positive and J_λ is positivity preserving, hence $u_{0\lambda} = J_\lambda u_0$ is also positive and, in particular, $u_{0\lambda}^-$ is equal to zero a.e. in $\mathcal{O}$. Let us introduce the (real) continuous local martingales $(M^n)_{n\in\mathbb{N}}$, M, defined as

$$\begin{aligned} M_t^{\lambda,n} &:= \int_0^t g_n'(u_\lambda(s))\bar{B}_\lambda(s)\,dW(s),\\ M_t^{\lambda} &:= -\int_0^t (u_\lambda^-(s)\bar{B}_\lambda(s)\,dW(s). \end{aligned}$$

One has, by the ideal property of Hilbert-Schmidt operators,

$$\begin{aligned} \bigl[M^{\lambda,n} - M^\lambda, M^{\lambda,n} - M^\lambda\bigr]_t &= \int_0^t \bigl\|(g_n'(u_\lambda(s)) + u_\lambda^-(s))\bar{B}_\lambda(s)\bigr\|^2_{\mathscr{L}^2(U,\mathbb{R})}\,ds\\ &\leq \int_0^t \bigl\|g_n'(u_\lambda(s)) + u_\lambda^-(s)\bigr\|^2\bigl\|\bar{B}_\lambda(s)\bigr\|^2_{\mathscr{L}^2(U,H)}\,ds \end{aligned}$$

for all $t \in [0,T]$. Recalling that $u_\lambda \in L^0(\Omega; C([0,T];H))$ and $g_n'(r) \to -r^-$ for every $r \in \mathbb{R}$, it follows by the dominated convergence theorem that $[M^{\lambda,n} - M^\lambda, M^{\lambda,n} - M^\lambda] \to 0$, hence that $M^{\lambda,n} \to M^\lambda$, as $n \to \infty$, i.e. that

$$\int_0^t g_n'(u_\lambda(s))\bar{B}_\lambda(s)\,dW(s) \longrightarrow -\int_0^t u_\lambda^-(s)\bar{B}_\lambda(s)\,dW(s)$$

for all $t \in [0,T]$. Similarly, the pathwise continuity of u_λ and the dominated convergence theorem yield

$$\int_0^t g_n'(u_\lambda(s))\bar{F}_\lambda(s)\,ds \longrightarrow -\int_0^t u_\lambda^-(s)\bar{F}_\lambda(s)\,ds$$

for all $t \in [0,T]$ as $n \to \infty$. Finally, the pointwise convergence $g_n'' \to 1\!\!1_{\mathbb{R}_-}$ and the dominated convergence theorem imply that

$$\int_0^t \sum_{j=0}^\infty \int_{\mathcal{O}} g_n''(u_\lambda(s))\bigl|\bar{B}_\lambda(s)e_j\bigr|^2\,ds \longrightarrow \int_0^t \sum_{j=0}^\infty \int_{\mathcal{O}} 1\!\!1_{\{u_\lambda(s)<0\}}\bigl|\bar{B}_\lambda(s)e_j\bigr|^2\,ds$$

for all $t \in [0,T]$ as $n \to \infty$. We are thus left with

$$\begin{aligned} \bigl\|u_\lambda^-(t)\bigr\|^2 \leq \int_0^t \Bigl(-2\langle u_\lambda^-(s), \bar{F}_\lambda(s)\rangle + \sum_{j=0}^\infty \int_{\mathcal{O}} 1\!\!1_{\{u_\lambda(s)<0\}}\bigl|\bar{B}_\lambda(s)e_j\bigr|^2\Bigr)\,ds\\ - \int_0^t u_\lambda^-(s)\bar{B}_\lambda(s)\,dW(s). \end{aligned}$$

Let us now take the limit as $\lambda \to 0$: if follows from the convergence property (5) and the continuous mapping theorem that

$$\bigl\|u_\lambda^-(t)\bigr\|^2 \longrightarrow \bigl\|\bar{u}^-(t)\bigr\|^2.$$

Recalling that $\bar{F}_\lambda = J_\lambda F(\bar{u})$, which converges pointwise to $F(\bar{u})$, one has

$$\int_0^t -2\big\langle u_\lambda^-(s), \bar{F}_\lambda(s)\big\rangle\,ds \longrightarrow \int_0^t -2\big\langle \bar{u}^-(s), F(s,\bar{u}(s))\big\rangle\,ds$$

Appealing again to (5), it is not difficult to check that

$$\mathbb{1}_{\{\bar{u}(s)<0\}} \leq \liminf_{\lambda\to 0} \mathbb{1}_{\{u_\lambda(s)<0\}} \quad \text{a.e. in } \mathcal{O} \quad \forall s \in [0,T].$$

Hence it follows from Fatou's lemma that

$$\int_0^t \sum_{j=0}^\infty \int_{\mathcal{O}} \mathbb{1}_{\{u_\lambda(s)<0\}} \big|\bar{B}_\lambda(s)e_j\big|^2\,ds \longrightarrow \int_0^t \sum_{j=0}^\infty \int_{\mathcal{O}} \mathbb{1}_{\{\bar{u}(s)<0\}} \big|B(s,\bar{u}(s))e_j\big|^2\,ds.$$

Let us define the real continuous local martingales $(M^\lambda)_{\lambda>0}$, M, defined as

$$\begin{aligned} M_t^\lambda &:= -\int_0^t u_\lambda^-(s)\bar{B}_\lambda(s)\,dW(s),\\ M_t &:= -\int_0^t \bar{u}^-(s)\bar{B}(s)\,dW(s). \end{aligned}$$

One has

$$\big[M^\lambda - M, M^\lambda - M\big]_t = \int_0^t \big\|u_\lambda^-(s)\bar{B}_\lambda(s) - \bar{u}^-(s)\bar{B}(s)\big\|^2_{\mathscr{L}^2(U,\mathbb{R})}\,ds,$$

where, by the ideal property of Hilbert-Schmidt operators and the contractivity of J_λ,

$$\begin{aligned} \big\|u_\lambda^-\bar{B}_\lambda - \bar{u}^-\bar{B}\big\|_{\mathscr{L}^2(U,\mathbb{R})} &\leq \big\|(u_\lambda^- - \bar{u}^-)\bar{B}_\lambda\big\|_{\mathscr{L}^2(U,\mathbb{R})} + \big\|\bar{u}^-(\bar{B}_\lambda - \bar{B})\big\|_{\mathscr{L}^2(U,\mathbb{R})}\\ &\leq \big\|u_\lambda^- - \bar{u}^-\big\|\big\|\bar{B}\big\|_{\mathscr{L}^2(U,H)} + \big\|\bar{u}^-\big\|\big\|\bar{B}_\lambda - \bar{B}\big\|_{\mathscr{L}^2(U,H)}. \end{aligned}$$

Since u_λ converges to $\bar{u}$ in the sense of (5) and, as already seen, $\bar{B}_\lambda \to \bar{B}$ in $L^0(\Omega; L^2(0,T;\mathscr{L}^2(U,H)))$, the dominated convergence theorem yields, for every $t \in [0,T]$,

$$\big[M^\lambda - M, M^\lambda - M\big]_t \longrightarrow 0,$$

thus also

$$\int_0^t u_\lambda^-(s)\bar{B}_\lambda(s)\,dW(s) \longrightarrow \int_0^t \bar{u}^-(s)B(s,\bar{u}(s))\,dW(s).$$

Recalling assumption (A3), one obtains, for every $t \in [0,T]$,

$$\|\bar{u}^-(t)\|^2 \leq 2C\int_0^t \|\bar{u}^-(s)\|^2\,ds - 2\int_0^t \bar{u}^-(s)B(s,\bar{u}(s))\,dW(s),$$

thus also, integrating by parts,

$$e^{-2Ct}\|\bar{u}^-(t)\|^2 \leq -2\int_0^t e^{-2Cs}\bar{u}^-(s)B(s,\bar{u}(s))\,dW(s) =: \tilde{M}_t.$$

The process $\tilde{M}$ is a positive local martingale, hence a supermartingale, with $\tilde{M}(0) = 0$, therefore M is identically equal to zero. This implies that $\|\bar{u}^-(t)\| = 0$ for all $t \in [0,T]$, hence, in particular, that $u(t)$ is positive a.e. in $\mathcal{O}$ for all $t \in [0,T]$. By definition of $\bar{u}$, we deduce that

$$u^\sigma \geq 0 \quad \text{a.e. in } \Omega \times [0,T] \times \mathcal{O}$$

for every $\sigma < \tau$. Since σ is arbitrary, this readily implies that

$$u \geq 0 \quad \text{a.e. in } [\![0,\tau[\![\times \mathcal{O},$$

thus completing the proof of Theorem 2.

Remark 6 *In [10] the substantially weaker assumption was made that $-A$ generates a positive semigroup. This was possible because F and B were assumed to be Lipschitz continuous. In fact, in this case the process u_λ, strong solution of the equation obtained by replacing A with its Yosida approximation A_λ in (1), i.e.*

$$du_\lambda + A_\lambda u_\lambda \, dt = F(u_\lambda)\, dt + B(u_\lambda)\, dW, \qquad u_\lambda(0) = u_0,$$

converges to the unique mild solution u to (1), and the positivity of u_λ, for every $\lambda > 0$, was shown. In the more general situation considered here, where F and B are not supposed to be Lipschitz continuous, it is not even clear whether the above regularized equation admits a solution at all. For this reason we introduced a different approximation scheme in §4.1, that implies the need for an estimate such as (7), which in turn is satisfied if $-A$ generates a sub-Markovian semigroup, rather than just a positive one.

5 Positivity of forward rates

Musiela's stochastic PDE can be written as

$$du + Au\, dt = \beta(t,u)\, dt + \sum_{k=1}^{\infty} \sigma_k(t,u)\, dw^k(t), \qquad u(0) = u_0, \tag{8}$$

where $-A$ is (formally, for the moment) the infinitesimal generator of the semigroup of translations, $(w^k)_{k\in\mathbb{N}}$ is a sequence of independent standard Wiener processes, σ_k is a random, time-dependent superposition operator for each $k \in \mathbb{N}$, as well as β, and u takes values in a space of continuous functions, so that $u(t,x) := [u(t)](x)$, $x \geq 0$, models the value of the forward rate prevailing at time t for delivery at time $t+x$. In order to exclude arbitrage (or, more precisely, in order for the corresponding discounted bond price process to be a local martingale), β needs to satisfy the so-called Heath-Jarrow-Morton no-arbitrage condition

$$\beta(t,v) = \sum_{k=1}^{\infty} \sigma_k(t,v) \int_0^{\cdot} [\sigma_k(t,v)](y)\, dy.$$

In order for (8) to admit a solution with continuous paths, a by now standard choice of state space is the Hilbert space H_α, $\alpha > 0$, which consists of absolutely continuous functions $\phi : \mathbb{R}_+ \to \mathbb{R}$ such that

$$\|\phi\|^2_{H_\alpha} := \phi(\infty)^2 + \int_0^\infty |\phi'(x)|^2 e^{\alpha x}\, dx < \infty.$$

Under measurability, local boundedness, and local Lipschitz continuity conditions on (σ_k), one can rewrite (8) as

$$du + Au\, dt = \beta(t,u)\, dt + B(t,u)\, dW(t), \qquad u(0) = u_0, \tag{9}$$

where A is the generator of the semigroup of translations on H_α, W is a cylindrical Wiener process on $U = \ell^2$, and $B : \Omega \times \mathbb{R}_+ \times H \to \mathscr{L}^2(U,H)$ is such that

$$\sum_{k=1}^\infty \int_0^\cdot \sigma_k(s, v(s))\, dw^k(s) = \int_0^\cdot B(s, v(s))\, dW(s).$$

Under such assumptions on (σ_k), (8) admits a unique local mild solution with values in H_α. If (σ_k) satisfy stronger (global) boundedness and Lipschitz continuity assumptions, then the local mild solution is in fact global. For details we refer to [3], as well as to [10].

Positivity of forward rates, i.e. of the mild solution to (8), is established in [10] by proving positivity of mild solutions in weighted L^2 spaces to regularized versions of (8). Such an approximation argument is employed because the conditions on (σ_k) ensuring (local) Lipschitz continuity of the coefficients in the associated stochastic evolution (9) equation in H_α do not imply (local) Lipschitz continuity of the coefficients if state space is changed to a weighted L^2 space.

Thanks to Theorem 2, we can give a much shorter, more direct proof of the (criterion for the) positivity of forward rates. Let $L^2_{-\alpha}$ denote the weighted space $L^2(\mathbb{R}_+, e^{-\alpha x}\, dx)$, and note that H_α is continuously embedded in $L^2_{-\alpha} =: H$. Let us check that assumptions (A1), (A2), and (A3) are satisfied. Assumption (A1) holds true with the choice $\mathcal{O} = \mathbb{R}_+$, endowed with the absolutely continuous measure $m(dx) := e^{-\alpha x}\, dx$. As far as assumption (A2) is concerned, a simple computation shows that $A + \alpha I$ is monotone on $L^2_{-\alpha}$, and, by standard ODE theory, one also verifies that the range of $A + \alpha I + I$ coincides with the whole space $L^2_{-\alpha}$, therefore $A + \alpha I$ is maximal monotone. Even though A itself is not maximal monotone, this is clearly not restrictive, as the "correction" term αI can be incorporated in β without loss of generality. To verify that the resolvent $J_\lambda \in \mathscr{L}(H)$ of $A + \alpha I$ is sub-Markovian, let $y \in H$, so that $J_\lambda y \in \mathsf{D}(A)$ is the unique solution y_λ to the problem

$$y_\lambda - \lambda y'_\lambda + \lambda \alpha y_\lambda = y.$$

If $0 \le y \le 1$ a.e. in $\mathbb{R}_+$, then we have, multiplying both sides by $(y_\lambda - 1)^+$, in the sense of the scalar product of H, that

$$\begin{aligned}(1+\lambda\alpha)\big\langle y_\lambda, (y_\lambda - 1)^+\big\rangle_{-\alpha} &- \lambda\big\langle y'_\lambda, (y_\lambda - 1)^+\big\rangle_{-\alpha}\\ &= \big\langle y, (y_\lambda - 1)^+\big\rangle_{-\alpha} \le \big\langle 1, (y_\lambda - 1)^+\big\rangle_{-\alpha}.\end{aligned}$$

Here and in the following we denote the scalar product and norm of $L^2_{-\alpha}$ simply by $\langle\cdot,\cdot\rangle_{-\alpha}$ and $\|\cdot\|_{-\alpha}$, respectively. Since

$$\big\langle y_\lambda,(y_\lambda-1)^+\big\rangle_{-\alpha}-\big\langle 1,(y_\lambda-1)^+\big\rangle_{-\alpha}=\big\|(y_\lambda-1)^+\big\|^2_{-\alpha}, \tag{10}$$

we obtain

$$\big\|(y_\lambda-1)^+\big\|^2_{-\alpha}-\frac{\lambda}{2}\int_0^\infty\frac{d}{dx}((y_\lambda-1)^+)^2(x)e^{-\alpha x}\,dx+\lambda\alpha\big\langle y_\lambda,(y_\lambda-1)^+\big\rangle_{-\alpha}\leq 0,$$

where, integrating by parts,

$$\begin{aligned}
-\frac{\lambda}{2}&\int_0^\infty\frac{d}{dx}((y_\lambda-1)^+)^2(x)e^{-\alpha x}\,dx\\
&=-\frac{\lambda\alpha}{2}\int_0^\infty((y_\lambda(x)-1)^+)^2e^{-\alpha x}\,dx+\frac{\lambda}{2}((y_\lambda(0)-1)^+)^2\\
&=-\frac{\lambda\alpha}{2}\big\langle y_\lambda,(y_\lambda-1)^+\big\rangle_{-\alpha}+\frac{\lambda\alpha}{2}\big\langle 1,(y_\lambda-1)^+\big\rangle_{-\alpha}+\frac{\lambda}{2}((y_\lambda(0)-1)^+)^2\\
&\geq-\frac{\lambda\alpha}{2}\big\langle y_\lambda,(y_\lambda-1)^+\big\rangle_{-\alpha}.
\end{aligned}$$

Rearranging the terms yields

$$\big\|(y_\lambda-1)^+\big\|^2_{-\alpha}+\frac{\lambda\alpha}{2}\big\langle y_\lambda,(y_\lambda-1)^+\big\rangle_{-\alpha}\leq 0,$$

where the second term on the left-hand side is positive by (10). Therefore $\|(y_\lambda-1)^+\|_{-\alpha}=0$, which implies that $y_\lambda\leq 1$ a.e. in $\mathbb{R}_+$. A completely similar argument, i.e. scalarly multiplying the resolvent equation by y_λ^-, also shows that $y_\lambda\geq 0$ a.e. in $\mathbb{R}_+$, thus completing the proof that J_λ is sub-Markovian. We still need to show that J_λ is contractive in $L^1_{-\alpha}$. Let $y,z\in H$ and $y_\lambda:=J_\lambda y$, $z_\lambda:=J_\lambda z$, so that

$$(y_\lambda-z_\lambda)-\lambda(y_\lambda-z_\lambda)'+\lambda\alpha(y_\lambda-z_\lambda)=y-z. \tag{11}$$

Define the sequences of functions $(\gamma_k),(\hat\gamma_k)\subset\mathbb{R}^\mathbb{R}$ as

$$\gamma_k:r\mapsto\tanh(kr),\qquad\hat\gamma_k:r\mapsto\int_0^r\gamma_k(s)\,ds,$$

and recall that, as $k\to\infty$, γ_k converges pointwise to the sign function, and $\hat\gamma_k$ converges pointwise to the absolute value function. Scalarly multiplying (11) with $\gamma_k(y_\lambda-z_\lambda)$ yields

$$\begin{aligned}
(1+\lambda\alpha)\big\langle y_\lambda-z_\lambda,\gamma_k(y_\lambda-z_\lambda)\big\rangle_{-\alpha}-\lambda\big\langle(y_\lambda-z_\lambda)',\gamma_k(y_\lambda-z_\lambda)\big\rangle_{-\alpha}\\
=\big\langle y-z,\gamma_k(y_\lambda-z_\lambda)\big\rangle_{-\alpha}\leq\big\|y-z\big\|_{L^1_{-\alpha}},
\end{aligned}$$

where, integrating by parts,

$$\begin{aligned}
&\big\langle (y_\lambda - z_\lambda)', \gamma_k(y_\lambda - z_\lambda)\big\rangle_{-\alpha}\\
&\qquad = \int_0^\infty \big(\gamma_k(y_\lambda - z_\lambda)(x)(y_\lambda - z_\lambda)'(x)\big)e^{-\alpha x}\,dx\\
&\qquad = \int_0^\infty \frac{d}{dx}\hat{\gamma}_k(y_\lambda - z_\lambda)(x)e^{-\alpha x}\,dx\\
&\qquad = -\hat{\gamma}_k(y_\lambda(0) - z_\lambda(0)) + \alpha\int_0^\infty \hat{\gamma}_k(y_\lambda - z_\lambda)(x)e^{-\alpha x}\,dx\\
&\qquad \leq \alpha\int_0^\infty \hat{\gamma}_k(y_\lambda - z_\lambda)(x)e^{-\alpha x}\,dx.
\end{aligned}$$

This implies

$$\begin{aligned}
&\big\langle y_\lambda - z_\lambda, \gamma_k(y_\lambda - z_\lambda)\big\rangle_{-\alpha}\\
&\qquad + \lambda\alpha\big\langle y_\lambda - z_\lambda, \gamma_k(y_\lambda - z_\lambda)\big\rangle_{-\alpha} - \lambda\alpha\int_0^\infty \hat{\gamma}_k(y_\lambda - z_\lambda)(x)e^{-\alpha x}\,dx\\
&\qquad \leq \big\|y - z\big\|_{L^1_{-\alpha}}.
\end{aligned}$$

Taking the limit as $k \to \infty$, the sum of the second and third term on the left-hand side converges to zero by the dominated convergence theorem, while the first term on the left-hand side converges to $\|y_\lambda - z_\lambda\|_{L^1_{-\alpha}}$, thus proving that

$$\big\|y_\lambda - z_\lambda\big\|_{L^1_{-\alpha}} \leq \big\|y - z\big\|_{L^1_{-\alpha}},$$

i.e. that the resolvent of $A + \alpha I$ is contractive in $L^1_{-\alpha}$. We have thus shown that assumption (A2) holds for $A + \alpha I$. Moreover, assumption (A3) is satisfied if, for example,

$$|\sigma_k(\omega, t, x, r)|\mathbb{1}_{\{r\leq 0\}} \lesssim r^-.$$

for all $k \in \mathbb{N}$ and $(\omega, t, x) \in \Omega \times \mathbb{R}^2_+$ (see [10], where also slightly more general sufficient conditions are provided). Since all integrability assumptions of Theorem 2 are satisfied, as it follows by inspection of the proof of well-posedness in H_α (see [3, 10, 11]), we conclude that, under the above assumptions on (σ_k), forward rates are positive at all times.

Acknowledgments. Large part of this paper was written while the first-named author was visiting the Interdisziplinäres Zentrum für Komplexe Systeme, Universität Bonn, Germany, whose hospitality is gratefully acknowledged. The second-named author was partially supported by the Vienna Science and Technology Fund (WWTF) through Project MA14-009 and by the Austrian Science Fund (FWF) through Project M 2876.

References

1. V. Barbu, G. Da Prato, and M. Röckner, *Existence and uniqueness of nonnegative solutions to the stochastic porous media equation*, Indiana Univ. Math. J. **57** (2008), no. 1, 187–211. MR 2400255
2. H. Brézis and W. A. Strauss, *Semi-linear second-order elliptic equations in L^1*, J. Math. Soc. Japan **25** (1973), 565–590. MR 0336050 (49 #826)
3. D. Filipović, *Consistency problems for Heath-Jarrow-Morton interest rate models*, Lecture Notes in Mathematics, vol. 1760, Springer-Verlag, Berlin, 2001. MR 1828523 (2002e:91001)
4. M. Haase, *Convexity inequalities for positive operators*, Positivity **11** (2007), no. 1, 57–68. MR 2297322 (2008d:39034)
5. D. C. Heath, R. A. Jarrow, and A. Morton, *Bond pricing and the term structure of interest rates: a new methodology for contingent claims valuation*, Econometrica **60** (1992), no. 1, 77–105.
6. T. Hytönen, J. van Neerven, M. Veraar, and L. Weis, *Analysis in Banach spaces. Vol. II*, Springer, Cham, 2017, Probabilistic methods and operator theory. MR 3752640
7. N. V. Krylov, *Maximum principle for SPDEs and its applications*, Stochastic differential equations: theory and applications, Interdiscip. Math. Sci., vol. 2, World Sci. Publ., Hackensack, NJ, 2007, pp. 311–338. MR 2393582
8. ———, *A relatively short proof of Itô's formula for SPDEs and its applications*, Stoch. Partial Differ. Equ. Anal. Comput. **1** (2013), no. 1, 152–174. MR 3327504
9. M. Kunze and J. van Neerven, *Continuous dependence on the coefficients and global existence for stochastic reaction diffusion equations*, J. Differential Equations **253** (2012), no. 3, 1036–1068. MR 2922662
10. C. Marinelli, *Positivity of mild solution to stochastic evolution equations with an application to forward rates*, arXiv preprint.
11. ———, *Local well-posedness of Musiela's SPDE with Lévy noise*, Math. Finance **20** (2010), no. 3, 341–363. MR 2667893
12. M. Métivier, *Semimartingales*, Walter de Gruyter & Co., Berlin, 1982. MR 688144 (84i:60002)
13. E. Pardoux, *Equations aux derivées partielles stochastiques nonlinéaires monotones*, Ph.D. thesis, Université Paris XI, 1975.
14. M. Röckner and Feng-Yu Wang, *Non-monotone stochastic generalized porous media equations*, J. Differential Equations **245** (2008), no. 12, 3898–3935. MR 2462709
15. L. Scarpa and U. Stefanelli, *An order approach to SPDEs with antimonotone terms*, Stoch. Partial Differ. Equ. Anal. Comput. (2019). DOI:10.1007/s40072-019-00161-7