AFFINE PROCESSES UNDER PARAMETER UNCERTAINTY

TOLULOPE FADINA, ARIEL NEUFELD, AND THORSTEN SCHMIDT

ABSTRACT. We develop a one-dimensional notion of affine processes under parameter uncertainty, which we call *non-linear affine processes*. This is done as follows: given a set Θ of parameters for the process, we construct a corresponding non-linear expectation on the path space of continuous processes. By a general dynamic programming principle we link this non-linear expectation to a variational form of the Kolmogorov equation, where the generator of a single affine process is replaced by the supremum over all corresponding generators of affine processes with parameters in Θ. This non-linear affine process yields a tractable model for Knightian uncertainty, especially for modelling interest rates under ambiguity.

We then develop an appropriate Itô-formula, the respective term-structure equations and study the non-linear versions of the Vasiček and the Cox-Ingersoll-Ross (CIR) model. Thereafter we introduce the non-linear Vasiček-CIR model. This model is particularly suitable for modelling interest rates when one does not want to restrict the state space a priori and hence the approach solves this modelling issue arising with negative interest rates.

Keywords: affine processes, Knightian uncertainty, Riccati equation, Vasiček model, Cox-Ingersoll-Ross model, non-linear Vasiček/CIR model, Heston model, Itô-formula, Kolmogorov equation, fully nonlinear PDE, semimartingale.

1. INTRODUCTION

The modelling of a dynamic and unpredictable phenomenon like stock markets or interest rate markets is often approached via chosing an appropriate stochastic model. In many cases, the choice of the model is a delicate and difficult question. In complex dynamic environments like financial markets it is rather the rule than the exception that unforeseen events lead to difficulties with the a-priori chosen model and improvements of the model have to be developed and implemented.

A promiment example in this direction is the role of affine short-rate models in the last 20 years: while around 2000, the property of the Vasiček model that interest rates can become negative was heavily critizied and the non-negative Cox-Ingersoll-Ross (CIR) model was preferred, the consequences of the financial crises 2007-2008 leading to negative interest rates in the Euro zone rendered the CIR model no longer applicable and led to a renaissance of the Vasiček model.

This example illustrates the important question of *model uncertainty*, which is one of the most important topics in applied sciences and in particular plays a prominent role in finance, not only since the financial crisis. The apparent risk of losses due to model mis-specification, called *model risk* fostered the development of strategies which are *robust* against model risk, typically leading to non-linear pricing rules. These robust strategies play a prominent role in the literature, see Denis and Martini (2006), Cont (2006), Eberlein et al. (2014), Madan (2016), Acciaio et al. (2016), Muhle-Karbe and Nutz (2018), Bielecki et al. (2018) and the book Guyon and Henry-Labordère (2013), to name just a few references in this direction.

A key observation in these works is that the single probability measure used in the classical approaches to specify a model has to be replaced by a family of probability measures (i.e. a full class of models). Such an approach is very natural from the statistical viewpoint: when a model has certain parameters to be estimated, the estimators carry statistical uncertainty

and one considers confidence intervals instead, corresponding to a family of probability measures. The latter formulation of model risk is typically referred to as *parameter uncertainty*, see Avellaneda et al. (1995), Wilmott and Oztukel (1998), Fouque and Ren (2014), and is a major motivation for our research.

Examples in this direction are the notions of g-Brownian motion and G-Brownian motion referring to a Brownian motion with drift or volatility uncertainty, see Peng (1997, 2007a,b) and references therein. Most recently, this theory has been extended to more general approaches, so-called non-linear Lévy processes, see Neufeld and Nutz (2017) and Denk et al. (2017) in this regard.

Here we generalize this notion to affine processes under parameter uncertainty (called non-linear affine processes). While a classical affine process corresponds to a single semi-martingale law, we represent the affine process under parameter uncertainty by a *family* of semimartingale laws whose differential characteristics are bounded from above and below by affine functions of the current states. Nonlinear Lévy processes constitute the special case where the bounds do not depend on the state of the process. It seems important to stretch that for affine processes the bounds on drift and volatility are allowed to depend on the state of the process (in an affine way, however). On the contrary, this state dependence leads to a number of additional difficulties and we therefore restrict ourselves to the simplest case, namely the one-dimensional case without jumps.

It is our aim to provide the appropriate tools for incorporating parameter uncertainty in the prominent class of affine models. This naturally leads to a non-linear version of affine processes and associated non-linear expectations. After having established a dynamic programming principle, we establish the connection to the non-linear Kolmogorov equation. This allows us to study a number of interesting further steps, a non-linear version of the Itô-formula, and non-linear affine term structure equations.

We also provide a number of examples: besides a non-linear variant of the Black-Scholes model and non-linear Vasiček and Cox-Ingersoll-Ross (CIR) models we also introduce a non-linear Vasiček-CIR model. In the latter model one can incorporate negative interest rates in combination with a CIR-like behaviour, solving the problem raised in many practical applications when the state space needed to be restricted to positive interest rates (see Carver (2012)).

The paper is organized as follows: in Section 2 we introduce non-linear affine processes. In Section 3 we prove that a dynamic programming principle holds. Section 4 provides the non-linear Kolmogorov equation and as examples the non-linear Vasiček model and the non-linear CIR model. In Section 5 we provide a non-linear Itô-formula together with some examples and in Section 6 we study (non-linear) affine term structure models. Section 7 studies the application to model risk and Section 8 concludes.

2. Setup

We begin with a short review of continuous affine processes in one dimension. For a detailed exposition we refer to Duffie et al. (2003) and Filipović (2009). Consider the canonical state space, which is either $\mathcal{X} = \mathbb{R}$ or $\mathcal{X} = \mathbb{R}_{>0}$. A (time-homogeneous) Markov process X with values in the state space $\mathcal{X}$ is called *affine* if the conditional characteristic function of X is exponential affine. This means that there exist $\mathbb{C}$-valued functions $\phi(t, u)$ and $\psi(t, u)$, respectively, such that

$$\mathbb{E}[e^{uX_T} \mid X_t] = e^{\phi(T-t,u)+\psi(T-t,u)X_t}$$

for all complex $u \in \{ix : x \in \mathbb{R}\}$, $0 \le t \le T$. The key for our non-linear formulation will be a characterization of X in terms of stochastic differential equations: more precisely, the affine process X is the unique strong solution of the stochastic differential equation

$$dX_t = (b^0 + b^1 X_t)dt + \sqrt{a^0 + a^1 X_t}dW_t, \qquad X_0 = x \qquad (1)$$

where the drift parameter $b^0 + b^1 X_t$ and the diffusion parameter $a^0 + a^1 X_t$ depend on the current value of X in an affine way. Here, the process W is a standard Brownian motion. It should be noted that, depending on the state space, not all parameter combinations are possible, but only those combinations which are admissible in the sense made precise in Theorem 10.2 in Filipović (2009). For our case this implies that if on the one side $\mathcal{X} = \mathbb{R}$ we necessarily have $a^1 = 0$ and $a^0 > 0$ and, on the other side, if $\mathcal{X} = \mathbb{R}_{>0}$ we obtain $a^0 = 0$, $a^1 > 0$ and $b^0 > 0$. In addition, the coefficients ϕ and ψ solve ODEs (classified as Riccati

equations) which is the essence for the high degree of tractability of affine processes in the sense that explicit calculations are possible or efficient numerical methods are obtainable; see Duffie et al. (2003), Filipović (2009) for details and applications in this regard.

2.1. **Non-linear affine processes.** In this section we introduce the necessary tools for defining affine processes under parameter uncertainty. To this end, fix a final time horizon $T > 0$ and let $\Omega = C([0,T])$ be the canonical space of continuous, one-dimensional paths. We endow Ω with the topology of uniform convergence and denote by $\mathscr{F}$ its Borel σ-field. Let X be the canonical process $X_t(\omega) = \omega_t$, and let $\mathbb{F} = (\mathscr{F}_t)_{t \geq 0}$ with $\mathscr{F}_t = \sigma(X_s, 0 \leq s \leq t)$ be the (raw) filtration generated by X.

As we are interested in semimartingale laws on Ω we begin by denoting by $\mathfrak{P}(\Omega)$ the Polish space of all probability measures on Ω equipped with the topology of weak convergence[1]. The process X will be called a (continuous) P-$\mathbb{F}$-semimartingale, for $P \in \mathfrak{P}(\Omega)$, if there exist processes $B = B^P$ and $M = M^P$ such that $X = X_0 + B + M$, where B has continuous paths of (locally) finite variation P-a.s., M is a continuous P-$\mathbb{F}$-local martingale and $B_0 = M_0 = 0$.

It will be important in the following that, by Proposition (2.2) in Neufeld and Nutz (2014), X is a P-$\mathbb{F}$-semimartingale if and only if it is a P-semimartingale with respect to the right-continuous filtration $\mathbb{F}_+ = (\mathscr{F}_{t+})_{t \geq 0}$ or with respect to the usual augmentation $\mathbb{F}_+^P$; here $\mathscr{F}_{t+} = \cap_{s>t} \mathscr{F}_s$. Hence, in the following we can consider semimartingales with respect to the raw filtration $\mathbb{F}$.

The P-$\mathbb{F}$-characteristics of a continuous semimartingale $X = X_0 + B^P + M^P$ in the above representation is the pair (B^P, C) where $C = \langle M^P \rangle$. The non-negative process C does not depend on P, as the quadratic variation is a path property[2]. For the following, we will focus on semimartingales where the semimartingale characteristics are absolutely continuous (a.c.), i.e. there exist predictable processes β^P and $\alpha \geq 0$, such that

$$B^P = \int_0^{\cdot} \beta_s^P ds, \quad C = \int_0^{\cdot} \alpha_s ds. \tag{2}$$

A probability measure $P \in \mathfrak{P}(\Omega)$ is called a *semimartingale law* for X, if X is a P-$\mathbb{F}$-semimartingale. We denote by

$$\mathfrak{P}_{\text{sem}}^{\text{ac}} = \{P \in \mathfrak{P}(\Omega) \mid X \text{ is a } P\text{-}\mathbb{F}\text{-semimartingale with a.c. characteristics}\}$$

the set of all semimartingale laws of X which have absolutely continuous characteristics. In the following we will always denote by (β^P, α) as in (2) the differential characteristics of X under $P \in \mathfrak{P}_{\text{sem}}^{\text{ac}}$.

Our main goal is to allow for a specific version of *model risk* in the sense that there is uncertainty on the parameter vector $\theta = (b^0, b^1, a^0, a^1)$ of the affine process. We assume that there is additional information on bounds on the parameter vector θ and denote these finite bounds by $\underline{b}^i, \bar{b}^i, \underline{a}^i, \bar{a}^i, i = 0, 1$, respectively. This leads to the compact set

$$\Theta = [\underline{b}^0, \bar{b}^0] \times [\underline{b}^1, \bar{b}^1] \times [\underline{a}^0, \bar{a}^0] \times [\underline{a}^1, \bar{a}^1] \subset \mathbb{R}^2 \times \mathbb{R}_{\geq 0}^2. \tag{3}$$

We are interested in the intervals generated by the associated affine functions. In this regard, let $B := [\underline{b}^0, \bar{b}^0] \times [\underline{b}^1, \bar{b}^1]$ and $A := [\underline{a}^0, \bar{a}^0] \times [\underline{a}^1, \bar{a}^1]$. Moreover, we denote for $a \in \mathbb{R}^2$, $a := (a^0, a^1)$ and similarly for $b \in \mathbb{R}^2$. Furthermore, let

$$\begin{aligned} b^*(x) &:= \{b^0 + b^1 x : b \in B\}, \\ a^*(x) &:= \{a^0 + a^1 x^+ : a \in A\} \end{aligned} \tag{4}$$

for $x \in \mathbb{R}$ denote the associated set-valued functions. As the state space will, in general, be $\mathbb{R}$ we have to ensure non-negativity of the quadratic variation which is achieved using $(\cdot)^+ := \max\{\cdot, 0\}$ in the definition of a^*. Due to the nice structure of Θ the sets are always intervals: indeed,

$$\begin{aligned} b^*(x) &= [\underline{b}^0 + (\underline{b}^1 \mathbb{1}_{\{x \geq 0\}} + \bar{b}^1 \mathbb{1}_{\{x < 0\}})x, \bar{b}^0 + (\bar{b}^1 \mathbb{1}_{\{x \geq 0\}} + \underline{b}^1 \mathbb{1}_{\{x < 0\}})x], \\ a^*(x) &= [\underline{a}^0 + \underline{a}^1 x^+, \bar{a}^0 + \bar{a}^1 x^+]. \end{aligned} \tag{5}$$

---

[1]The weak topology is the topology induced by the bounded continuous functions on Ω. Then, $\mathfrak{P}(\Omega)$ is a separable metric space and we denote the associated Borel σ-field by $\mathscr{B}(\mathfrak{P}(\Omega))$.

[2]This is because C can be constructed as a single process not depending on P; that is, two measures under which X has different diffusion are necessarily singular, see Proposition 6.6 in Neufeld and Nutz (2014), for the construction of C.

Clearly, it would be possible to consider more general Θ, which is, however, not our focus here.

Definition 2.1. Let Θ be a set as in (3) with associated a^*, b^* as in (4), let $t \in [0, T]$, and let $P \in \mathfrak{P}_{\text{sem}}^{\text{ac}}$ be a semimartingale law. We call P *affine-dominated on $(t, T]$* by Θ, if (β^P, α) satisfy

$$\beta_s^P \in b^*(X_s), \quad \text{and} \quad \alpha_s \in a^*(X_s), \tag{6}$$

for $dP \otimes dt$-almost all $(\omega, s) \in \Omega \times (t, T]$. If $t = 0$, we call P *affine-dominated* by Θ.

An affine process can uniquely be characterized by its transition probabilities. We will make use of this fact for characterizing affine processes under model uncertainty. Moreover, we denote by $\mathscr{O}$ the considered state space, which will be either $\mathbb{R}$, $\mathbb{R}_{\geq 0}$ or $\mathbb{R}_{>0}$.

Definition 2.2. Let Θ be a set as in (3) with associated a^*, b^* as in (4). A *non-linear affine process* starting at $x \in \mathscr{O}$ is the family of semimartingale laws $P \in \mathfrak{P}_{\text{sem}}^{\text{ac}}$, such that

 (i) $P(X_0 = x) = 1$,
 (ii) P is affine-dominated by Θ.

As explained in the introduction, parameter uncertainty is represented by a family of models replacing the one single model in the approaches without uncertainty: according to Definition 2.2, the affine process under parameter uncertainty is represented by a *family* of semimartingale laws instead of a single one. We denote the semimartingale laws $P \in \mathfrak{P}_{\text{sem}}^{\text{ac}}$, satisfying $P(X_0 = x) = 1$ and being affine-dominated by Θ by $\mathcal{A}(x, \Theta)$. Intuitively, this corresponds to a non-linear affine process starting in x.

It is well-known that the state space $\mathscr{O}$ needs to be chosen in correspondence with the choice of Θ: indeed, the squared Bessel process is an affine process with state space $\mathbb{R}_{>0}$ (see, for example Karatzas and Shreve (1988), Prop. 3.22 of Ch. 3) and the set $\mathcal{A}(x, \Theta)$ will be empty for $x < 0$. To exclude additional difficulties in this direction, we call a family of non-linear affine processes $(\mathcal{A}(x, \Theta))_{x \in \mathscr{O}}$ with state space $\mathscr{O}$ *proper*, if either $a^0 > 0$ holds, or $\underline{a}^0 = \bar{a}^0 = 0$ and $\underline{b}^0 \geq \bar{a}^1/2 > 0$.

It is clear that in the case with $\mathscr{O} = \mathbb{R}$ the assumption $\underline{a}^0 > 0$ is sufficient for reaching the full state space. The case with non-negative state spaces is more delicate. We concentrate on the case $\mathscr{O} = \mathbb{R}_{>0}$. The following proposition gives a sufficient condition in this regard. It moreover shows that the non-linear affine process does not reach zero, in the sense that the event of reaching 0 has zero probability under all $P \in \mathcal{A}(x, \Theta)$.

Proposition 2.3. *Let $x > 0$, and assume that $\underline{a}^0 = \bar{a}^0 = 0$ and $\underline{b}^0 \geq \bar{a}^1/2 > 0$. Then for any $P \in \mathcal{A}(x, \Theta)$ it holds that*

$$P(X_t > 0, \ 0 \leq t \leq T) = 1.$$

Proof. Let $P \in \mathcal{A}(x, \Theta)$, denote by β^P and α the associated processes from Equation (2), and denote by M^P the P-local martingale part of the P-semimartingale X. Moreover, for any $c \geq 0$ define

$$\tau_c := \inf\{t \geq 0 \colon X_t = c\}.$$

We need to show for any time $\mathcal{T} > 0$ that $P[\tau_0 \leq \mathcal{T}] = 0$. To that end, fix an arbitrary $\mathcal{T} > 0$. We adopt the method in Gikhman (2011) to our setting: let ε such that $0 < \varepsilon < x$. Notice that by continuity of the paths of X, we have that $X \geq \varepsilon > 0$ on $[\![0, \tau_\varepsilon]\!]$. Moreover, by assumption,

$$m := \frac{2\underline{b}^0 - \bar{a}^1}{\bar{a}^1} > 0. \tag{7}$$

Itô's formula yields for any $t \geq 0$ that

$$(X_{\tau_\varepsilon \wedge t})^{-m} = (X_0)^{-m} - \int_0^{\tau_\varepsilon \wedge t} m X_s^{-(m+1)} \beta_s^P \, ds + \frac{1}{2} \int_0^{\tau_\varepsilon \wedge t} m(m+1) X_s^{-(m+2)} \alpha_s \, ds$$

$$- \int_0^{\tau_\varepsilon \wedge t} m X_s^{-(m+1)} \, dM_s^P$$

Define the processes M and $\tilde{M}$ by

$$M_t = -\int_0^t m X_s^{-(m+1)} \, dM_s^P, \qquad t \geq 0,$$

$$\tilde{M}_t = \int_0^t (X_{\tau_\varepsilon \wedge s})^{-(m+1)} dM_s^P, \qquad t \geq 0.$$

Clearly, both M and $\tilde{M}$ are local martingales and $M = -m\tilde{M}$ on $[\![0, \tau_\varepsilon]\!]$. We now show that $\tilde{M}$ is a true martingale. By the Burkholder-Davis-Gundy inequality, since $X \geq \varepsilon$ on $[\![0, \tau_\varepsilon]\!]$, we obtain that

$$E\left[\sup_{0 \leq s \leq T} |\tilde{M}_s| \right] \leq CE\left[\left(\int_0^T X_{\tau_\varepsilon \wedge s}^{-2(m+1)} \alpha_{\tau_\varepsilon \wedge s} ds \right)^{1/2} \right]$$

$$\leq CE\left[\left(\bar{a}^1 \int_0^T X_{\tau_\varepsilon \wedge s}^{-2(m+1)+1} ds \right)^{1/2} \right]$$

$$\leq C(\bar{a}^1 \varepsilon^{-2(m+1)+1} T)^{1/2} < \infty$$

and hence $\tilde{M}$ is indeed a martingale. Therefore, M is a true martingale on $[\![0, \tau_\varepsilon]\!]$. Next, since $P \in \mathcal{A}(x, \Theta)$ and $m > 0$, we obtain the estimate

$$X_{\tau_\varepsilon \wedge t}^{-m}$$

$$\leq x^{-m} - \int_0^{\tau_\varepsilon \wedge t} m X_s^{-(m+1)} (\underline{b}^0 + \underline{b}^1 X_s) \, ds + \frac{1}{2} \int_0^{\tau_\varepsilon \wedge t} m(m+1) X_s^{-(m+2)} \bar{a}^1 X_s \, ds + M_{\tau_\varepsilon \wedge t}$$

$$= x^{-m} - m\underline{b}^1 \int_0^{\tau_\varepsilon \wedge t} X_s^{-m} \, ds + \int_0^{\tau_\varepsilon \wedge t} \left(\tfrac{m(m+1)\bar{a}^1}{2} - m\underline{b}^0 \right) X_s^{-(m+1)} \, ds + M_{\tau_\varepsilon \wedge t}.$$

Taking expectations and using that $X \geq \varepsilon > 0$ on $[\![0, \tau_\varepsilon]\!]$ yields, in view of (7), for all $t \geq 0$ that

$$E[(X_{\tau_\varepsilon \wedge t})^{-m}] \leq x^{-m} - m\underline{b}^1 \int_0^{\tau_\varepsilon \wedge t} E[(X_s)^{-m}] \, ds \quad \leq x^{-m} + m|\underline{b}^1| \int_0^{\tau_\varepsilon \wedge t} E[(X_s)^{-m}] \, ds \tag{8}$$

Moreover, as $X \geq \varepsilon > 0$ on $[\![0, \tau_\varepsilon]\!]$, we obtain that

$$m|\underline{b}^1| \int_0^{\tau_\varepsilon \wedge t} E[(X_s)^{-m}] \, ds = m|\underline{b}^1| \int_0^t \mathbb{1}_{\{s \leq \tau_\epsilon\}} E[(X_{\tau_\varepsilon \wedge s})^{-m}] \, ds \leq m|\underline{b}^1| \int_0^t E[(X_{\tau_\varepsilon \wedge s})^{-m}] \, ds.$$

This together with (8) yields for any $t \geq 0$ that

$$E[(X_{\tau_\varepsilon \wedge t})^{-m}] \leq x^{-m} + m|\underline{b}^1| \int_0^t E[(X_{\tau_\varepsilon \wedge s})^{-m}] \, ds.$$

By means of Gronwall's inequality we obtain for all $t \geq 0$ that

$$E[(X_{t \wedge \tau_\varepsilon})^{-m}] \leq x^{-m} e^{m|\underline{b}^1| t}. \tag{9}$$

In particular, we obtain by Tschebyscheffs inequality for our fixed time $\mathcal{T} > 0$ that

$$P(\tau_\varepsilon < \mathcal{T}) = P(X_{\tau_\varepsilon \wedge \mathcal{T}} \leq \varepsilon) = P((X_{\tau_\varepsilon \wedge \mathcal{T}})^{-m} \geq \varepsilon^{-m}) \leq \varepsilon^m E[(X_{\tau_\varepsilon \wedge \mathcal{T}})^{-m}].$$

Inserting (9) and letting ε tend to zero yields the claim. The proof of Proposition 2.3 is thus complete. $\qquad \square$

3. DYNAMIC PROGRAMMING

One of the key insights for Markov processes is the deep link between Markov processes and their expectations to partial differential equations given by the *Kolmogorov* equation. In this section we generalize this relation to the case with parameter uncertainty, i.e. we develop the relation of the non-linear affine process to a non-linear version of the Kolmogorov equation. The path we detail in this section uses dynamic programming and the results obtained in Nutz and van Handel (2013) and El Karoui and Tan (2013a). The key to dynamic programming is a certain stability property under conditioning and pasting. As the non-linear affine processes we considered up to now always start from time $t = 0$, we introduce the appropriate conditional formulations first.

For the remainder of the section, fix Θ as in (3) with associated a^* and b^* as in (4). Denote by $\mathcal{A}(t, x, \Theta)$ the semimartingale laws $P \in \mathfrak{P}_{\text{sem}}^{\text{ac}}$, such that

(i) $P(X_t = x) = 1$, and

(ii) P is affine-dominated on $(t, T]$ by Θ.

The following result yields measurability of the non-linear affine process starting at t in $\omega(t)$.

Lemma 3.1. *The set $\big\{(\omega, t, P) \in \Omega \times [0, T] \times \mathfrak{P}(\Omega) \,\big|\, P \in \mathcal{A}(t, \omega(t), \Theta)\big\}$ is Borel.*

Proof. By Theorem 2.6 in Neufeld and Nutz (2014), the set $\mathfrak{P}^{\mathrm{ac}}_{\mathrm{sem}}$ is Borel, which proves that

$$\big\{(\omega, t, P) \in \Omega \times [0, T] \times \mathfrak{P}^{\mathrm{ac}}_{\mathrm{sem}} \,\big|\, P(X_t = \omega(t)) = 1\big\}$$

is Borel. Moreover, Theorem 2.6 in Neufeld and Nutz (2014), also grants the existence of a Borel-measurable map

$$(P, \widetilde{\omega}, s) \mapsto (\beta^P_s(\widetilde{\omega}), \alpha_s(\widetilde{\omega}))$$

such that (β^P, α) are the differential characteristics of X under P. Therefore, we obtain the Borel measurability of the set

$$G := \bigg\{(\omega, t, P, \widetilde{\omega}, s) \in \Omega \times [0, T] \times \mathfrak{P}^{\mathrm{ac}}_{\mathrm{sem}} \times \Omega \times [0, T] \,\bigg|\, s > t,$$

$$P(X_t = \omega(t)) = 1, (\beta^P_s(\widetilde{\omega}), \alpha_s(\widetilde{\omega})) \in b^*(X_s(\widetilde{\omega})) \times a^*(X_s(\widetilde{\omega}))\bigg\}.$$

Applying Fubini's theorem yields the Borel measurability of the set

$$G' := \bigg\{(\omega, t, P, \widetilde{\omega}) \in \Omega \times [0, T] \times \mathfrak{P}^{\mathrm{ac}}_{\mathrm{sem}} \times \Omega \,\bigg|\, \int_0^T \mathbf{1}_G(\omega, t, P, \widetilde{\omega}, s) \, \mathbf{1}_{(t,T]}(s) \, ds = T - t\bigg\}.$$

Now, observe that

$$\big\{(\omega, t, P) \in \Omega \times [0, T] \times \mathfrak{P}(\Omega) \,\big|\, P \in \mathcal{A}(t, \omega(t), \Theta)\big\}$$
$$= \big\{(\omega, t, P) \in \Omega \times [0, T] \times \mathfrak{P}^{\mathrm{ac}}_{\mathrm{sem}} \,\big|\, E^P[\mathbf{1}_{G'}(\omega, t, P, \cdot)] = 1\big\}.$$

and the right hand side is Borel measurable due to a monotone class argument as in (Neufeld and Nutz, 2014, Lemma 3.1). $\square$

Lemma 3.2. *Fix $(x, t) \in \mathbb{R} \times [0, T]$, $P \in \mathcal{A}(t, x, \Theta)$ and a stopping time τ taking values in $[t, T]$.*

(i) There exists a family of conditional probabilities $(P_\omega)_{\omega \in \Omega}$ of P with respect to $\mathscr{F}_\tau$ such that $P_\omega \in \mathcal{A}(\tau(\omega), \omega_{\tau(\omega)}, \Theta)$ for P-a.e. $\omega \in \Omega$.

(ii) Assume that there exists a family of probability measures $(Q_\omega)_{\omega \in \Omega}$ such that $Q_\omega \in \mathcal{A}(\tau(\omega), \omega_{\tau(\omega)}, \Theta)$ for P-a.e. $\omega \in \Omega$, and the map $\omega \to Q_\omega$ is $\mathscr{F}_\tau$ measurable. Then, the probability measure $P \otimes Q$ defined by

$$P \otimes Q(\cdot) = \int_\Omega Q_\omega(\cdot) \, P(d\omega)$$

is an element of $\mathcal{A}(t, x, \Theta)$.

Proof. This follows directly from (Neufeld and Nutz, 2017, Theorem 2.1), see also (Guo et al., 2017, Lemma 4.6). $\square$

Now, fix a Borel measurable function $\psi : \mathscr{O} \to \mathbb{R}$ and define the value function $v : [0, T] \times \mathscr{O} \to \mathbb{R}$ by

$$v(t, x) := \sup_{P \in \mathcal{A}(t, x, \Theta)} E^P[\psi(X_T)].$$

Using the above results, we obtain the following dynamic programming principle.

Proposition 3.3. *Consider a proper family of non-linear affine processes with state space $\mathscr{O}$. For any $(t, x) \in [0, T] \times \mathscr{O}$ and any stopping time τ taking values in $[t, T]$, we obtain*

$$v(t, x) = \sup_{P \in \mathcal{A}(t, x, \Theta)} E^P\big[v(\tau, X_\tau)\big].$$

Proof. The result follows from Theorem 2.1 in El Karoui and Tan (2013b) (with $\overline{\mathcal{P}}_{t,x} = \mathcal{A}(t, x, \Theta)$ in the notation of El Karoui and Tan (2013b)) by noting that analyticity is implied by measurability as shown in Lemma 3.1 and the required stability assumptions have been shown in Lemma 3.2. $\square$

3.1. **Continuity of the value function.** In the following, we show the continuity of the value function $v(t, x)$. To this end, introduce the constant

$$\mathcal{K} = |\underline{b}^0| + |\underline{b}^1| + |\bar{b}^0| + |\bar{b}^1| + \bar{a}^0 + \bar{a}^1, \tag{10}$$

which is finite by Assumption (3). The following inequality is the cornerstone of the results in this section.

Lemma 3.4. *Consider a proper family of non-linear affine processes with state space $\mathcal{O}$ and let $q \geq 1$. There exists an $0 < \varepsilon \equiv \varepsilon(q) < 1$ such that for all $0 < h \leq \varepsilon$, all $t \in [0, T - h]$ and all $x \in \mathcal{O}$, we have*

$$\sup_{P \in \mathcal{A}(t,x,\Theta)} E^P \left[\sup_{0 \leq s \leq h} |X_{t+s} - x|^q \right] \leq C(h^q + h^{q/2})$$

for some constant $C = C(x, q) > 0$ which may depend on x, but is independent of h and t.

Proof. Let $q \in [1, \infty)$. Consider $P \in \mathcal{A}(t, x, \Theta)$ and denote by $X_s = x + B_s^P + M_s^P$, $s \geq t$, the semimartingale representation of X with predictable finite-variation part B^P and local martingale M^P. In the following we will repeatedly use the elementary inequality that

$$(a_1 + a_2)^q \leq 2^{q-1}(a_1^q + a_2^q) \tag{11}$$

and denote $c_q := 2^{q-1}$.

The Burkholder–Davis–Gundy (BDG) inequality (see Theorem 4.1 in Revuz and Yor (1999)) together with Jensen's inequality and (11) yields for any $h \in [0, T - t]$ that

$$E^P \left[\sup_{0 \leq s \leq h} |X_{t+s} - x|^q \right] \leq c_q E^P \left[\sup_{0 \leq s \leq h} |M_{t+s}^P|^q \right] + c_q E^P \left[\sup_{0 \leq s \leq h} |B_{t+s}^P|^q \right] \tag{12}$$

$$\leq c_q \widetilde{C}_q E^P \left[\left(\int_t^{t+h} \alpha_u \, du \right)^{q/2} \right] + c_q E^P \left[\left(\int_t^{t+h} |\beta_u^P| \, du \right)^q \right].$$

Note that the constant $\widetilde{C}_q \geq 1$ from the BDG inequality does depend on q only.

Let $\widetilde{\mathcal{K}} := 1 + \mathcal{K}$, where $\mathcal{K}$ is the constant defined in (10). Choose any $0 < \varepsilon = \varepsilon(q) < 1$ small enough such that it satisfies

$$1 - c_q^3 \widetilde{C}_q \widetilde{\mathcal{K}}^q (\varepsilon^q + \varepsilon^{q/2}) > 0. \tag{13}$$

Let us verify that such a fixed ε satisfies the desired property: by the very definition of $P \in \mathcal{A}(t, x, \Theta)$, we have on $[t, t + h]$ that both α and $|\beta^P|$ are bounded from above by $\widetilde{\mathcal{K}} + \widetilde{\mathcal{K}} \sup_{0 \leq s \leq h} |X_{t+s}| \geq 1$ since they are affine dominated. This, together with Jensen's inequality yields that

$$E^P \left[\left(\int_t^{t+h} \alpha_u \, du \right)^{q/2} \right] \leq h^{q/2} E^P \left[\left(\widetilde{\mathcal{K}} + \widetilde{\mathcal{K}} \sup_{0 \leq s \leq h} |X_{t+s}| \right)^{q/2} \right]$$

$$\leq h^{q/2} E^P \left[\left(\widetilde{\mathcal{K}} + \widetilde{\mathcal{K}} \sup_{0 \leq s \leq h} |X_{t+s}| \right)^q \right]$$

$$\leq h^{q/2} c_q \left(\widetilde{\mathcal{K}}^q + \widetilde{\mathcal{K}}^q E^P \left[\left(\sup_{0 \leq s \leq h} |X_{t+s}| \right)^q \right] \right)$$

$$\leq h^{q/2} c_q^2 \left(\widetilde{\mathcal{K}}^q + \widetilde{\mathcal{K}}^q |x|^q + \widetilde{\mathcal{K}}^q E^P \left[\left(\sup_{0 \leq s \leq h} |X_{t+s} - x| \right)^q \right] \right).$$

In a similar way we obtain

$$E^P \left[\left(\int_t^{t+h} |\beta_u^P| \, du \right)^q \right] \leq h^q E^P \left[\left(\widetilde{\mathcal{K}} + \widetilde{\mathcal{K}} \sup_{0 \leq s \leq h} |X_{t+s}| \right)^q \right]$$

$$\leq h^q c_q^2 \left(\widetilde{\mathcal{K}}^q + \widetilde{\mathcal{K}}^q |x|^q + \widetilde{\mathcal{K}}^q E^P \left[\left(\sup_{0 \leq s \leq h} |X_{t+s} - x| \right)^q \right] \right).$$

Inserting these inequalities into (12), considering $h \leq \varepsilon$ and noting that $\tilde{C}_q \geq 1$ implies that

$$E^P \Big[\sup_{0 \leq s \leq h} |X_{t+s} - x|^q \Big] \tag{14}$$

$$\leq c_q^3 \, \widetilde{C}_q \widetilde{\mathcal{K}}^q (h^{q/2} + h^q) E^P \Big[\sup_{0 \leq s \leq h} |X_{t+s} - x|^q \Big] + c_q^3 \, \widetilde{C}_q \widetilde{\mathcal{K}}^q (1 + |x|^q)(h^{q/2} + h^q)$$

$$\leq c_q^3 \, \widetilde{C}_q \widetilde{\mathcal{K}}^q (\varepsilon^{q/2} + \varepsilon^q) E^P \Big[\sup_{0 \leq s \leq h} |X_{t+s} - x|^q \Big] + c_q^3 \, \widetilde{C}_q \widetilde{\mathcal{K}}^q (1 + |x|^q)(h^{q/2} + h^q).$$

Since $h \leq \varepsilon$ and we chose $0 < \varepsilon < 1$ such that (13) holds, we obtain for the constant $C := \frac{c_q^3 \, \widetilde{C}_q \widetilde{\mathcal{K}}^q (1 + |x|^q)}{1 - c_q^3 \, \widetilde{C}_q \widetilde{\mathcal{K}}^q (\varepsilon^{q/2} + \varepsilon^q)} > 0$, being independent of t, h, P, that

$$E^P \Big[\sup_{0 \leq s \leq h} |X_{t+s} - x|^q \Big] \leq C \left(h^{q/2} + h^q \right).$$

As $P \in \mathcal{A}(t, x, \Theta)$ was chosen arbitrarily, the claim is proven. $\qquad \square$

Remark 3.5. *The proof of Lemma 3.4, actually shows that for the corresponding $0 < \varepsilon < 1$, we have for all $0 < h \leq \varepsilon$, all $t \in [0, T - h]$ and all $x \in \mathbb{R}$, that the local martingale part $(M_{t+s}^P)_{0 \leq s \leq h}$ restricted on $[t, t+h]$ is a true martingale for any $P \in \mathcal{A}(t, x, \Theta)$.*

Lemma 3.6. *Consider a proper family of non-linear affine processes with state space $\mathscr{O}$ and let $\psi : \mathscr{O} \to \mathbb{R}$ be Lipschitz continuous with Lipschitz-constant L_ψ. Then the value function*

$$v(t, x) := \sup_{P \in \mathcal{A}(t, x, \Theta)} E^P \big[\psi(X_T) \big]$$

is jointly continuous. In particular, $v(t, \cdot)$ is Lipschitz continuous with constant L_ψ and $v(\cdot, x)$ is locally $1/2$-Hölder continuous.

Proof. For the Lipschitz-continuity of $v(t, \cdot)$, observe that for any t

$$|v(t, x) - v(t, y)| \leq \sup_{P \in \mathcal{A}(t, x, \Theta)} E^P \Big[|\psi(X_T) - \psi(y - x + X_T)| \Big] \leq L_\psi |y - x|.$$

For the locally $1/2$-Hölder continuity, let $t \in [0, T)$ and $0 \leq u \leq T - t$ be small enough. Then the dynamic programming principle derived in Proposition 3.3, the Lipschitz continuity of $v(t, \cdot)$ and Lemma 3.4 yield

$$|v(t, x) - v(t + u, x)| = \Big| \sup_{P \in \mathcal{A}(t, x, \Theta)} E^P \big[v(t + u, X_{t+u}) - v(t + u, x) \big] \Big|$$

$$\leq L_\psi \sup_{P \in \mathcal{A}(t, x, \Theta)} E^P \big[|X_{t+u} - x| \big]$$

$$\leq L_\psi C(x, 1) \left(u^{1/2} + u \right)$$

with constant $C(x, 1)$ from Lemma 3.4. Finally, consider a sequence (t_n, x_n) converging to (t, x). Then

$$|v(t_n, x_n) - v(t, x)| \leq |v(t_n, x_n) - v(t_n, x)| + |v(t_n, x) - v(t, x)|$$

$$\leq L_\psi |x_n - x| + L_\psi C(x, 1) \big(|t_n - t|^{1/2} + |t_n - t| \big).$$

Letting n go to infinity yields the result. $\qquad \square$

4. THE KOLMOGOROV EQUATION

In this section we provide the link between the non-linear affine process and the associated non-linear Kolmogorov equation. More precisely, we relate the non-linear affine process to a (fully) non-linear partial differential equation (PDE). This is achieved by a probabilistic construction involving an optimal control problem on the canonical space of continuous paths where the controls are laws of affine-dominated semimartingales. To this end, note that the affine process X given in Equation (1) is uniquely characterized by its infinitesimal generator,

$$\mathscr{L}^\theta f(x) = (b^0 + b^1 x) \partial_x f(x) + \frac{1}{2}(a^0 + a^1 x^+) \partial_{xx} f(x). \tag{15}$$

This is equivalent to $f(X_t) - f(x) - \int_0^t \mathscr{L}^\theta f(X_s) ds$, $0 \leq t \leq T$ solving the martingale problem, see, e.g., Theorem 21.7 in Kallenberg (2002).

Consider the state space $\mathscr{O}$ which will be either $\mathbb{R}$, $\mathbb{R}_{\geq 0}$ or $\mathbb{R}_{>0}$. Fix $\psi : \mathscr{O} \to \mathbb{R}$ and consider the fully non-linear PDE

$$\begin{cases} -\partial_t v(t,x) - G\big(x, \partial_x v(t,x), \partial_{xx} v(t,x)\big) = 0 & \text{on } [0,T) \times \mathscr{O}, \\ \qquad\qquad\qquad\qquad v(T,x) = \psi(x) & x \in \mathscr{O}, \end{cases} \tag{16}$$

where $G : \mathscr{O} \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is defined by

$$G(x,p,q) := \sup_{(b^0,b^1,a^0,a^1)\in\Theta} \left\{ (b^0 + b^1 x)p + \frac{1}{2}(a^0 + a^1 x^+)q \right\}. \tag{17}$$

The function $-G$ satisfies the degenerate ellipticity condition and as Θ is compact, it is also continuous. Observe that the PDE defined in (16) can be seen as non-linear affine PDE, since for $\theta := (b^0, b^1, a^0, a^1)$ it is of the form

$$\begin{cases} -\partial_t v(t,x) - \sup_{\theta\in\Theta} \mathscr{L}^\theta v(t,x) = 0 & \text{on } [0,T) \times \mathscr{O}, \\ \qquad\qquad\qquad v(T,x) = \psi(x) & x \in \mathscr{O}. \end{cases}$$

We write $C_b^{2,3}([0,T) \times \mathscr{O})$ for the set of functions on $[0,T) \times \mathscr{O}$ having bounded continuous derivatives up to the second and third order in t and x, respectively. An upper semicontinuous function u on $[0,T) \times \mathscr{O}$ will be called a viscosity subsolution of (16) if $u(T,\cdot) \leq \psi(\cdot)$ and

$$-\partial_t \varphi(t,x) - G\big(x, D_x \varphi(t,x), D_{xx}^2 \varphi(t,x)\big) \leq 0$$

whenever $\varphi \in C_b^{2,3}([0,T) \times \mathscr{O})$ is such that $\varphi \geq u$ on $[0,T) \times \mathscr{O}$ and $\varphi(t,x) = u(t,x)$. The definition of a viscosity supersolution is obtained by reversing the inequalities and the semicontinuity. Finally, a continuous function is a viscosity solution if it is both sub- and supersolution.

We obtain a stochastic representation for the non-linear affine PDE (16).

Theorem 4.1. *Consider a proper family of non-linear affine processes with state space $\mathscr{O}$ and let $\psi : \mathscr{O} \to \mathbb{R}$ be Lipschitz continuous. Then*

$$v(t,x) := \sup_{P\in\mathcal{A}(t,x,\Theta)} E^P\big[\psi(X_T)\big], \quad x \in \mathscr{O}$$

is a viscosity solution of the non-linear PDE in (16).

Proof. The proof essentially follows the well-known standard arguments in stochastic control, see e.g. the proof of (Neufeld and Nutz, 2017, Proposition 5.4).

By Lemma 3.6, $v(t,x)$ is continuous on $[0,T) \times \mathbb{R}$, and we have $v(T,x) = \psi(x)$ by the definition of v. We show that v is a viscosity subsolution of the non-linear affine PDE defined in (16); the supersolution property is proved similarly. We remark that in the subsequent lines within this proof, $C > 0$ is a constant whose values may change from line to line.

Let $(t,x) \in [0,T) \times \mathbb{R}$ and let $\varphi \in C_b^{2,3}([0,T) \times \mathbb{R}^d)$ be such that $\varphi \geq v$ and $\varphi(t,x) = v(t,x)$. By the dynamic programming principle obtained in Proposition 3.3 we have for any $0 < u < T - t$ that

$$\begin{aligned} 0 &= \sup_{P\in\mathcal{A}(t,x,\Theta)} E^P\big[v(t+u, X_{t+u}) - v(t,x)\big] \\ &\leq \sup_{P\in\mathcal{A}(t,x,\Theta)} E^P\big[\varphi(t+u, X_{t+u}) - \varphi(t,x)\big]. \end{aligned} \tag{18}$$

Fix any $P \in \mathcal{A}(t,x,\Theta)$, denote as above by (β^P, α) the differential characteristics of the continuous semimartingale X under P, and denote by M^P the P-local martingale part of the P-semimartingale X. Then, Itô's formula yields

$$\begin{aligned} \varphi(t+u, X_{t+u}) - \varphi(t,x) &= \int_0^u \partial_t \varphi(t+s, X_{t+s})\, ds + \int_0^u \partial_x \varphi(t+s, X_{t+s})\, dM_{t+s}^P \\ &\quad + \int_0^u \partial_x \varphi(t+s, X_{t+s}) \beta_{t+s}^P\, ds + \frac{1}{2} \int_0^u \partial_{xx} \varphi(t+s, X_{t+s}) \alpha_{t+s}\, ds. \end{aligned} \tag{19}$$

As $\varphi \in C_b^{2,3}([0,T) \times \mathbb{R})$, $\partial_x \varphi$ is uniformly bounded, thus by Remark 3.5, we see that for small enough $0 < u < T - t$ the local martingale part in (19) is in fact a true martingale, starting

at 0. In particular, its expectation vanishes. The next step is to estimate the expectation of the other terms. In this regard, note that

$$E^P\left[\int_0^u \partial_x\varphi(t+s,X_{t+s})\beta_{t+s}^P\,ds\right]$$
$$\leq \int_0^u E^P\left[\left|\partial_x\varphi(t+s,X_{t+s})-\partial_x\varphi(t,x)\right||\beta_{t+s}^P|+\partial_x\varphi(t,x)\beta_{t+s}^P\right]ds. \qquad (20)$$

Since $\varphi \in C_b^{2,3}$, $\partial_x\varphi$ is Lipschitz. Hence we obtain with the constant $\mathcal{K}$ from Equation (10) together with Lemma 3.4 that for small enough u,

$$\int_0^u E^P\left[\left|\partial_x\varphi(t+s,X_{t+s})-\partial_x\varphi(t,x)\right|\cdot|\beta_{t+s}^P|\right]ds$$
$$\leq C\int_0^u E^P\left[\left(s+\sup_{0\leq v\leq u}|X_{t+v}-x|\right)\cdot|\beta_{t+s}^P|\right]ds$$
$$\leq C\int_0^u E^P\left[\left(s+\sup_{0\leq v\leq u}|X_{t+v}-x|\right)\left(\mathcal{K}+\mathcal{K}\sup_{0\leq v\leq u}|X_{t+v}|\right)\right]ds$$
$$\leq C\int_0^u E^P\left[\left(s+\sup_{0\leq v\leq u}|X_{t+v}-x|\right)\left(\mathcal{K}+\mathcal{K}|x|+\mathcal{K}\sup_{0\leq v\leq u}|X_{t+v}-x|\right)\right]ds$$
$$\leq C\left(u^3+u^{5/2}+u^2+u^{3/2}\right). \qquad (21)$$

Inserting (21) into (20) yields

$$E^P\left[\int_0^u \partial_x\varphi(t+s,X_{t+s})\beta_{t+s}^P\,ds\right]$$
$$\leq \int_0^u E^P\left[\partial_x\varphi(t,x)\,\beta_{t+s}^P\right]ds+C\left(u^3+u^{5/2}+u^2+u^{3/2}\right). \qquad (22)$$

The same argument applied to $\partial_{xx}\varphi$ leads to

$$E^P\left[\int_0^u \partial_{xx}\varphi(t+s,X_{t+s})\,\alpha_{t+s}\,ds\right]$$
$$\leq \int_0^u E^P\left[\partial_{xx}\varphi(t,x)\,\alpha_{t+s}\right]ds+C\left(u^3+u^{5/2}+u^2+u^{3/2}\right). \qquad (23)$$

Moreover, by a similar calculation, we have

$$E^P\left[\int_0^u \partial_t\varphi(t+s,X_{t+s})\,ds\right]$$
$$\leq \int_0^u \partial_t\varphi(t,x)\,ds+\int_0^u E^P\left[\left|\partial_t\varphi(t+s,X_{t+s})-\partial_t\varphi(t,x)\right|\right]ds$$
$$\leq \int_0^u \partial_t\varphi(t,x)\,ds+C\int_0^u E^P\left[s+\sup_{0\leq v\leq u}|X_{t+v}-x|\right]ds$$
$$\leq \int_0^u \partial_t\varphi(t,x)\,ds+C\left(u^2+u^{3/2}\right). \qquad (24)$$

As above we write $\theta:=(b^0,b^1,a^0,a^1)$ for an element in Θ. Then, by taking expectations in (19) and using (20)–(24) yields

$$E^P\left[\varphi(t+u,X_{t+u})-\varphi(t,x)\right]\leq C\left(u^3+u^{5/2}+u^2+u^{3/2}\right)$$
$$+\int_0^u\left(\partial_t\varphi(t,x)+E^P\left[\partial_x\varphi(t,x)\,\beta_{t+s}^P+\partial_{xx}\varphi(t,x)\,\alpha_{t+s}\right]\right)ds$$
$$\leq C\left(u^3+u^{5/2}+u^2+u^{3/2}\right)+u\partial_t\varphi(t,x)$$
$$+\int_0^u E^P\left[\sup_{\theta\in\Theta}\left\{(b^0+b^1X_{t+s})\,\partial_x\varphi(t,x)+\frac{1}{2}(a^0+a^1X_{t+s}^+)\,\partial_{xx}\varphi(t,x)\right\}\right]. \qquad (25)$$

Here the supremum turns out to be $G(X_{t+s},\partial_x\varphi(t,x),\partial_{xx}\varphi(t,x))$. Note that by the very definition of G,

$$G(X_{t+s},p,q)\leq G(x,p,q)+\sup_{\theta\in\Theta}\left\{|b^1|\,|X_{t+s}-x|\,|p|+|a^1|\,|X_{t+s}-x|\,|q|\right\}.$$

Therefore, by using that $\varphi \in C_b^{2,3}$, the definition of the constant $\mathcal{K}$ in (10) and Lemma 3.4, we have

$$
\int_0^u E^P\Big[G(X_{t+s}, \partial_x \varphi(t,x), \partial_{xx}\varphi(t,x))\}\Big]\,ds
$$
$$
\leq uG(x, \partial_x \varphi(t,x), \partial_{xx}\varphi(t,x)) + uC\mathcal{K}E^P\big[|X_{t+s} - x|\big]
$$
$$
\leq uG(x, \partial_x \varphi(t,x), \partial_{xx}\varphi(t,x)) + C\mathcal{K}\big(u^2 + u^{3/2}\big). \tag{26}
$$

Combining (25)–(26) yields

$$
E^P\Big[\varphi(t+u, X_{t+u}) - \varphi(t,x)\Big] \leq u\partial_t \varphi(t,x) + uG(x, \partial_x \varphi(t,x), \partial_{xx}\varphi(t,x))
$$
$$
+ C\big(u^3 + u^{5/2} + u^2 + u^{3/2}\big). \tag{27}
$$

for some constant $C > 0$ which is independent of P. As the choice of $P \in \mathcal{A}(t,x,\Theta)$ was arbitrary, we deduce from (18) that

$$
0 \leq \sup_{P \in \mathcal{A}(t,x,\Theta)} E^P\big[\varphi(t+u, X_{t+u}) - \varphi(t,x)\big]
$$
$$
\leq u\partial_t \varphi(t,x) + uG(x, \partial_x \varphi(t,x), \partial_{xx}\varphi(t,x)) + C\big(u^3 + u^{5/2} + u^2 + u^{3/2}\big). \tag{28}
$$

By dividing first in (28) by $-u$ and then let u go to zero, we obtain that

$$
-\partial_t \varphi(t,x) - G(x, \partial_x \varphi(t,x), \partial_{xx}\varphi(t,x)) \leq 0,
$$

which proves that v is indeed a viscosity subsolution as desired. $\qquad\square$

4.1. **Uniqueness.** Uniqueness in our framework is not covered by standard arguments as in Fleming and Soner (2006) or Crandall et al. (1992), since the diffusion coefficient does not satisfy a global Lipschitz condition. This is a well-known difficulty already discussed in Feller (1951). To overcome this, we have to distinguish the two cases where the state space is either $\mathbb{R}$ or $\mathbb{R}_{>0}$.

We begin with the general case which covers the non-linear Vasiček-CIR model discussed in detail in Section 6.1. In this case we do not decide a priori whether a Vasiček or a Cox-Ingersoll-Ross (CIR) model takes place and therefore consider the full space $\mathbb{R}$ as state space. To achieve this, we assume throughout a non-vanishing volatility, i.e. $\underline{a}^0 > 0$. When the process reaches zero from above, this avoids that a deterministic behaviour on $\mathbb{R}_{<0}$ takes place and that singularities arise. On the other side, we do not need any assumptions on a^1.

Proposition 4.2. *Assume that $\underline{a}^0 > 0$ and let $\mathcal{O} = \mathbb{R}$. Then, for any given continuous and bounded function $\psi : \mathbb{R} \to \mathbb{R}$, the non-linear PDE introduced in (16)–(17) admits at most one viscosity solution $v(t,x)$ on $[0,T] \times \mathbb{R}$ satisfying the terminal conditions*

$$
v(T,x) = \psi(x), \quad x \in \mathbb{R}.
$$

Proof. Uniqueness in this case follows by the observation that the coefficients are Lipschitz once $\underline{a}^0 > 0$. Indeed, then following the standard procedure detailed in Section V.9 in Fleming and Soner (2006) allows to extend the uniqueness results from Corollary V.8.1 therein to an unbounded domain as considered here. $\qquad\square$

On the other side if we consider the case where $\mathbb{R}_{>0}$ is the state space, we will necessarily require $\underline{a}^0 = \bar{a}^0 = 0$. Then, the (bounds on the) diffusion coefficient do no longer satisfy a global Lipschitz property and the standard methodology can not be applied. To the best of our knowledge, only Costantini et al. (2012) and Amadori (2007) treat this setting while we will apply here the techniques from the second article.

In this regard, let

$$
h_\epsilon(x) = \begin{cases} \max\left(\frac{1}{\log(x)}, 1\right) & \epsilon = 0 \\ \max\left(\frac{1}{x^\epsilon}, 1\right) & \text{otherwise.} \end{cases} \tag{29}
$$

Proposition 4.3. *Assume that $\underline{a}^0 = \bar{a}^0 = 0$, $\underline{b}^0 \geq \bar{a}^1/2 > 0$, let $\mathcal{O} = \mathbb{R}_{>0}$, and consider a Lipschitz-continuous $\psi : \mathcal{O} \to \mathbb{R}$.*

(i) Then, (16)–(17) has a one and only one viscosity solution v on $[0,T] \times \mathcal{O}$ satisfying

$$
\sup_{(t,x) \in [0,T] \times \mathcal{O}} \frac{|v(t,x)|}{1+x} < \infty. \tag{30}
$$

(ii) Let $\epsilon = 2\underline{b}^0/\bar{a}^1 - 1 > 0$ and suppose that there is $\rho, \rho' > 0$ and $\epsilon' \in [0, \epsilon)$ such that $\frac{|\psi(x)|}{1+|x|}$ is bounded on $(0, \infty)$ and

$$x \mapsto \frac{|\psi(x)|}{h_{\epsilon'}(x)} \tag{31}$$

is bounded from above respectively in the set (ρ, ∞) and $(0, \rho')$. Then (16)–(17) has one and only one viscosity solution $v(t, x)$ on $[0, T] \times \mathcal{O}$ satisfying

$$\limsup_{x \to \infty} \sup_{t \in [0,T]} \frac{|v(t,x)|}{x^2} = 0, \quad \lim_{x \to 0} \sup_{t \in [0,T]} \frac{|v(t,x)|}{h_\epsilon(x)} = 0. \tag{32}$$

Proof. Positivity follows readily from Proposition 2.3. This will allow to treat the non-linear PDE in (16)–(17) on the state space $\mathcal{O} = \mathbb{R}_{>0}$. The key for uniqueness is that on compact subsets of $\mathbb{R}_{>0}$ the square-root *is* Lipschitz continuous.

To begin with, note that existence of a solution under the Lipschitz property of ψ follows from Theorem 4.1. Furthermore, Theorem 4 in Amadori (2007) yields the desired uniqueness in our case. We recall this result together with its assumptions in Appendix B. The claim the follows from Theorem B.1. $\square$

Remark 4.4. *Already the results in the case without uncertainty in Costantini et al. (2012) show that these results do not generalize to $\mathbb{R}_{\geq 0}$, because then the Lipschitz property on compact subsets which is crucially used in the proof will no longer be satisfied. Moreover, Example 1 in Amadori (2007) shows that the condition $\underline{b}^0 \geq \bar{a}^1/2 > 0$ is indeed necessary for uniqueness.*

It is time to accommodate the established results with some first motivating examples. We will begin with classical affine processes under parameter uncertainty, leading to non-linear affine processes. Thereafter we show how to extend beyond this case and introduce classes of non-linear affine processes which do not have a classical counterpart.

4.2. **The non-linear Vasiček model.** The first example of an affine model is the so-called Vasiček model, see for example Filipović (2009). It is a Gaussian Ornstein-Uhlenbeck process and is obtained as the strong solution of the SDE in Equation (1) by considering $a^1 = 0$. Introducing parameter uncertainty we arrive at the non-linear affine process with $[\underline{a}^1, \bar{a}^1] = \{0\}$. We call this case the *non-linear Vasiček model*.

While in the case with no parameter uncertainty, this model can be characterized efficiently by its Fourier transforms and the associated Riccati equation, this is no longer possible here (except for the special case where $\underline{b}^1 = \bar{b}^1$, see Remark 4.7 below) and one has to rely on numerical techniques. In one dimension, this does not at all pose a problem, see for example Heider (2010). To illustrate this, we solve the equation for the simplest pay-off, $f(x) = e^x$. The result is shown in Figure 1.

This example combines and encompasses the following two well-known non-linear processes: g-Brownian motion and G-Brownian motion, see Peng (1997) and Peng (2007a), for example, and Neufeld and Nutz (2017) for the case with jumps. In the following we will also elaborate on the case where explicit solutions can be obtained, see Proposition 4.6 and Remark 4.7.

4.3. **The non-linear CIR model.** While the Gaussianity of the Vasiček model immediately implies that the process becomes negative with positive probability, this is inappropriate for various applications, e.g. in credit risk. There, the considered affine process models an intensity, which by definition has to be non-negative. Also positive interest rates were a must of a model before the recent crises in 2008-2010. The Cox-Ingersoll-Ross (CIR) model serves as an affine model with state space $\mathbb{R}_{>0}$ and therefore satisfies these needs. It is obtained by choosing $a^0 = 0$ and $a^1 > 0$ in the SDE in (1).

The CIR model under parameter uncertainty, which we call the non-linear CIR model, is obtained by considering the state space $\mathcal{O} = \mathbb{R}_{>0}$ and assuming that $\underline{a}^0 = \bar{a}^0 = 0$, $\underline{b}^0 \geq \bar{a}^1/2 > 0$. It is remarkable how far-reaching the positivity of X will be: indeed, a first look at the non-linear Kolmogorov equation already reveals that for increasing and convex functions the supremum in Equation (17) will be attained by the upper bounds of the coefficients.

The Fourier-transform which is lacking monotonicity (as e^{iux} does) turns out to loose tractability in the non-linear case considered here. The Laplace transform does not suffer

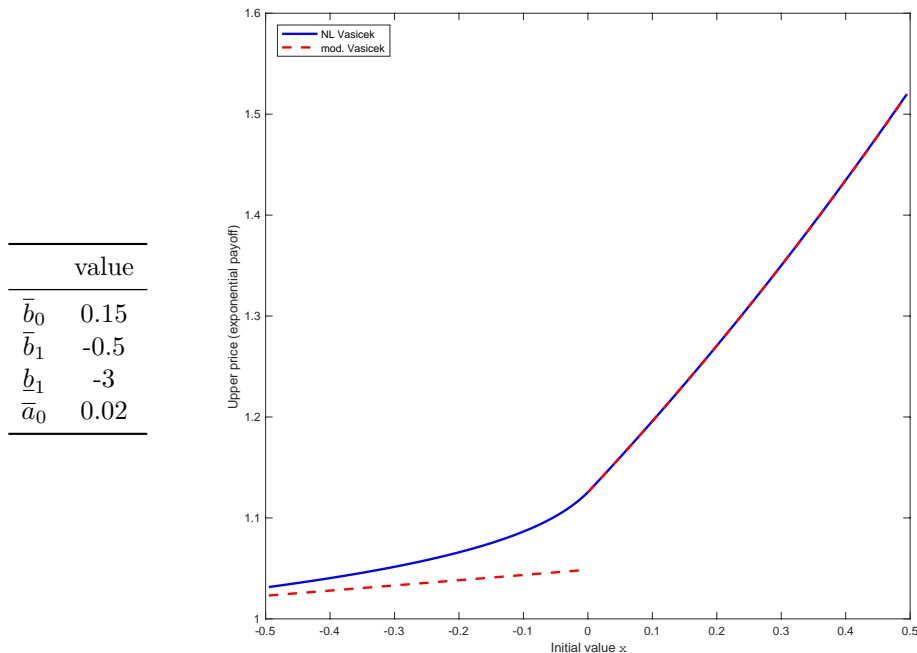|	value
$\bar{b}_0$	0.15
$\bar{b}_1$	-0.5
$\underline{b}_1$	-3
$\bar{a}_0$	0.02

FIGURE 1. This figure shows the solution of the non-linear Kolmogorov equation for the non-linear Vasiček-model with boundary condition $f(x) = e^x$. The dashed line shows the solutions for the Vasiček model with parameter-set $(\bar{b}_0, \underline{b}_1, \bar{a}_0)$ (on $\mathbb{R}_{<0}$) and $(\bar{b}_0, \bar{b}_1, \bar{a}_0)$ (on $\mathbb{R}_{\geq 0}$) and illustrate the non-linearity in the solution due to the parameter uncertainty. Note that on the positive halfline the curves overlap in the figure.

from this problem and we show in the following that in important special cases it can be computed explicitly. However, inversion techniques can no longer be applied in the non-linear setting and the Laplace transform merely serves as a prime example of an increasing convex function for which the non-linear expectations can be computed explicitly in special cases (and for decreasing concave functions of course).

Remark 4.5. *In principle, the non-linear CIR model can be extended to the whole space, i.e. choosing $\mathscr{O} = \mathbb{R}$ is possible, since in Equation* (4), *negative values pose no problem. In the negative halfline the dynamics of the model have no diffusive part and only moves through the drift (which of course can still be stochastic). This could happen with a positive (upper) mean-reversion level and by starting from a negative value.*

We begin by a general result for affine processes which classifies when the classical representation via Riccati equations still holds.

Recall from (5), that $b^*(x)$ and $a^*(x)$ are actually intervals. Define the upper bound of $a^*(x)$ by

$$\bar{a}(x) := \bar{a}^0 + \bar{a}^1 x^+. \tag{33}$$

Note that the upper bound of $b^*(x)$ is given by

$$\bar{b}(x) := \bar{b}^0 + \underline{b}^1 x \mathbb{1}_{\{x<0\}} + \bar{b}^1 x \mathbb{1}_{\{x \geq 0\}}. \tag{34}$$

Letting $\bar{B}^{1,x} = \underline{b}^1 \mathbb{1}_{\{x<0\}} + \bar{b}^1 \mathbb{1}_{\{x \geq 0\}}$, we obtain that $\bar{b}(x) = \bar{b}^0 + \bar{B}^{1,x} x$. Note that for $x \neq y$ we may no longer have that $\bar{b}(x) = \bar{b}^0 + \bar{B}^{1,y} x$.

Proposition 4.6. *Consider a non-linear affine process $\mathcal{A}(x, \Theta)$ and assume that for all $P \in \mathcal{A}(x, \Theta)$*

$$\beta_t^P \leq \bar{b}^0 + \bar{B}^{1,x} X_t, \tag{35}$$

$dP \otimes dt$-almost everywhere for $0 \leq t \leq T$. Moreover, assume either that $\underline{a}_1 = \bar{a}_1 = 0$ or that for all $P \in \mathcal{A}(x, \Theta)$, $X_t \geq 0$ $P \otimes dt$-a.e. If there exists a $\bar{P} \in \mathcal{A}(x, \Theta)$ and a $\bar{P}$-$\mathbb{F}$-Brownian

motion W such that the canonical process under $\bar{P}$ is the unique strong solution of

$$dX_t = (\bar{b}^0 + \bar{B}^{1,x}X_t)dt + \sqrt{\bar{a}(X_t)}dW_t, \qquad X_0 = x, \tag{36}$$

then, for all $u \geq 0$ and $0 \leq t \leq T$,

$$\sup_{P \in \mathcal{A}(x,\Theta)} E^P\big[e^{uX_t}\big] = \exp\big(\phi(t,u) + \psi(t,u)x\big),$$

where ϕ, ψ solve the Riccati equations

$$\partial_t \phi(t,u) = \frac{1}{2}\bar{a}^0 \psi(t,u)^2 + \bar{b}^0 \psi(t,u), \qquad \phi(0,u) = 0 \tag{37}$$

$$\partial_t \psi(t,u) = \frac{1}{2}\bar{a}^1 \psi(t,u)^2 + \bar{B}^{1,x}\psi(t,u), \qquad \psi(0,u) = u. \tag{38}$$

This result is obtained by showing that the supremum in the non-linear expectation is obtained by the maximal semimartingale law $\bar{P}$ which corresponds to an affine process with parameters $\bar{a}^0, \bar{a}^1, \bar{b}^0, \bar{B}^{1,x}$. Inspection of the proof shows that this property also holds when e^{ux} is replaced by any other increasing and convex function (a Call payoff, for example). An analogous formulation for $u < 0$ of course also holds. Furthermore, it is interesting to see that the Riccati equations in Equations (37) and (38) can be replaced by respective versions thereof.

Note that the assumption $X_t \geq 0$ $P \otimes dt$-a.e. is implied by assuming $x > 0$ together with $\underline{a}^0 = \bar{a}^0 = 0$ and $\underline{b}^0 \geq \bar{a}^1/2 > 0$ due to Proposition 2.3.

Proof. The claim follows by an application of Theorem 2.2 in Bergenthum and Rüschendorf (2007). This needs validity of the so-called propagation of order (PO) property and we give a detailed account of this in Appendix A. Proposition A.4 in particular yields that the PO property is satisfied for increasing and convex (decreasing and concave) functions when compared to an affine process.

Let $P \in \mathcal{A}(x,\Theta)$. By assumption there exists a $\bar{P} \in \mathcal{A}(x,\Theta)$ and a $\bar{P}$-$\mathbb{F}$-Brownian motion W, such that X is the unique strong solution of the stochastic differential equation (36) under $\bar{P}$. Since e^{ux} for $u \geq 0$ is increasing and convex and (35) holds, Theorem 2.2 in Bergenthum and Rüschendorf (2007) yields that

$$E^P\big[e^{uX_t}\big] \leq E^{\bar{P}}\big[e^{uX_t}\big].$$

Since $P \in \mathcal{A}(x,\Theta)$ was arbitrary and, under $\bar{P}$, X is an affine process with parameters $\bar{a}^0, \bar{a}^1, \bar{b}^0, \bar{B}^{1,x}$, the affine representation follows and the Riccati equations in (37) can be obtained directly from Theorem 10.1 in Filipović (2009). $\qquad\square$

Remark 4.7. *It can be easily checked that the conditions for Proposition 4.6 hold in the following two cases:*

(1) The non-linear Vasiček model with state space $\mathscr{O} = \mathbb{R}$ and $\bar{b}^1 = \underline{b}^1$,

(2) the non-linear CIR model on the state space $\mathscr{O} = \mathbb{R}_{>0}$.

By the above result we can use the classical Fourier-inversion technique for these affine processes when pricing increasing and convex payoffs (like Call options) or decreasing and concave ones, see Section 10.3 in Filipović (2009) for examples and details in this direction. In Example 5.6 we will sketch an application in a Heston model with parameter uncertainty in the stochastic volatility.

5. AN ITÔ-FORMULA FOR NON-LINEAR AFFINE PROCESSES

In this section we will construct new processes from non-linear affine processes by simple transformations. The main tool for this will be a suitable formulation of the Itô-formula in our setting.

Consider a twice continuously differentiable function $F \in C^2(\mathbb{R})$. If we start from a non-linear affine process $\mathcal{A}(x,\Theta)$ and consider $\tilde{X} := F(X)$ then for any $P \in \mathcal{A}(x,\Theta)$ the process $\tilde{X}$ is a P-semimartingale and we denote its (differential) semimartingale characteristics by $\tilde{\alpha}$ and $\tilde{\beta}^P$ (starting from α and β^P from Equality (2)). In this section we answer the question if the non-linear process $\tilde{X}$ itself, i.e. the associated semimartingale laws can be studied independently of X. This corresponds one-to-one to the question if there exist an independent formulation of the non-linear process $\tilde{X}$. The following proposition gives a positive answer to this question.

We define the interval-valued functions a^F and b^F by

$$a^F(x) := [(F'(x))^2(\underline{a}^0 + \underline{a}^1 x^+), (F'(x))^2(\bar{a}^0 + \bar{a}^1 x^+)] \tag{39}$$

and

$$b^F(x) := \left[\inf_{(\beta,\alpha)\in b^*(x)\times a^*(x)} \left(F'(x)\beta + \frac{1}{2}F''(x)\alpha \right), \sup_{(\beta,\alpha)\in b^*(x)\times a^*(x)} \left(F'(x)\beta + \frac{1}{2}F''(x)\alpha \right) \right]. \tag{40}$$

The non-linear process $\tilde{X}$ inherits certain bounds from X which is characterized in the following proposition.

Proposition 5.1. *Let $\mathcal{A}(x,\Theta)$ be a non-linear affine process and $F \in C^2$. Then, for every $P \in \mathcal{A}(x,\Theta)$, $\tilde{X} = F(X)$ is a P-semimartingale with differential characteristics $\tilde{\alpha}$ and $\tilde{\beta}^P$ satisfying*

$$\tilde{\alpha}_s \in a^F(X_s), \tag{41}$$

$$\tilde{\beta}_s^P \in b^F(X_s). \tag{42}$$

Proof. Let $t \in [0, T]$. By definition, $P \in \mathcal{A}(x, \Theta)$ implies that

$$X_t = X_0 + \int_0^t \beta_s^P ds + M_t^P$$

where $\beta_s^P \in b^*(X_s)$, $\alpha_s \in a^*(X_s)$ and M^P is the continuous local martingale part in the P-semimartingale decomposition of X. As previously, we denote $\langle M^P \rangle = \int_0^{\cdot} \alpha_s ds$. Since $F \in C^2(\mathbb{R})$, the Itô formula yields that

$$F(X_t) = F(X_0) + \int_0^t \left(F'(X_s)\beta_s^P + \frac{1}{2}F''(X_s)\alpha_s \right) ds + \int_0^t F'(X_s) dM_s^P$$

and, hence,

$$\tilde{\beta}_s^P = F'(X_s)\beta_s^P + \frac{1}{2}F''(X_s)\alpha_s \tag{43}$$

$$\tilde{\alpha}_s = (F'(X_s))^2 \alpha_s.$$

Now the properties $\tilde{\alpha}_s \in a^F(X_s)$ and $\tilde{\beta}_s^P \in b^F(X_s)$ can be checked directly. $\square$

Remark 5.2. *Intuitively, the above result allows to construct the non-linear process $\tilde{X} = F(X)$ when X is non-linear affine. The new bounds for the (differential) semimartingale characteristics are given by $b^F(X)$ and $a^F(X)$, respectively. However, the drift and volatility of $\tilde{X}$ now relate to each other, which often gives a substantially smaller class in comparison to all semimartingale laws whose drift and volatility stay in $b^F(X_t)$ and $a^F(X_t)$.*

In general, (non-linear) affine processes are stable under affine transformation. The following example shows that, we may even consider the non-linear transformation $F(x) = x^2$, at least in some special cases.

Example 5.3. Let $\mathcal{A}(x,\Theta)$ be a non-linear Vasiček model satisfying $\bar{b}^0 = \underline{b}^0 = 0$, and $\tilde{X} = F(X) = X^2$. We apply Proposition 5.1: first, note that since $F'' = 2 > 0$,

$$b^F(x) = [2x^2 \underline{b}^1 + \underline{a}^0, 2x^2 \bar{b}^1 + \bar{a}^0],$$

and $a^F(x) = [4x^2 \underline{a}^0, 4x^2 \bar{a}^0]$. Then, b^F and a^F can even be written as functions of $\tilde{X} = X^2$. This would not be the case if $\bar{b}^0 = \underline{b}^0 = 0$ does not hold, since b^F would depend on x (which is not a function of x^2). Under this observation we may directly study the semimartingale characteristics of $\tilde{X}$. Replacing x^2 by $\tilde{x}$ in b^F and a^F we indeed observe an affine structure and it is tempting to conjecture that we obtained a non-linear CIR model. In general, this is not the case: for simplicity, choose $\underline{a}^0 = \underline{b}^1 = 0$ and $\bar{a}^0 = \bar{b}^1 = 1$ and $x = 1$. Then $b^F(1) = [0, 3]$ and $a^F(1) = [0, 4]$. For a non-linear CIR model, any choices of $(\tilde{\beta}, \tilde{\alpha})$ in $b^F \times a^F$ should be possible. Now choose, say, $\tilde{\alpha} = 4$ (corresponding to a maximar volatility of $\alpha = 1$ in the original model). Then not all choices of $\tilde{\beta} \in [0, 3]$ are reached by the original model: indeed, one immediately obtains from (43) that $\tilde{\beta}$ needs to lie in $[1, 3]$.

In the choice where only one parameter (either α or β) carries uncertainty, this problem of course vanishes. This is the case for the existing transformations of g- and $G-$Brownian motion in the literature and we provide further examples in this direction below.

The above example also illustrates, that non-linear transformations of processes under ambiguity should be handled with care. The following example shows how to obtain a geometric kind of dynamics, which allows us to obtain the *non-linear Black-Scholes* model as considered in (Epstein and Ji, 2013, Example 3) and Vorbrink (2014). Both works consider the case where there is only volatility uncertainty.

Example 5.4. Let $\mathcal{A}(x, \Theta)$ be a non-linear affine process and consider $F(X) = e^X$. Again, we apply Proposition 5.1. First, note that with $\tilde{x} = e^x$,

$$a^F(x) = (e^x)^2 a^*(x).$$

Moreover, since $a^*(x) = [\underline{a}^0 + \underline{a}^1 x^+, \bar{a}^0 + \bar{a}^1 x^+]$, we obtain

$$a^F(x) = (\tilde{x})^2 [\underline{a}^0 + \underline{a}^1 (\log \tilde{x})^+, \bar{a}^0 + \bar{a}^1 (\log \tilde{x})^+] = \tilde{a}(\tilde{x}), \qquad (44)$$

and we already computed $\tilde{a}$. In a similar manner, one obtains $\tilde{b}$ from (40) noting that

$$b^F(x) = \begin{cases} [e^x(\underline{b}^0 + \underline{b}^1 x) + \frac{1}{2}e^x(\underline{a}^0 + \underline{a}^1 x^+), e^x(\bar{b}^0 + \bar{b}^1 x) + \frac{1}{2}e^x(\bar{a}^0 + \bar{a}^1 x^+)], & x \geq 0 \\ [e^x(\underline{b}^0 + \bar{b}^1 x) + \frac{1}{2}e^x\underline{a}^0, e^x(\bar{b}^0 + \underline{b}^1 x) + \frac{1}{2}e^x\bar{a}^0], & x < 0. \end{cases}$$

The state space of e^X is of course $\mathbb{R}_{>0}$.

Example 5.5 (The non-linear Black-Scholes model)**.** Allowing for *drift* and *volatility* uncertainty in the log-price of a stock, one arrives at a non-linear Black-Scholes model. We consider a Brownian motion with drift and volatility uncertainty, which is in our language a non-linear Vasiček model with $\underline{b}^1 = \bar{b}^1 = 0$. Furthermore, we assume that the stock price is given by $S = \exp(X)$, i.e. $F(x) = e^x$. Then, the calculations from the previous example immediately yield that the stock price is given by the non-linear process $\tilde{X}$ where $a^F(x) = \tilde{a}(F(x))$ and $b^F(x) = \tilde{b}(F(x))$ with

$$\tilde{a}(x) = x^2 [\underline{a}^0, \bar{a}^0]$$

and

$$\tilde{b}(x) = [x\underline{b}^0 + \frac{1}{2}x\underline{a}^0, x\bar{b}^0 + \frac{1}{2}x\bar{a}^0].$$

Option pricing for monotone convex (concave) pay-offs can immediately be done by Proposition 4.6, see Example 3 in Epstein and Ji (2013) for explicit formulae for call options (with no uncertainty of the drift). The article Vorbrink (2014) excludes drift uncertainty by arguing that under risk-neutral pricing the drift is known.

Example 5.6 (The Heston model with uncertainty in the volatility parameters)**.**
The model put forward in Heston (1993) is one of the most popular models for stochastic volatility, which also is heavily used in foreign exchange markets. Model and calibration risk is an important issue, see for example Guillaume and Schoutens (2012) in this regard. Here we give a short outline how a non-linear version could be constructed, allowing for parameter uncertainty in volatility only (and not in the drift of the stock price or in the correlation of volatility and stock price). In this regard, we extend Ω in the classical way to construct an additional (independent) Brownian motion $\tilde{V}$ which allows us to construct two correlated Brownian motions V and W. The correlation is fixed and denoted by ρ. Each $P \in \mathfrak{P}(\Omega)$ is extended by leaving $\tilde{V}$ untouched, such that (V, W) will be a two-dimensional Brownian motion where V and W have correlation ρ and we denote this new semimartingale law again by P.

Consider a non-linear CIR process $\mathcal{A}(x, \Theta)$ with state space $\mathcal{O} = \mathbb{R}_{>0}$ as introduced in Section 4.3. The stock price S is given by the strong solution of the SDE

$$dS_t = S_t X_t dV_t, \qquad 0 \leq t \leq T$$

where X is the canonical process on Ω (and hence a non-linear process). Hence the volatility X stems from a non-linear CIR model which means intuitively, that we have a CIR model with parameter uncertainty with upper and lower bounds $\bar{b}^0, \bar{b}^1, \bar{a}^1$ and $\underline{b}^0, \underline{b}^1, \underline{a}^1$, respectively. For simplicity we chose a vanishing risk-free rate of interest.

We show how to compute a call-price in this non-linear Heston model in the following. The call price $C(T, K)$ for maturity $T \leq T^*$ and strike $K > 0$ is given by the supremum of the expectations $E^P[(S_T - K)^+]$ over all (extended) semimartingale laws P from $\mathcal{A}(x, \Theta)$.

Since the pay-off function $(s - K)^+$ is increasing and convex, the arguments of Proposition 4.6 apply and

$$C(T, K) = E^{\bar{P}}[(S_T - K^+)],$$

where $\bar{P}$ is the worst-case semimartingale law which achieves the supremum. Again from the proof of Proposition 4.6 we find that under $\bar{P}$, X is a (classical) CIR-process with parameters $\bar{b}^0, \bar{b}^1, \bar{a}^1$. The call price formula can be found in Heston (1993), see also Section 10.3.3. in Filipović (2009) for a derivation using Fourier inversion techniques.

6. AFFINE TERM STRUCTURE MODELS

One of the most important application of affine models is in term structure models. In this regard, we provide in the following a term-structure equation for non-linear affine models implying prices for derivatives or bond-prices.

Consider a payoff $f(X_T)$ taking place at time $T > 0$. In the classical setting, arbitrage-free prices are given by expectations of the discounted pay-off under a risk-neutral measure. According to the superhedging duality in Biagini et al. (2017) (Theorem 5.1), in the case we consider here – when there is a family of such measures – upper bounds of these price processes (and hence the smallest superhedging price) given $X_t = x$ are given by

$$F(T - t, x) = \sup_{P \in \mathcal{A}(t,x,\Theta)} E^P\left[e^{-\int_t^T X_s ds} f(X_T) | X_t = x\right], \quad 0 \le t \le T. \tag{45}$$

The following result states the non-linear term-structure equation for the pay-off $f(X_T)$.

Proposition 6.1. *Assume that f is Lipschitz-continuous. Then $F(t, x)$ is a viscosity solution of*

$$\partial_t F(t, x) - \sup_{\theta \in \theta} \mathscr{L}^\theta F(t, x) + x F(t, x) = 0, \tag{46}$$

with boundary condition $F(0, x) = f(x)$. If in addition,

(i) $\underline{a}^0 > 0$, $\mathscr{O} = \mathbb{R}$, and f is bounded, then $F(t, x)$ is the unique solution of (46), or

(ii) if $\underline{a}^0 = \bar{a}^0 = 0$, $\underline{b}^0 \ge \bar{a}^1/2 > 0$ and $\mathscr{O} = \mathbb{R}_{>0}$, then $F(t, x)$ is the only viscosity solution, such that

$$\sup_{(t,x) \in [0,T] \times \mathbb{R}_{>0}} \frac{|F(t,x)|}{1 + x} < \infty. \tag{47}$$

For a proof of this result one can argue the same way as in the proof of Theorem 4.1. More precisely, dynamic programming yields for any stopping time τ taking values in $[t, T]$ that

$$F(T - t, x) = \sup_{P \in \mathcal{A}(t,x,\Theta)} E^P\left[e^{-\int_t^\tau X_s ds} F(T - \tau, X_\tau)\right].$$

Then, following the arguments in Theorem 4.1 leads to the desired result. Alternatively, one could also enlarge the state space to transform this control problem which is in Lagrange form to one of the Mayer form like in Proposition 3.3 and Theorem 4.1 (see for example Remark 3.10 in Bouchard and Touzi (2011)).

The term-structure equation now allows to obtain the bond prices by considering the pay-off $f(X_T) = 1$. We illustrate how an extension of the state space can be used to achieve a result similar to Proposition 4.6 leading to closed-form bond prices in special cases.

In our approach, upper bond prices under the non-linear affine term structure model $\mathcal{A}(t, x, \Theta)$, $x \in \mathscr{O}$, are given by

$$\bar{p}(t, T, x) = \sup_{P \in \mathcal{A}(t,x,\Theta)} E^P\left[e^{-\int_t^T X_s ds} | X_t = x\right], \quad 0 \le t \le T, \tag{48}$$

conditional on $X_t = x$. For arbitrary $\omega \in \Omega$ one obtains the bond price as $\bar{p}(t, T)(\omega) = \bar{p}(t, T, \omega_t)$. The respective lower bond price $\underline{p}(t, T, x)$ is obtained by replacing the supremum with an infimum. The following proposition shows that in important special cases these prices can be obtained in closed form. Again, for $x \in \mathscr{O}$ recall from (33) and (34) the upper bound of $a^*(x)$ and $b^*(x)$ defined by

$$\bar{a}(x) := \bar{a}^0 + \bar{a}^1 x^+, \qquad \bar{b}(x) := \bar{b}^0 + \underline{b}^1 x \mathbb{1}_{\{x<0\}} + \bar{b}^1 x \mathbb{1}_{\{x \ge 0\}},$$

and recall $\bar{B}^{1,x} = \underline{b}^1 \mathbb{1}_{\{x<0\}} + \bar{b}^1 \mathbb{1}_{\{x \ge 0\}}$, so we obtain that $\bar{b}(x) = \bar{b}^0 + \bar{B}^{1,x} x$. Moreover, recall from (36) the affine process with coefficients $\bar{a}^0, \bar{a}^1, \bar{b}^0, \bar{B}^{1,x}$.

Proposition 6.2. *Consider a non-linear affine process $\mathcal{A}(x, \Theta)$, assume for all $P \in \mathcal{A}(x, \Theta)$ that*

$$\beta_t^P \leq \bar{b}^0 + \bar{B}^{1,x} X_t \tag{49}$$

$dP \otimes dt$-almost everywhere for $0 \leq t \leq T$ and assume that either $\underline{a}^1 = \bar{a}^1 = 0$ or that for all $P \in \mathcal{A}(x, \Theta)$, $X_t \geq 0$ $P \otimes dt$-a.s. If there exists $\bar{P} \in \mathcal{A}(x, \Theta)$ and a $\bar{P}$-$\mathbb{F}$-Brownian motion W such that the canonical process under $\bar{P}$ is the unique strong solution of (36), *then, for all $u \geq 0$ and $0 \leq t \leq T$,*

$$\bar{p}(t, T, x) = \exp\big(\phi(t, u) + \psi(t, u)x\big),$$

where ϕ, ψ solve the Riccati equations

$$\partial_t \phi(t, u) = \frac{1}{2}\bar{a}^0 \psi(t, u)^2 + \bar{b}^0 \psi(t, u), \qquad \phi(0, u) = 0, \tag{50}$$

$$\partial_t \psi(t, u) = \frac{1}{2}\bar{a}^1 \psi(t, u)^2 + \bar{B}^{1,x} \psi(t, u) - 1, \qquad \psi(0, u) = u. \tag{51}$$

Proof. This results also follows by using semimartingale comparison. In this regard let $P \in \mathcal{A}(x, \Theta)$ and consider the two-dimensional process $Y = (Y^1, Y^2)$ where $Y^1 = -\int_0^{\cdot} X_s ds$ and $Y^2 = X$. Then, there is no parameter uncertainty with respect to the dynamics of Y^1 since its differential semimartingale characteristics are obtained from $dY_t^1 = -X_t dt$.

By assumption there exists a $\bar{P} \in \mathcal{A}(x, \Theta)$ and a $\bar{P}$-$\mathbb{F}$-Brownian motion W, such that X is the unique strong solution of the stochastic differential equation (36). Denote by β, α the differential semimartingale characteristics under P of the two-dimensional semimartingale Y and by $\bar{\beta}, \bar{\alpha}$ those of Y under $\bar{P}$. It is easily verified that $\beta_t \leq \bar{\beta}_t$ and also that $\alpha_t \leq_{\text{psd}} \bar{\alpha}_t$ in the positive semidefinite order[3].

Since $(y^1, y^2) \mapsto e^{u^1 y^1}$ for $u \geq 0$ is increasing and convex, Theorem 2.2 in Bergenthum and Rüschendorf (2007) (the propagation of order (PO) property is shown in Appendix A) yields that,

$$E^P\big[e^{-\int_0^t X_s ds}\big] \leq E^{\bar{P}}\big[e^{-\int_0^t X_s ds}\big].$$

Since, $P \in \mathcal{A}(x, \Theta)$ was chosen arbitrarily and under $\bar{P}$, X is an affine process with parameters $\bar{a}^0, \bar{a}^1, \bar{b}^0, \bar{B}^{1,x}$, the affine representation follows and the Riccati equations in (50)-(51) can be obtained directly from Theorem 10.4 in Filipović (2009). $\qquad \square$

Remark 6.3. *Again, as in Remark 4.7, it can be easily checked that the conditions for Proposition 6.2 hold in the following two cases:*

(1) The non-linear Vasiček model with state space $\mathscr{O} = \mathbb{R}$ and $\bar{b}^1 = \underline{b}^1$,

(2) the non-linear CIR model on the state space $\mathscr{O} = \mathbb{R}_{>0}$.

It is remarkable, that in the general non-linear Vasiček model with parameter uncertainty on the speed of mean reversion, the classical exponential affine bond pricing formula ceases to hold. Thus, in this model, the interval of parameters can not directly be backed out from a interpolation of bid and ask prices with a standard Vasiček model. This however is the case in the CIR model and in the Vasiček model where b^1 is known.

The previous results and examples directly allow the treatment of term-structure models based on the non-linear Vasiček model and on the non-linear CIR model (see Sections 4.2 and 4.3). An important difficulty for the modeller in practical situations is that she has to make her choice between these models with strong implications: for example, the state space can allow for negative values or can strictly exclude them, a well-documented difficulty of the affine models after the European crisis, see Carver (2012). The following example shows that in a non-linear setting one is able to mix these two type of models and the modeller no longer has to decide a priori if she allows or excludes negative values.

6.1. **The non-linear Vasiček-CIR model.** If one is not able to restrict the state space a priori, one can consider the following non-linear affine model: assume that both parameters a_0 and a_1 are subject to parameter uncertainty (or at least one of them with the other parameter not vanishing). Intuitively, this means that the model may switch between a Vasiček-like or CIR-like behaviour. In particular, when one is not able to restrict the state space a priori to $\mathbb{R}_{\geq 0}$, this non-linear model allows to incorporate both model approaches in a robust (i.e. non-linear) sense.
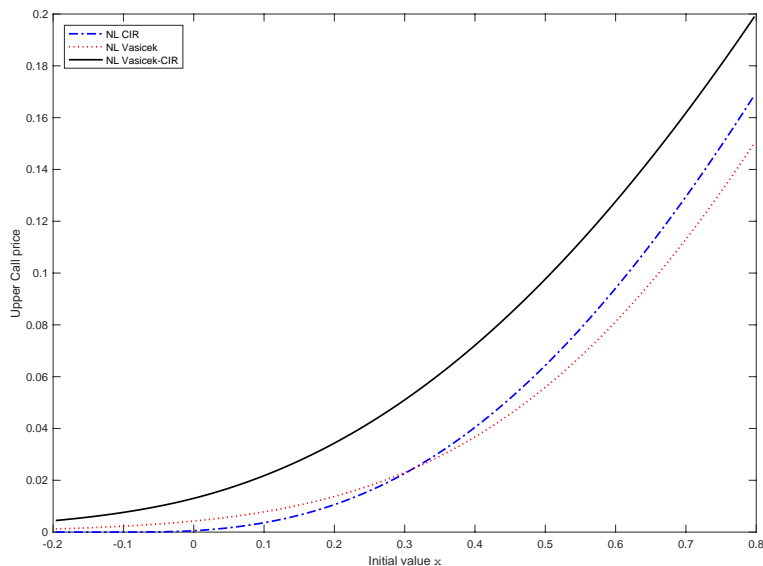
---

[3] Here, $A \leq_{\text{psd}} B$ means $x^{\top}(B - A)x \geq 0$ for all $x \in \mathbb{R}^2$.

In the interest rate markets in the early years after 2000, market participants believed in positive interest rates and thus favoured the CIR-model. The credit crises led to decreasing interest rates and the Vasiček model came back, as it allows for negative interest rates. This effect is well-known and its implications for banking are quite important, see for example Carver (2012), Patel et al. (2017), Orlando et al. (2016), Russo and Fabozzi (2017). In the near future, however, when interest rates may rise again one could be interested in deviating again from the Vasiček-model. With the non-linear Vasiček-CIR model such a switch is no longer necessary and one is able to behave consistently through such seemingly different time periods.

More precisely, assume that $\underline{a}^0 > 0$ and consider the state space $\mathscr{O} = \mathbb{R}$. Uniqueness for the non-linear PDE (16) follows readily from Proposition 4.2. In this general case, there will be no explicit solutions like for example in Proposition 6.2 above and we have to rely on numerical techniques.

Figure 2 illustrates the differences in option pricing of the non-linear Vasiček-CIR model to both the Vasiček and the CIR model. While in the non-linear CIR model, this price can still be computed explicitly, this no longer holds for the other two models. The parameters of the plot are chosen for best illustration purposes. The plot shows the (upper) price of a Call option under a non-linear affine model with the parameters stated in the plot. While the non-linear Vasiček model puts clearly weight on $\mathbb{R}_{<0}$, the CIR model shifts the mass towards the right, leading to higher option values for $x > 0.3$.

The graph further highlights the mentioned difficulty of the modeler to choose from one model: starting from positive values (i.e. here $x > 0.3$), he possibly would prefer to choose the (classical/non-linear) CIR model, while for falling values prices become prohibitively low, illustrating the need to switch to the Vasiček model. The non-linear Vasiček-CIR model clearly does not suffer from this problem (it is clear by definition that the parameters of the non-linear Vasiček-CIR model can be chosen to reproduce the option prices of the two other models).

Parameters	$\overline{a}_0$	$\underline{a}_0$	$\overline{a}_1$	$\underline{a}_1$	$\overline{b}_0$	$\underline{b}_0$	$\overline{b}_1$	$\underline{b}_1$
	0.08	0.00	0.2	0.00	0.15	0.05	-0.5	-1

FIGURE 2. This figure shows the upper Call price for the non-linear Vasiček, the non-linear CIR and the non-linear Vasiček-CIR model. The first two models are obtained from the latter by simply letting $\bar{a}^1 = 0$ ($\bar{a}^0 = 0$, respectively). The Call price has strike 0.5 and the parameters are given in the table above.

7. Model risk

In financial applications, model risk is an important factor for risk management. In the remarkable work Cont (2006), a systematic framework for the management of model risk has been proposed which we recall shortly and thereafter apply to the non-linear affine models. The importance of this topic is illustrated by the intensive research in this area, see for example Bannör and Scherer (2013), Guillaume and Schoutens (2012), Breuer and Csiszár (2016), Barrieu and Scandolo (2015), da Fonseca and Grasselli (2011) among many others.

In this approach, the market contains a number of benchmark instruments which are liquidly traded instruments and the observation consists in bid and ask prices thereof. Moreover, there is a set of arbitrage-free pricing models $\mathcal{Q}$ which is consistent with the observations of the benchmark instruments.

In our framework, both can be described through a non-linear affine model: the non-linear affine model specifies a set $\mathcal{Q}$ of pricing measures, as for example in the non-linear affine term-structure approach studied in Section 6. Consistent bid and ask prices can be obtained by suprema and infima over these pricing measures, exactly as it was done for $\underline{p}(t, T)$ and $\bar{p}(t, T)$ in the previous section.

A *coherent* measure of model uncertainty for a payoff function $\psi : \mathbb{R} \to \mathbb{R}$ with $\psi(X_T)$ denoting the payoff at time T can be computed from the upper and lower price bounds

$$\bar{\pi}(\psi) = \sup_{Q \in \mathcal{Q}} E^Q[\psi(X_T)], \qquad \underline{\pi}(\psi) = \inf_{Q \in \mathcal{Q}} E^Q[\psi(X_T)]. \tag{52}$$

The measure of model uncertainty on the derivative ψ is given by

$$\mu_{\mathcal{Q}}(\psi) = \bar{\pi}(\psi) - \underline{\pi}(\psi).$$

Some examples are provided in Cont (2006), including the non-linear Black-Scholes model. We illustrate the application of non-linear affine models in this framework with a short study of model risk in the non-linear Vasiček model.

7.1. **The non-linear Vasiček model.** As an example we consider the non-linear Vasiček-model introduced in Section 4.2. Recall that this model is characterized by the assumption that $[\underline{a}^1, \bar{a}^1] = \{0\}$. In the non-linear case, non-linear expectations are given by the solution of the non-linear Kolmogorov equation (15). Proposition 4.6 allows to trace this solution back to existing solutions for affine models if the payoffs are increasing and convex (decreasing and concave), see Example 5.6. For more general payoffs we rely on numerical methods which we illustrate now.

7.1.1. *Options.* The model risk for options in this model is illustrated in the following two pictures, where we price a call and a butterfly. To construct the set Θ, we take estimated parameter values together with their 95% confidence intervals from the literature. The results is shown in Figure 3. While for the Call option the model risk increases monotonically with the initial value, the maximal model risk for the butterfly is attained for the initial value x directly at the maximal payoff.

8. Conclusion

In this paper we introduced affine processes under uncertainty. This extends the existing class of non-linear Lévy processes to Markov processes where the interval for the parameter uncertainty may depend on the current state (in an affine way, however). We obtained a dynamic programming principle implying a non-linear Kolmogorov equation which can be used to price options in a fast and efficient way. Many existing models can be embedded into a setting with parameter uncertainty which we illustrate with a number of examples. However, the non-linear framework also allows for new model variations which did not exist in the classical approach and we illustrate this with a term-structure Vasiček-CIR model, where the modeler does not need to decide a priori if the state space should include negative rates or not, a strong restriction in existing models. The generalization to higher dimensions or to the case with jumps is left for future research. Here, we concentrated on the conceptual introduction of state-dependent parameter uncertainty and chose the simplest but still highly interesting example for the illustration of our ideas.
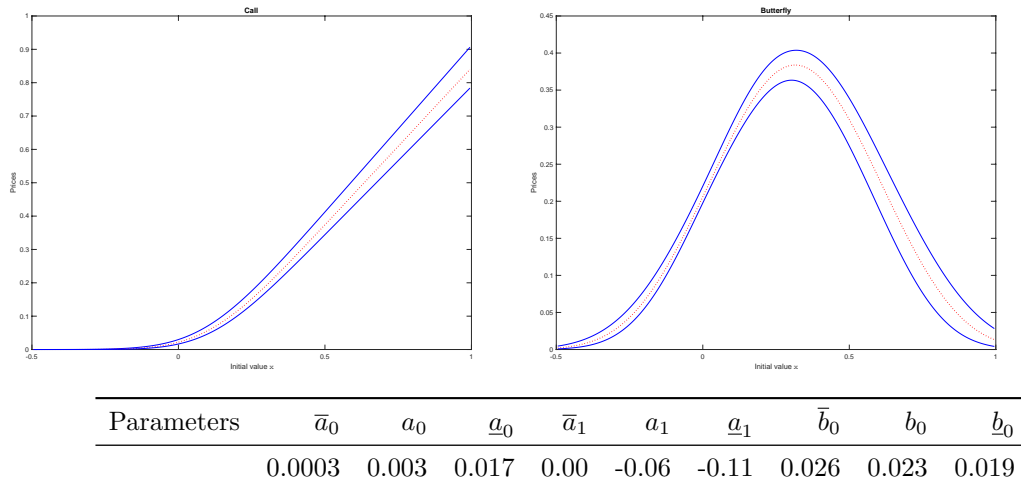
Parameters	$\overline{a}_0$	a_0	$\underline{a}_0$	$\overline{a}_1$	a_1	$\underline{a}_1$	$\overline{b}_0$	b_0	$\underline{b}_0$
	0.0003	0.003	0.017	0.00	-0.06	-0.11	0.026	0.023	0.019

FIGURE 3. This figure shows the solution of the non-linear Kolmogorov equation for the non-linear Vasiček-model with boundary condition $f(x) = (x - K)^+$ and $f(x) = (x - K_1)^+ - 2(x - K_2)^+ + (x - K_3)^+$, $K = 0.1$, $K_1 = -0.2$, $K_2 = 0.3$, and $K_3 = 0.8$. For each of the figures, the dashed lines show the solution of the Vasiček model with no uncertainty. The upper and lower solid lines gives the upper and lower price bounds. The parameters for the computation are given in the table below the pictures.

APPENDIX A. SEMIMARTINGALE COMPARISON

In this section we recall the main result in Bergenthum and Rüschendorf (2007) on comparison of semimartingales with Markov processes and show that the crucial *propagation of order* (PO) property is satisfied for a large class of Markov processes, in particular for affine processes. To the best of our knowledge, existing results in the literature require Lipschitz assumptions on the coefficient, which will not hold in our case.

As function class $\mathcal{F}$ we will consider increasing and convex functions. For a real-valued Markov process S^* and a terminal time T we define the propagation operator

$$\mathcal{G}^g(t, x) = E[g(S_T^*)|S_t^* = x].$$

Assumption A.1. *For some function class $\mathcal{F}$ and some Markov process S^* we say that $PO(S^*, \mathcal{F})$ holds if $\mathcal{G}^g(t, \cdot) \in \mathcal{F}$ for all $0 \le t \le T$ and for all $g \in \mathcal{F}$.*

Propagation of monotonicity and convexity follows in a very elegant way through total positivity of the transition densities of continuous Markov processes.

Proposition 3.1 in Kijima (2002) immediately yields the following result.

Proposition A.2. *Assume that S^* is a strong Markov process having continuous sample paths and that g is increasing (decreasing). Then $\mathcal{G}^g(0, \cdot)$ is increasing (decreasing).*

For the propagation of convexity we need an additional step, because the considered processes in Kijima (2002) are in fact martingales. The proof crucially uses the variation-diminishing property of totally positive functions.

Proposition A.3. *Assume that S^* is a strong Markov process with state space S having continuous sample paths and that there exist π_0, $\pi_1 \in \mathbb{R}$ with $\pi_1 \ne 0$, such that*

$$E[S_T^*|S_0^* = x] = \pi_0 + x\pi_1, \qquad x \in S. \tag{53}$$

Then for convex (concave) functions g it holds that $\mathcal{G}^g(0, \cdot)$ is convex (concave).

Proof. We modify the first step in Proposition 3.2 in Kijima (2002). In this regard, note that Equation (53) yields that

$$x = \frac{E[S_T^*|S_0^* = x] - \pi_0}{\pi_1}. \tag{54}$$

Moreover, we follow the notation in Kijima (2002) and denote by $q_T(x, y)$ the transition density of the Markov process S^*, i.e.

$$\mathcal{G}^g(0, x) = E[g(S_T^*)|S_0^* = x] = \int_S g(y) \, q_T(x, y) dy. \tag{55}$$

Let $\tilde{\alpha}_1, \tilde{\alpha}_2 \in \mathbb{R}$. From (54) and (55) we obtain that

$$\mathcal{G}(0, x) - \tilde{\alpha}_1 x - \tilde{\alpha}_2 = \int_{\mathbb{R}} q_T(x, y)\Big\{g(y) - \tilde{\alpha}_1 x - \tilde{\alpha}_2\Big\}dy$$

$$= \int_{\mathbb{R}} q_T(x, y)\Big\{g(y) - \frac{\tilde{\alpha}_1}{\pi_1}y + \frac{\tilde{\alpha}_1 \pi_0}{\pi_1} - \tilde{\alpha}_2\Big\}dy.$$

Since $\frac{\tilde{\alpha}_1}{\pi_1}y + \frac{\tilde{\alpha}_1 \pi_0}{\pi_1} - \tilde{\alpha}_2 =: \alpha_1 y + \alpha_2$ is an affine function of y, the claim follows by repeating the arguments in Proposition 3.2 in Kijima (2002). □

Consider in addition a different Markov process S, possibly on a different probability space and denote by

$$\mathcal{F}_{icx} := \{f : \mathbb{R} \to \mathbb{R}, \text{increasing and convex}\}.$$

A combination of Proposition A.2 with A.3 yields the following result.

Proposition A.4. *Let S^* be a strong and homogeneous Markov process with continuous sample paths and that*

$$E[S_t^*|S_0^* = x] = \pi_0(t) + x\pi_1(t), \qquad 0 \le t \le T \tag{56}$$

holds with $\pi_1(t) \ne 0$ for all $t \in [0, T]$. Then $PO(S^*, \mathcal{F}_{icx})$ holds.

Proof. Let $g \in \mathcal{F}_{icx}$. First, Proposition A.2 yields that $\mathcal{G}^g(0, \cdot)$ is increasing. As the choice of T was arbitrary, we obtain that also $\mathcal{G}^g(t, \cdot)$ is increasing by repeating the argument of Proposition A.2 and using homogeneity of S^*.

Second, Proposition A.3 yields that $\mathcal{G}^g(0, \cdot)$ is convex. Denote

$$\mathcal{H}^g(t, x) = E[g(S_t^*)|S_0^* = x].$$

Then, $\mathcal{G}^g(t, x) = \mathcal{H}^g(T - t, x)$ since the Markov process is homogeneous. Now Proposition A.3 yields that $\mathcal{H}^g(t, \cdot)$ is convex since the choice of T was arbitrary and the claim follows. □

We remark that for a continuous, one-dimensional affine process X with $[\underline{a}^1, \bar{a}^1] = \{0\}$ or $\mathcal{O} = \mathbb{R}_{>0}$, affinity of the expectation in (53) is satisfied which follows directly from the exponential-affine structure of the Laplace transform $E[e^{uX_t}|X_0 = x] = \exp(\phi(t, u) + \psi(t, u)x)$. Indeed, an application of Theorem 13.2 in Jacod and Protter (2004) yields that

$$E[X_t|X_0 = x] = \partial_u \phi(t, 0) + x\partial_u \psi(t, 0),$$

where we use the explicit expressions for ϕ and ψ from Section 10.3.2 in Filipović (2009).

APPENDIX B. COMPARISON RESULTS

In this section we recall the comparison results from Amadori (2007) in our notation. Again, the crucial point for this results is that Lipschitz assumptions on the full domain do not hold. Note that we only consider the one-dimensional, time-homogeneous case here, which simplifies the matter significantly. While minimization is the core topic of Amadori (2007), the financial applications mainly treat maximization, such that we concentrate on the maximization. The stated results follow from the original results by replacing ψ with $-\psi$.

Fix the state space $\mathcal{O} = \mathbb{R}_{>0}$ and consider the controlled diffusion $X = X^\theta$

$$dX_s = b(X_s, \theta_s)ds + \sqrt{a(X_s, \theta_s)}dW_s, \quad s > t \tag{57}$$

with initial condition $X_t = x \in \mathcal{O}$. Our application will be in Proposition 4.3, which considers the non-linear CIR-modell. Since then $\underline{a}^0 = \bar{a}^0 = 0$ we consider $\Theta = [\underline{b}^0, \bar{b}^0] \times [\underline{b}^1, \bar{b}^1] \times \{0\} \times [\underline{a}^1, \bar{a}^1]$ and an adapted process (θ_s) taking values in Θ. Moreover, for each $x \in \mathcal{O}, \theta \in \Theta, b(x, \theta) = b^0 + b^1 x$ and $\sqrt{a(x, \theta)} = \sqrt{a^1 x}$ which is not Lipschitz at 0.

Now, fix a continuous function $\psi : \mathcal{O} \to \mathbb{R}$ and define the value function $v : [0, T] \times \mathcal{O} \to \mathbb{R}$ by

$$v(t, x) := \sup_{(\theta_s)} E_{x,t}[\psi(X_T)],$$

where the supremum ranges over all adapted processes (θ_s) taking values in Θ, X is the controlled diffusion satisfying (57), and $E_{t,x}$ refers to the conditional expectation conditioning on $X_t = x$. The associated *Hamilton-Jacobi-Bellman* equation is given in (16).

First, Assumption 1 in Amadori (2007) holds. Indeed, note that the functions f and r therein equal to zero in our case, that the functions b and $\sqrt{a}$ are Lipschitz-continuous on $[\epsilon, \infty)$ for all $\epsilon > 0$ and all $\theta \in \Theta$. Moreover, let $\| \cdot \|_{C^{0,1}([\epsilon,\infty))}$ denote the Lipschitz[4] coefficient on $[\epsilon, \infty)$, then clearly

$$\sup\{\| \sqrt{a(\cdot, \theta)} \|_{C^{0,1}([\epsilon,\infty))} \colon \theta \in \Theta\} < \infty,$$

and, similarly, for b and all conditions of Assumption 1 hold.

Second, Assumption 4 is implied by the Feller condition $\underline{b}^0 \geq \bar{a}^1/2 > 0$. Indeed, Assumption 4 requires that

$$\limsup_{x \to 0} \sup_{\theta \in \Theta} \Big(\frac{1}{x} - \frac{2b(x,\theta)}{a(x,\theta)}\Big) < \infty.$$

Since $\underline{b}^0 \geq \bar{a}^1/2 > 0$ implies that

$$\frac{1}{x} - \frac{2b(x,\theta)}{a(x,\theta)} = \frac{1}{x}\Big(1 - \frac{2b^0}{a^1}\Big) - \frac{2b^1}{a^1} \leq -\frac{2b^1}{a^1},$$

Assumption 4 also holds. In our notation, the uniqueness result following immediately from the comparison principle given in Theorem 4 in Amadori (2007) reads as follows. We refer to the Section 4 for the definitions of viscosity solutions, super- and subsolutions.

Theorem B.1. *Assume that $\underline{b}^0 \geq \bar{a}^1/2 > 0$. Let v be a continuous locally bounded viscosity solution of* (16).

(i) *Assume*

$$\sup_{\mathscr{O} \times [0,T]} \frac{|v(t,x)|}{1+x} < \infty. \tag{58}$$

Let u be a locally bounded viscosity solution satisfying (58), *then $u = v$.*

(ii) *Assume*

$$\limsup_{x \to \infty} \sup_{t \in [0,T]} \frac{|v(t,x)|}{x^2} = 0, \quad \lim_{x \to 0} \sup_{t \in [0,T]} \frac{|v(t,x)|}{h_\epsilon(x)} = 0. \tag{59}$$

Let u be a locally bounded viscosity solution satisfying (59), *then $u = v$.*

Proof. Recall that a viscosity solution is a supersolution and a subsolution. Since u and v are both viscosity solutions, $u = v$ holds on $\mathscr{O} \times \{T\}$. For (i), note that the conditions hold with $\gamma = 1$ both for u and v, such that applying Theorem 4 in Amadori (2007) twice (once as a supersolution and once as a subsolution) yields $u = v$ on $\mathscr{O} \times [0,T]$. The claim (ii) follows similarly. $\square$

REFERENCES

Acciaio, B., Beiglböck, M., Penkner, F. and Schachermayer, W. (2016), 'A model-free version of the fundamental theorem of asset pricing and the super-replication theorem', *Math. Finance* **26**(2), 233–251.

Amadori, A. L. (2007), 'Uniqueness and comparison properties of the viscosity solution to some singular HJB equations', *Nonlinear Differential Equations and Applications NoDEA* **14**(3-4), 391–409.

Avellaneda, M., Levy, A. and Parás, A. (1995), 'Pricing and hedging derivative securities in markets with uncertain volatilities', *Appl. Math. Finance* **2**(2), 73–88.

Bannör, K. F. and Scherer, M. (2013), 'Capturing parameter risk with convex risk measures', *European Actuarial Journal* (1), 97–132.

Barrieu, P. and Scandolo, G. (2015), 'Assessing financial model risk', *European Journal of Operational Research* **242**(2), 546 – 556.

Bergenthum, J. and Rüschendorf, L. (2007), 'Comparison of semimartingales and Lévy processes', *Ann. Probab.* **35**(1), 228–254.

Biagini, S., Bouchard, B., Kardaras, C. and Nutz, M. (2017), 'Robust fundamental theorem for continuous processes', *Mathematical Finance* **27**(4), 963–987.

---

[4]This is the Hölder coefficient for exponent α with $\alpha = 1$, see Section 5.1 in Evans (2010).

Bielecki, T. R., Cialenco, I. and Rutkowski, M. (2018), 'Arbitrage-free pricing of derivatives in nonlinear market models', *Probability, Uncertainty and Quantitative Risk* **3**(1), 2.

Bouchard, B. and Touzi, N. (2011), 'Weak dynamic programming principle for viscosity solutions', *SIAM Journal on Control and Optimization* **49**(3), 948–962.

Breuer, T. and Csiszár, I. (2016), 'Measuring distribution model risk', *Mathematical Finance* **26**(2), 395–411.

Carver, L. (2012), 'Negative rates: Dealers struggle to price 0% floors', *Risk magazine* .

Cont, R. (2006), 'Model uncertainty and its impact on the pricing of derivative instruments', *Mathematical Finance* **16**, 519–542.

Costantini, C., Papi, M. and D'Ippoliti, F. (2012), 'Singular risk-neutral valuation equations', *Finance and Stochastics* **16**(2), 249–274.

Crandall, M. G., Ishii, H. and Lions, P.-L. (1992), 'User's guide to viscosity solutions of second order partial differential equations', *Bull. Amer. Math. Soc.* **27**(1), 1–67.

da Fonseca, J. and Grasselli, M. (2011), 'Riding on the smiles', *Quantitative Finance* **11**(11), 1609–1632.

Denis, L. and Martini, C. (2006), 'A theoretical framework for the pricing of contingent claims in the presence of model uncertainty', *The Annals of Applied Probability* **16**(2), 827–852.

Denk, R., Kupper, M. and Nendel, M. (2017), 'A semigroup approach to nonlinear Lévy processes', *arXiv:1710.08130v1* .

Duffie, D., Filipović, D. and Schachermayer, W. (2003), 'Affine processes and applications in finance', *Annals of Applied Probability* **13**, 984–1053.

Eberlein, E., Madan, D. B., Pistorius, M. and Yor, M. (2014), 'Bid and ask prices as nonlinear continuous time G-expectations based on distortions', *Mathematics and Financial Economics* **8**(3), 265–289.

El Karoui, N. and Tan, X. (2013*a*), 'Capacities, measurable selection and dynamic programming part I: Abstract framework', *arXiv:1310.3363v1* .

El Karoui, N. and Tan, X. (2013*b*), 'Capacities, measurable selection and dynamic programming part II: Application in stochastic control problems', *arXiv:1310.3363v1* .

Epstein, L. G. and Ji, S. (2013), 'Ambiguous volatility and asset pricing in continuous time', *The Review of Financial Studies* **26**(7), 1740–1786.

Evans, L. C. (2010), *Partial differential equations*, American Mathematical Society.

Feller, W. (1951), 'Two singular diffusion problems', *Annals of Mathematics* pp. 173–182.

Filipović, D. (2009), *Term Structure Models: A Graduate Course*, Springer Verlag. Berlin Heidelberg New York.

Fleming, W. H. and Soner, H. M. (2006), *Controlled Markov Processes and Viscosity Solutions*, 2nd edn, Springer, New York.

Fouque, J.-P. and Ren, B. (2014), 'Approximation for option prices under uncertain volatility', *SIAM Journal on Financial Mathematics* **5**(1), 360–383.

Gikhman, I. (2011), 'A short remark on Fellers square root condition', *Available on SSRN* .

Guillaume, F. and Schoutens, W. (2012), 'Calibration risk: Illustrating the impact of calibration risk under the Heston model', *Review of Derivatives Research* **15**(1), 57–79.

Guo, G., Tan, X. and Touzi, N. (2017), 'Tightness and duality of martingale transport on the Skorokhod space', *Stochastic Processes and their Applications* **127**(3), 927–956.

Guyon, J. and Henry-Labordère, P. (2013), *Nonlinear option pricing*, CRC Press.

Heider, P. (2010), 'Numerical methods for non-linear Black-Scholes equations', *Applied Mathematical Finance* **17**(1), 59–81.

Heston, S. (1993), 'A closed-form solution for options with stochastic volatility and applications to bond and currency options', *Review of Financial Studies* **6**, 327–343.

Jacod, J. and Protter, P. (2004), *Probability essentials*, Springer Science & Business Media.

Kallenberg, O. (2002), *Foundations of modern probability*, Probability and its Applications (New York), second edn, Springer-Verlag, New York.

Karatzas, I. and Shreve, S. E. (1988), *Brownian Motion and Stochastic Calculus*, Springer Verlag. Berlin Heidelberg New York.

Kijima, M. (2002), 'Monotonicity and convexity of option prices revisited', *Mathematical Finance* **12**(4), 411–425.

Madan, D. B. (2016), 'Benchmarking in two price financial markets', *Annals of Finance* **12**(2), 201–219.

Muhle-Karbe, J. and Nutz, M. (2018), 'A risk-neutral equilibrium leading to uncertain volatility pricing', *Finance and Stochastics* **22**(2), 281–295.

Neufeld, A. and Nutz, M. (2014), 'Measurability of semimartingale characteristics with respect to the probability law', *Stochastic Processes and their Applications* **124**(11), 3819 – 3845.

Neufeld, A. and Nutz, M. (2017), 'Nonlinear Lévy processes and their characteristics', *Transactions of the American Mathematical Society* **369**(1), 69–95.

Nutz, M. and van Handel, R. (2013), 'Constructing sublinear expectations on path space', *Stochastic Processes and their Applications* **123**(8), 3100 – 3121.

Orlando, G., Mininni, R. M. and Bufalo, M. (2016), 'A revised approach to CIR short-term interest rates model'.

Patel, J., Russo, V. and Fabozzi, F. J. (2017), 'Using the right implied volatility quotes in times of low interest rates: An empirical analysis across different currencies', *Finance Research Letters* .

Peng, S. (1997), Backward SDE and related *g*-expectation, *in* 'Backward stochastic differential equations', Vol. 364 of *Pitman Res. Notes Math. Ser.*, Longman, pp. 141–159.

Peng, S. (2007*a*), '*G*-Brownian motion and dynamic risk measure under volatility uncertainty', *Lecture Notes* .

Peng, S. (2007*b*), '*G*-expectation, *G*-Brownian motion and related stochastic calculus of Itô type', *Stochastic Analysis and Applications* **2**, 541–567.

Revuz, D. and Yor, M. (1999), *Continuous martingales and Brownian motion*, Springer Verlag. Berlin.

Russo, V. and Fabozzi, F. J. (2017), 'Calibrating short interest rate models in negative rate environments', *The Journal of Derivatives* **24**(4), 80–92.

Vorbrink, J. (2014), 'Financial markets with volatility uncertainty.', *Journal of Mathematical Economics* **53**, 64–78.

Wilmott, P. and Oztukel, A. (1998), 'Uncertain parameters, an empirical stochastic volatility model and confidence limits', *International Journal of Theoretical and Applied Finance* **1**(1), 175–189.

Freiburg University, www.stochastik.uni-freiburg.de
E-mail address: `tolulope.fadina@stochastik.uni-freiburg.de`

Nanyang Technological University, Division of Mathematical Sciences, Singapore
E-mail address: `ariel.neufeld@ntu.edu.sg`

Freiburg Institute of Advanced Studies (FRIAS), Germany. University of Strasbourg Institute for Advanced Study (USIAS), France. Department of Mathematical Stochastics, University of Freiburg, Eckerstr.1, 79104 Freiburg, Germany.
E-mail address: `thorsten.schmidt@stochastik.uni-freiburg.de`