ANNALES
POLONICI MATHEMATICI
127.3 (2021)

The investor problem based on the HJM model

Szymon Peszat and Dariusz Zawisza (Kraków)

Abstract. We consider a consumption-investment problem (both on a finite and an infinite time horizon) in which the investor has access to a bond market. In our approach prices of bonds with different maturities are described by the general HJM factor model. We assume that the bond market consists of the whole family of rolling bonds and the investment strategy is a general signed measure distributed on all real numbers representing time to maturity specifications for different rolling bonds. In particular, we can consider a portfolio of coupon bonds. The investor's objective is to maximize the time-additive HARA utility of the consumption process. We solve the problem by means of the HJB equation for which we prove the required regularity of solution and all required estimates to ensure applicability of the verification theorem. Explicit calculations for affine models are presented.

1. Introduction. The famous Merton problem of maximizing the expected utility of a consumption stream has a long tradition and has been solved in many different settings. In the original formulation (see e.g. [28]), we have a market with two possible investments: a safe investment (bank account) with price dynamics

$$\frac{\mathrm{d}B(t)}{B(t)} = r\,\mathrm{d}t, \qquad B(s) = b,$$

and a risky investment (stocks) with price dynamics

$$\frac{\mathrm{d}P(t)}{P(t)} = \mu\,\mathrm{d}t + \sigma\,\mathrm{d}W(t), \qquad P(s) = p,$$

where r, μ and σ are constants, and W is a real-valued Wiener process. The investor at any time t chooses a consumption rate $C(t)$ and can transfer money from one investment to the other without transaction costs. Then

2020 *Mathematics Subject Classification*: 60H30, 93E20, 91G30, 91G10.
Key words and phrases: Merton problem, HJM model, portfolio immunization, optimal consumption, interest rate, rolling bond, G2++ model.
Received 29 April 2021; revised 3 September 2021.
Published online 9 December 2021.

DOI: 10.4064/ap210429-12-11

the dynamics of the corresponding capital $z^{\psi,C}$ is given by

$$\frac{\mathrm{d}z^{\psi,C}(t)}{z^{\psi,C}(t)} = [r(1-\psi_t)+\mu\psi_t - C(t)]\,\mathrm{d}t + \psi_t\sigma\,\mathrm{d}W(t), \qquad z^{\psi,C}(s)=z,$$

where ψ_t is the fraction of the capital invested in stocks at time t. Given a horizon T of investments, a discount factor $\gamma \geq 0$, $\alpha \in (-\infty,0)\cup(0,1)$ and $a, b \geq 0$, the goal is to find a strategy $(\widehat{\psi}, \widehat{C})$ which maximizes the satisfaction (or reward) functional

$$(1.1)\qquad J_T(z,s,\psi,C) = \frac{1}{\alpha}\mathbb{E}\Big[a\int_s^T \mathrm{e}^{-\gamma(t-s)}(C(t)z^{\psi,C}(t))^{\alpha}\,\mathrm{d}t + b\,\mathrm{e}^{-\gamma(T-t)}(z^{\psi,C}(T))^{\alpha}\Big].$$

In the present paper, the investor has access to a market of so-called rolling bonds $U(t)(x)$, $t\in[s,T]$ and $x\in[0,T^*]$. Here T^* is the maximal time to maturity. Given x, the rolling bond $U(\cdot)(x)$ is a self-financing investment in a bond with fixed time to maturity x and therefore is a tradable instrument (for the precise definition see Rutkowski [32] and the Appendix). The dynamics of the rolling bonds is given by the SDE

$$(1.2)\qquad \frac{\mathrm{d}U(t)(x)}{U(t)(x)} = \overline{r}(t)\,\mathrm{d}t - \langle\tilde{\sigma}(t)(x), \lambda(t)\,\mathrm{d}t + \mathrm{d}W(t)\rangle,$$

where $\overline{r}$ is a (random) short rate, $\tilde{\sigma}$ is a random field, λ is a random process taking values in $\mathbb{R}^m$ and W is an m-dimensional Wiener process (see (3.1) and (3.4)).

The investment strategy ψ_t is a normalized signed measure on $[0,T^*]$. The dynamics of the capital $z^{\psi,C}$ has the form

$$(1.3)\qquad \frac{\mathrm{d}z^{\psi,C}(t)}{z^{\psi,C}(t)} = \int_0^{T^*} \frac{\mathrm{d}U(t)(x)}{U(t)(x)}\,\psi_t(\mathrm{d}x) - C(t)\,\mathrm{d}t.$$

The goal is to maximize the satisfaction (or reward) functional

$$J_T(u(s)(\cdot), z^{\psi,C}(s), s)$$

given by the right hand side of (1.1). We will restrict our attention to the Heath–Jarrow–Merton factor model with rolling bonds as the basic instrument. Recall that, in the HJM factor model, the short rate can be represented as $\widehat{r}(t) = \varphi(t, Y(t))$, where Y is a diffusion on $\mathbb{R}^n$. This enables us to solve the problem by means of the general HJB theory. In addition to the general theory we provide calculations for the Vasicek, Cox–Ingersoll–Ross and G2++ models. The last mentioned model has never been considered in the context of the optimal portfolio selection problem. In our main result we show that the corresponding HJB equation has a solution, this solution admits a Feynman–Kac representation, and we can check the hypotheses of the general verification theorem. Finally, we provide formulae for the optimal consumption rate $\widehat{C}$ and the investment strategy $\widehat{\psi}$.

Now let us briefly recall the state of the literature and mark a few distinctive features of our approach. There are many papers in which the consumption-investment problem is solved under short rate dynamics and without taking care of other HJM models (see Wachter [35], Guasoni and Wang [19], [20], Korn and Kraft [24], Chang and Chang [9], Detemple and Rindisbacher [12], Deelstra et al. [11], Hata et al. [21], Synowiec [33], Trybuła [34]). Moreover, many of them focus on particular interest rate models such as the Vasicek model or the Cox–Ingersoll–Ross model. We consider the general factor model with rather weak assumptions on coefficients regularity, which covers many other short term interest rate models. We consider finite and infinite time horizon problems. It should be noticed that Guasoni and Wang [20] provided the solution for the infinite horizon incomplete market model and on domains for the stochastic factor being a subset of $\mathbb{R}^n$. Further, Hata et al. [21] considered the finite horizon problem with a general factor model with factor dynamics operating on the entire $\mathbb{R}^n$. Both papers use different arguments than we do and do not take into account the delicate nature of the bond market.

Ringer and Tehranchi [31] and Palczewski [29] considered the problem of maximizing the utility of the terminal wealth in great generality but without assuming any consumption stream and without introducing the concept of rolling bond.

Many authors have considered rolling bonds with various objectives, but in the framework limited either to a finite number of rolling bonds and very specific affine factor dynamics (see the risk sensitive criterion of Bielecki et al. [5], Bielecki and Pliska [6]) or to one rolling bond with one or two specific factors. Let us mention only a few articles dedicated to the pension management problem: Battochio and Menoncin [3], Guan and Liang [17], [18], Zhang et al. [38]. Apart from the general factor model we consider a general normalized signed measure as the investor's strategy, which can be useful when dealing with a changing number of bonds available on the market.

Munk and Sørensen [26] solved the consumption investment problem under the interest rate risk by duality methods (see Cox and Huang [10], Karatzas et al. [23]). Their solution brings some knowledge about the investment position in general HJM models. However, they did not use rolling bonds, did not consider the infinite horizon case and did not present any detailed solution in the affine framework.

It is worth mentioning that our analysis is conducted in the stochastic environment with unbounded coefficients, which allows us to develop many technical contributions. For example, we present a novel method to prove the verification theorem on the infinite horizon. This should be compared with the methods proposed for example by Nagai [27] and Guasoni and Wang [20].

The paper is organized as follows. In the next two sections we introduce the concept of a controlled investment process and an HJM factor model. The main result for the finite horizon case (Theorem 4.1) is formulated and proved in Sections 4 and 5, respectively. Sections 6 and 7 are devoted to the solution for some particular models. In Section 8 we study the infinite horizon case (see Theorem 8.4). In the Appendix we recall the basic definitions and properties of a bond market.

2. Controlled processes. Let $(\Omega, \mathfrak{F}, \mathbb{P})$ be a probability space, and let $T^* \in (0, +\infty)$ be the maximal time to maturity available on the market. Given a finite time investment horizon $T \in (0, +\infty)$, we denote by $\mathcal{M}_T$ the space of all (weakly) predictable processes ψ taking values in the space of all signed measures with finite total variation norm satisfying

$$\int_0^{T^*} \psi_t(\mathrm{d}x) = 1, \qquad \forall t \in [0, T],$$

$$\int_0^T \int_0^{T^*} |\langle \tilde{\sigma}(t)(x), \lambda(t)\rangle| \, \|\psi_t\|_{\mathrm{Var}}(\mathrm{d}x) \, \mathrm{d}t < +\infty,$$

$$\int_0^T \Big(\int_0^{T^*} \|\tilde{\sigma}(t)(x)\| \, \|\psi_t(\mathrm{d}x)\|_{\mathrm{Var}} \Big)^2 \, \mathrm{d}t < +\infty,$$

where $\|\cdot\|_{\mathrm{Var}}$ stands for the total variation norm.

Let $\mathcal{C}_T$ be the space of all non-negative predictable processes. We call $\mathcal{A}_T = \mathcal{M}_T \times \mathcal{C}_T$ the set of *admissible strategies*. Let $(\psi, C) \in \mathcal{A}_T$. Let $z^{\psi,C}(t)$ denote the capital of an investor whose consumption rate at time t is $C(t) z^{\psi,C}(t)$ and who can invest in (rolling) bonds with an investment strategy ψ. Then $z^{\psi,C}$ satisfies (1.3). Combining (1.2) and (1.3) we obtain

$$\frac{\mathrm{d}z^{\psi,C}(t)}{z^{\psi,C}(t)} = [\overline{r}(t) - C(t)] \, \mathrm{d}t - \int_0^{T^*} \langle \tilde{\sigma}(t)(x)\psi_t(\mathrm{d}x), \lambda(t) \, \mathrm{d}t + \mathrm{d}W(t)\rangle. \tag{2.1}$$

Note that the measure ψ_t admits negative values. In the particular case of

$$\psi_t = \sum_{i=1}^n [\eta_{1,t_i}\delta_{x_1} + \eta_{2,t_i}\delta_{x_2} + \cdots + \eta_{l_i,t_i}\delta_{x_{t_i}}] \chi_{(t_i,t_{i+1}]}(t), \tag{2.2}$$

where $(\eta_{1,t_i}, \eta_{2,t_i}, \ldots, \eta_{l_i,t_i})$ are random vectors such that $\sum_{k=1}^{l_i} \eta_{k,t_i} = 1$, each point mass measure δ_{x_k} corresponds to a rolling bond with fixed time to maturity x_k. Note that this setting reflects the fact that in practice on each time segment $(t_i, t_{i+1}]$ we have different numbers and types of maturities (bonds) available on the market. On the other hand, this is a convenient way to place different coupon bonds in the portfolio. Namely, we can take

$$\psi_t = \sum_{i=1}^{n} [\eta_{1,t_i}\psi_{t,1} + \eta_{2,t_i}\psi_{t,2} + \cdots + \eta_{l_i,t_i}\psi_{t,l_i}]\chi_{(t_i,t_{i+1}]}(t),$$

where $\psi_{t,1}, \psi_{t,2}, \ldots, \psi_{t,l_i}$ represent different coupon bonds and are of the form (2.2).

3. HJM factor model. In this paper we restrict our attention to the so-called HJM factor model. Namely, we assume that the short rate has the form $\overline{r}(t) = \varphi(t, Y(t))$, where

$$\text{(3.1)} \quad \mathrm{d}Y(t) = [B(t, Y(t)) + \Sigma(t, Y(t))\lambda(t, Y(t))]\, \mathrm{d}t + \Sigma(t, Y(t))\, \mathrm{d}W(t),$$

$B\colon [0, T] \times \mathbb{R}^n \to \mathbb{R}^n$, $\lambda\colon [0, T] \times \mathbb{R}^n \to \mathbb{R}^m$, $\varphi\colon [0, T] \times \mathbb{R}^n \to \mathbb{R}$, $\Sigma\colon [0, T] \times \mathbb{R}^n \to \mathbb{R}^n \otimes \mathbb{R}^m$, and $W = (W_1, \ldots, W_m)$ is an m-dimensional Wiener process defined on a filtered probability space $(\Omega, \mathfrak{F}, (\mathfrak{F}_t), \mathbb{P})$.

We assume that the hypotheses given below are satisfied. The first one concerns the regularity of B, λ, and φ. For the existence and uniqueness of a solution to SDE (3.1) we need at least local Lipschitz continuity in y; unfortunately, we also need at least Hölder continuity in t and y for the existence of a classical solution to certain PDEs with coefficients B, Σ, and φ, and their Feynman–Kac representations (see e.g. Friedman [15]).

Hypothesis 1. Either

(i) the mappings B, λ are bounded and locally Lipschitz continuous with respect to t and y, while φ is locally Lipschitz continuous in t and y and satisfies the linear growth condition in y uniformly in t, or
(ii) the mappings B, φ, λ are Lipschitz continuous in t and y.

The second hypothesis is on boundedness, Lipschitz continuity and uniform ellipticity of Σ. This is our most restrictive assumption. In particular, it does not allow us to cover directly the important Cox–Ingersoll–Ross model. In this case we treat the problem separately and present some explicit calculations together with other particular examples focused mainly on affine models.

Hypothesis 2. The mapping Σ is bounded, Lipschitz continuous in t and y, and there exists a constant $c > 0$ such that

$$\sum_{i,j=1}^{n} \left(\Sigma(t, y)\Sigma(t, y)^{\top}\right)_{i,j} \xi_i \xi_j \geq c\|\xi\|^2, \qquad \xi, y \in \mathbb{R}^n,\ t \in [0, T].$$

The last hypothesis is a no-arbitrage assumption.

HYPOTHESIS 3. $\lambda\colon [0,T] \times \mathbb{R}^n \to \mathbb{R}^n$ is a bounded, locally Lipschitz continuous mapping and

$$\mathrm{d}\mathbb{P}^* = \mathrm{e}^{\int_0^T \langle \lambda(t,Y(t)),\mathrm{d}W(t)\rangle - \frac{1}{2}\int_0^T \|\lambda(t,Y(t))\|^2\,\mathrm{d}t}\,\mathrm{d}\mathbb{P}$$

is a martingale measure (see e.g. Appendix).

REMARK 3.1. Hypothesis 3 gives the form of the volatility coefficient $\tilde\sigma$ appearing in (1.2). Indeed, the price $P(t,S)$ at time t of the zero coupon bond with maturity S is

$$P(t,S) = \mathbb{E}^*\{\mathrm{e}^{-\int_t^S \varphi(u,Y(u))\,\mathrm{d}u}|\mathfrak{F}_t\} = \mathbb{E}^*\{\mathrm{e}^{-\int_t^S \varphi(u,Y(u))\,\mathrm{d}u}|Y(t)\}.$$

Hence $P(t,S) = F(t,S,Y(t))$, where $F = F(t,S,y)$ is a continuous function of $t \in [0,T]$, $t \le S \le T^*$, and $y \in \mathbb{R}^n$. Moreover, for fixed S, $F(\cdot,S,\cdot) \in C^{1,2}([0,S]\times\mathbb{R}^n)$ is the unique solution to the equation

$$\frac{\partial}{\partial t}F(t,S,y) + L_0F(t,S,y) - \varphi(t,y)F(t,S,y) = 0, \qquad F(S,S,y) = 1, \tag{3.2}$$

where

$$L_0F(t,S,y) := \tfrac{1}{2}\operatorname{Tr}\big(\varSigma(t,y)\varSigma(t,y)^\top D^2_{yy}F(t,S,y)\big) + \langle B(t,y), D_yF(t,S,y)\rangle. \tag{3.3}$$

Therefore (see Appendix), the no-arbitrage condition yields

$$\tilde\sigma(t)(x) = -\varSigma(t,Y(t))^\top D_y \log F(t,t+x,Y(t)). \tag{3.4}$$

Given a signed measure ψ on $[0,T^*]$ set

$$A(\psi)(t,y) := \varSigma(t,y)^\top \int_0^{T^*} D_y \log F(t,t+x,y)\,\psi(\mathrm{d}x). \tag{3.5}$$

Let $\psi \in \mathcal{M}_T$. Taking into account (2.1), (3.4), and (3.5), we infer that in the HJM factor model the wealth dynamics has the form

$$\frac{\mathrm{d}z^{\psi,C}(t)}{z^{\psi,C}(t)} = [\varphi(t,Y(t)) - C(t)]\,\mathrm{d}t + \langle A(\psi_t)(t,Y(t)), \lambda(t,Y(t)\,\mathrm{d}t + \mathrm{d}W(t)\rangle. \tag{3.6}$$

In the present paper we consider finite and infinite horizon optimal consumption problems. In the finite horizon case the objective of the investor is to maximize

$$J_T(z,y,s,\psi,C) = \frac{1}{\alpha}\mathbb{E}\Big[a\int_s^T \mathrm{e}^{-\gamma(t-s)}(C(t)z^{\psi,C}(t))^\alpha\,\mathrm{d}t + b\,\mathrm{e}^{-\gamma(T-t)}(z^{\psi,C}(T))^\alpha\Big], \tag{3.7}$$

where z and y are the values of $z^{\psi,C}$ and Y at a given initial time $s \in [0,T]$, $\gamma \ge 0$ is a discount factor, $\alpha \in (0,1)$ and $a, b \ge 0$. Therefore, the investor can decide how to distribute his preferences between the consumption stream and the terminal weight by controlling the parameters a and b. In fact, some of our results hold true also for $\alpha < 0$ (see Remark 5.1). Let

$$V_T(z,y,s) := \sup_{(\psi,C)\in\mathcal{A}_T} J_T(z,y,s,\psi,C)$$

be the *value function*. Our main result concerning the finite horizon problem is Theorem 4.1 below.

In the infinite horizon case we assume that the functions B, Σ, λ and φ do not depend on the time variable. The reward functional is

$$J(z,y,\psi,C) = \frac{1}{\alpha}\mathbb{E}\int_0^{+\infty} \mathrm{e}^{-\gamma t}(C(t)z^{\psi,C}(t))^{\alpha}\,\mathrm{d}t. \tag{3.8}$$

Here z and y are the values of $z^{\psi,C}$ and Y at initial time 0. Our main results concerning the infinite horizon case are formulated and proved in Section 8.

4. Main result concerning the finite horizon problem. Recall that L_0 is defined by (3.3). Let

$$Lf(t,y) = L_0 f(t,y) + \langle \Sigma(t,y)\lambda(t,y), D_y f(t,y)\rangle \tag{4.1}$$

be the generator of the diffusion (3.1). Let

$$g(t,y) := \frac{1}{1-\alpha}\left[\alpha\varphi(t,y) + \frac{\alpha}{2(1-\alpha)}\|\lambda(t,y)\|^2 - \gamma\right]. \tag{4.2}$$

Consider the linear PDE

$$\frac{\partial G}{\partial t} + LG + \frac{\alpha}{1-\alpha}\langle \Sigma\lambda, D_y G\rangle + gG + a^{\frac{1}{1-\alpha}} = 0, \qquad G(T,y) = b^{\frac{1}{1-\alpha}}. \tag{4.3}$$

For the existence of an optimal strategy we will need the following hypothesis.

HYPOTHESIS 4. The fraction $\frac{D_y G(t,y)}{G(t,y)}$ is globally bounded and there is a weakly measurable mapping $\widehat{\psi}\colon [0,T]\times\mathbb{R}^n \to \mathcal{M}_T$ such that for all $t\in[0,T]$ and $y\in\mathbb{R}^n$,

$$A(\widehat{\psi}_{t,y})(t,y) = \frac{\lambda(t,y)}{1-\alpha} + \Sigma(t,y)^{\top}\frac{D_y G(t,y)}{G(t,y)}.$$

THEOREM 4.1.

(i) *Assume Hypotheses* 1 *to* 3. *Then there exists a unique classical solution* $G\in C^{1,2}[0,T)\times\mathbb{R}^n)\cap C([0,T]\times\mathbb{R}^n)$ *of* (4.3) *satisfying the exponential growth condition* $|G(t,y)| \le A\,\mathrm{e}^{B\|y\|}$. *Moreover, G admits the Feynman–Kac representation*

$$G(t,y) = \mathbb{E}\,a^{\frac{1}{1-\alpha}}\int_t^T \mathrm{e}^{\int_t^u g(k,\widetilde{Y}(k))\,\mathrm{d}k}\,\mathrm{d}u + \mathbb{E}\,b^{\frac{1}{1-\alpha}}\,\mathrm{e}^{\int_t^T g(k,\widetilde{Y}(k))\,\mathrm{d}k}, \tag{4.4}$$

where g is given by (4.2) *and* $\widetilde{Y}$ *solves*

$$
\begin{aligned}
\mathrm{d}\widetilde{Y}(k) &= \left[B(k,\widetilde{Y}(k)) + \frac{1}{1-\alpha}\Sigma(k,\widetilde{Y}(k))\lambda(k,\widetilde{Y}(k))\right]\mathrm{d}k \\
&\quad + \Sigma(k,\widetilde{Y}(k))\,\mathrm{d}W(k), \\
\widetilde{Y}(t) &= y.
\end{aligned}
\tag{4.5}
$$

(ii) *Assume that additionally Hypothesis* 4 *holds. Then*

$$V_T(z,y,s) = G(s,y)^{1-\alpha}\mathrm{e}^{-\gamma s}z^{\alpha} = J_T(z,y,s,\widehat{\psi},\widehat{C}),$$

where the optimal investment policy $\widehat{\psi}$ *and the optimal consumption* $\widehat{C}$ *are given by* $\widehat{\psi}_t = \widehat{\psi}_{t,Y(t)}$ *and* $\widehat{C}(t) = G(t,Y(t))^{-1}$ *for* $t \in [s,T]$.

Theorem 4.1 is proved in Section 5.

4.1. Some auxiliary results

REMARK 4.2. Given t and y, consider the following equation for a signed measure ψ:

$$A(\psi)(t,y) = \frac{\lambda(t,y)}{1-\alpha} + \Sigma(t,y)^{\top}\frac{D_yG(t,y)}{G(t,y)}, \qquad \int_0^{T^*}\psi(\mathrm{d}x) = 1. \tag{4.6}$$

This equation appears in Hypothesis 4. The existence of its solution is crucial for the existence of an optimal investment strategy. Note that the right hand side of the first equation of (4.6) is a vector in $\mathbb{R}^m$. Let $\mathbf{x} = (x_1,\ldots,x_l) \in [0,T^*]^l$, where l is large enough. We are looking for ψ of the form

$$\psi = \sum_{k=1}^{l}\eta_k(t,y)\delta_{x_k},$$

where $\eta(t,y) = (\eta_1(t,y),\ldots,\eta_l(t,y))^{\top} \in \mathbb{R}^l$. Since $A(\psi)(t,y)$ is given by (3.5) we have the following system of linear equations for the column vector $\eta(t,y)$:

$$
\begin{cases}
\mathcal{F}(t,y,\mathbf{x})\eta(t,y) = \dfrac{\lambda(t,y)}{1-\alpha} + \Sigma(t,y)^{\top}\dfrac{D_yG(t,y)}{G(t,y)}, \\
\displaystyle\sum_{k=1}^{l}\eta_k(t,y) = 1,
\end{cases}
$$

where $\mathcal{F}(t,y,\mathbf{x})$ is the $m\times l$ matrix with columns

$$\Sigma(t,y)^{\top}D_y\log F(t,t+x_1,y),\ldots,\Sigma(t,y)^{\top}D_y\log F(t,t+x_{m+1},y).$$

Here the derivative $D_y\log F(t,t+x_k,y)$ is understood as a column vector. Let

$$e = (1,\ldots,1).$$

A solution exists provided that for the sequence $\mathbf{x}$ the $(m+1) \times l$ matrix

$$\overline{\mathcal{F}(t,y,\mathbf{x})} = \begin{bmatrix} \mathcal{F}(t,y,\mathbf{x}) \\ e \end{bmatrix}$$

has rank $m+1$. Summing up, if there are l and a vector $\mathbf{x} \in \mathbb{R}^l$ such that for all t and y, $\overline{\mathcal{F}(t,y,\mathbf{x})}$ has rank $m+1$, then condition (4.6) is fulfilled. Moreover, one can choose an optimal investment strategy in the form

$$\widehat{\psi}_t = \sum_{k=1}^{l} \eta_k(t, Y(t))\delta_{x_k}.$$

REMARK 4.3. Note that (3.5) has much in common with the standard *duration* of the portfolio of bonds used frequently in static bond portfolio immunization. So we might say that Theorem 4.1 gives a recipe for dynamic portfolio immunization.

5. Proof of Theorem 4.1. First of all, note that (4.3) is a linear equation and under Hypotheses 1 and 3 it has a unique smooth classical solution in $C^{1,2}([0,T) \times \mathbb{R}^n \times [0,T)) \cap C([0,T] \times \mathbb{R}^n)$ (see Zawisza [36, Theorem 3.3]) such that $|G(y,t)| \le A\,\mathrm{e}^{B\|y\|}$ (the latter follows easily from Zawisza [36, Lemma 3.2 and proof of Theorem 3.3]). Moreover, G admits the Feynman–Kac representation (4.4), (4.5).

Assume that Hypothesis 4 is fulfilled. To solve the optimization problem we will use the HJB approach. As usual we will try to find the function V in the form

$$V(z,y,s) = \frac{1}{\alpha} K(s,y)\,\mathrm{e}^{-\gamma s} z^{\alpha}.$$

Recall that $\alpha \in (0,1)$; for $\alpha < 0$ one would need to exchange sup with inf in the HJB equation. Let us write the HJB equations for the function K:

$$\frac{\partial K}{\partial t} + LK + (\alpha\varphi - \gamma)K + \sup_{c \ge 0}[-\alpha K c + a c^{\alpha}] \\ + \alpha \sup_{\psi} \left\{ \frac{\|A(\psi)\|^2(\alpha - 1)}{2} K + \langle A(\psi), \lambda K + \Sigma^{\top} D_y K \rangle \right\} = 0,$$

with the terminal condition $K(T,y) = b$.

Note that

$$\sup_{c \ge 0}[-\alpha K c + a c^{\alpha}] = a(1-\alpha)\left(\frac{K}{a}\right)^{\frac{\alpha}{\alpha-1}}$$

and the supremum is attained at $\widehat{c} = (K/a)^{\frac{1}{\alpha-1}}$.

Next note that

$$\alpha \sup_{A\in\mathbb{R}^m}\left\{\frac{\|A\|^2(\alpha-1)}{2}K+\langle A,\lambda K+\Sigma^\top D_yK\rangle\right\}$$
$$=\frac{\alpha}{2(1-\alpha)}\frac{1}{K}\|\lambda K+\Sigma^\top D_yK\|^2$$

and the supremum equals

$$\overline{A}:=\frac{1}{1-\alpha}\left(\lambda+\Sigma^\top\frac{D_yK}{K}\right).$$

We will show that $K(t,y)^{1-\alpha}=G(t,y)$, therefore Hypothesis 4 ensures that given t and y there is a signed measure $\widehat{\psi}_t(\mathrm{d}x)(y)$ such that

$$A(\widehat{\psi}_t(\cdot)(y))=\frac{1}{1-\alpha}\left(\lambda(t,y)+\Sigma(t,y)^\top\frac{D_yK(t,y)}{K(t,y)}\right),\qquad \int_0^{T^*}\widehat{\psi}_t(\mathrm{d}x)(y)=1.$$

Hence, we eventually arrive at the HJB equation

$$\begin{aligned}0&=\frac{\partial K}{\partial t}+LK+(\alpha\varphi-\gamma)K+a(1-\alpha)\left(\frac{K}{a}\right)^{\frac{\alpha}{\alpha-1}}\\&\quad+\frac{\alpha}{2(1-\alpha)}\frac{1}{K}\|\lambda K+\Sigma^\top D_yK\|^2,\\b&=K(T,y).\end{aligned}$$

The proof of the theorem will be completed as soon as we show that:

- $G(t,y):=K(t,y)^{1-\alpha}$ satisfies (4.4).
- We have
$$\frac{D_yK}{K}=(1-\alpha)\frac{D_yG}{G}.$$
- We can conduct the verification reasoning for the function $\mathrm{e}^{-\gamma t}K(t,y)z^\alpha$.

An elementary verification of the first two items is left to the reader. Hypothesis 4 guarantees boundedness of D_yG/G. Therefore, under Hypotheses 1–3, the optimal state process can be rewritten as

$$z^{\hat{\psi},\hat{C}}(t)=z\,\mathrm{e}^{\int_s^t h(\widetilde{Y}(u))\,\mathrm{d}u}Z(t),$$

where h satisfies the linear growth condition, while Z is a square integrable martingale. Next, we can use the fact that

$$\mathbb{E}\sup_{s\le t\le T}\mathrm{e}^{A\|\widetilde{Y}(t)\|}<+\infty \tag{5.1}$$

(see Zawisza [36, Lemma 3.2]), which guarantees the uniform integrability condition for a certain family of random variables and ensures that we can use the verification theorem to prove that $(\widehat{\psi},\widehat{C})$ is an optimal control. More precisely, applying the Itô formula and taking the expectation of both sides, we get

$$\mathbb{E}\,\mathrm{e}^{-\gamma(S\wedge\tau_n-t)}V\big(z^{\widehat{\psi},\widehat{C}}(S\wedge\tau_n),Y(S\wedge\tau_n),S\wedge\tau_n\big)$$
$$=V(z,y,t)-\frac{1}{\alpha}\mathbb{E}\,a\int_t^{S\wedge\tau_n}\mathrm{e}^{-\gamma(s-t)}(\widehat{C}(s)z^{\widehat{\psi},\widehat{C}}(s))^{\alpha}\,\mathrm{d}s,$$

where $(\tau_n,n\in\mathbb{N})$ is a localizing sequence of stopping times and $S<T$ is a positive constant.

Condition (5.1) justifies passing to the limit under the expectation sign. Eventually, we arrive at

$$V(z,y,t)=\frac{1}{\alpha}\mathbb{E}\Big[a\int_t^T\mathrm{e}^{-\gamma(s-t)}(\widehat{C}(s)z^{\widehat{\psi},\widehat{C}}(s))^{\alpha}\,\mathrm{d}s+b\,\mathrm{e}^{-\gamma(T-t)}(z^{\widehat{\psi},\widehat{C}}(T))^{\alpha}\Big].$$

Next, we need to show that the value function V dominates the value for other admissible strategies. Repeating the procedure for any admissible strategy (ψ,C), we get

$$\mathbb{E}\,\mathrm{e}^{-\gamma(S\wedge\tau_n-t)}V\big(z^{\psi,C}(S\wedge\tau_n),Y(S\wedge\tau_n),(S\wedge\tau_n)\big)$$
$$\leq V(z,y,t)-\frac{1}{\alpha}\mathbb{E}\,a\int_t^{S\wedge\tau_n}\mathrm{e}^{-\gamma(s-t)}(C(s)z^{\psi,C}(s))^{\alpha}\,\mathrm{d}s.$$

By the positivity of α, we can use the Fatou lemma to obtain

$$V(z,y,t)\geq\frac{1}{\alpha}\mathbb{E}\Big[a\int_t^T\mathrm{e}^{-\gamma(s-t)}(C(s)z^{\psi,C}(s))^{\alpha}\,\mathrm{d}s+b\,\mathrm{e}^{-\gamma(T-t)}(z^{\psi,C}(T))^{\alpha}\Big].\quad\blacksquare \tag{5.2}$$

REMARK 5.1. If $\alpha<0$, then Theorem 4.1(i) still holds true. We do not know how to show that the value function V dominates the value for other admissible strategies. The main problem is to show the convergence of the term

$$\mathbb{E}\,\mathrm{e}^{-\gamma(S\wedge\tau_n-t)}V\big(z^{\psi,C}(S\wedge\tau_n),Y(S\wedge\tau_n),S\wedge\tau_n\big).$$

REMARK 5.2. Assume that we additionally have the possibility to allocate our resources in the stock market $S(t)\in\mathbb{R}^N$, with dynamics

$$\mathrm{d}S(t)=\operatorname{diag}(S(t))[\overline{r}(t)+\Sigma_S(t,Y(t))\lambda(t,Y(t))]\,\mathrm{d}t+\Sigma_S(t,Y(t))\,\mathrm{d}W(t).$$

Then an investment policy is a pair (ψ,π), where ψ is a signed measure on $[0,T^*]$ and $\pi\in\mathbb{R}^N$ with

$$\int_0^{T^*}\psi(\mathrm{d}x)+\sum_{j=1}^N\pi_j=1.$$

The optimal investment policy $(\widehat{\psi}_t, \widehat{\pi}_t)$, $t \in [0,T]$, should solve the system

$$A(\widehat{\psi}_t)(t, Y(t)) + \Sigma_S(t, Y(t))^\top \widehat{\pi}_t = \frac{\lambda(t, Y(t))}{1-\alpha} + \Sigma(t, Y(t))^\top \frac{D_y G(t, Y(t))}{G(t, Y(t))}.$$

The existence of a sequence $\mathbf{x} = (x_1, \ldots, x_l)$ such that the matrix

$$\begin{bmatrix} \mathcal{F}(t, y, \mathbf{x}), \Sigma_S(t, y)^\top \\ e \end{bmatrix}$$

has rank $m+1$ ensures the existence of an optimal control $(\widehat{\psi}, \widehat{\pi})$ with $\widehat{\psi}_t$ being a point measure. We are aware that the processes S and Y share the same Wiener process and this restricts generality, but the analysis of the space with another independent Wiener process in the dynamics of S or Y is out of the scope of this paper. Partial results for the general problem can be found for example in Hata et al. [21] and Zawisza [37]. But our formulation is sufficient for example to cover the bond-stock mix problem of Brennan and Xia [8].

6. Examples. Our assumptions allow us to consider the following important models with practical implementations. Detailed calculations for specific affine examples of the models presented below are given in the next section.

EXAMPLE 6.1 (Consistent HJM models). We assume that the forward rate in the Musiela parametrization is given by $r(t)(x) = \phi(x, Y(t))$, where $\phi \in C^{1,2}([0, T^*] \times \mathbb{R}^n)$, and Y is given by (3.1). Note that for the short rate we have $\overline{r}(t) = \phi(0, Y(t))$. Then (see Filipović [13, Proposition 9.1]) in the case of B and Σ independent of t, under Hypotheses 1 and 2, the consistency condition in Hypothesis 3 holds if and only if $\Phi(x, y) = \int_0^x \phi(u, y)\, \mathrm{d}u$ satisfies

$$\frac{\partial \Phi}{\partial x}(x, y) = \phi(0, y) + \langle B(y), D_y \Phi(x, y)\rangle + \frac{1}{2} \sum_{i,j=1}^{n} (\Sigma(y)\Sigma(y)^\top)_{i,j} \left[\frac{\partial^2 \Phi}{\partial y_i \partial y_j}(x, y) - \frac{\partial \Phi}{\partial y_i}(x, y)\frac{\partial \Phi}{\partial y_j}(x, y)\right].$$

EXAMPLE 6.2 (Short term interest rate models). Assume that

$$\mathrm{d}\overline{r}(t) = B(t, \overline{r}(t))\, \mathrm{d}t + \Sigma(t, \overline{r}(t))[\lambda(t)\, \mathrm{d}t + \mathrm{d}W(t)],$$

where B and Σ are functions satisfying Hypotheses 1 and 2 and λ is a deterministic bounded measurable function. Then, in the framework of the HJM factor model, we take $Y = \overline{r}$.

EXAMPLE 6.3 (Gaussian-separable HJM model, also known as Ritchken–Sankarasubramanian model). Assume that the volatility has the form $\sigma(t)(x) = \beta(t)\nu(x+t)$, where β and ν are deterministic functions. In this case, assuming some mild regularity of the function ν and $f(0,t)$, we have

$$\mathrm{d}\overline{r}(t) = \left[\frac{\partial f(0,t)}{\partial t} - f(0,t)\frac{\nu'(t)}{\nu(t)} + \int_0^t \beta(u)\nu(u)\,\mathrm{d}u + \overline{r}(t)\frac{\nu'(t)}{\nu(t)}\right]\mathrm{d}t + \beta(t)\nu(t)\,\mathrm{d}W^*(t).$$

Thus the problem can be easily reduced to short rate models. This can be further generalized to the so-called Cheyette models

$$\sigma(t)(x) = \sum_{i=1}^{N} \beta_i(t)\frac{\nu_i(t)}{\nu_i(x+t)}.$$

For more details we refer to Beyna [4].

EXAMPLE 6.4 (Quasi-Gaussian HJM model). Let

$$\sigma(t)(x) = \beta(t,\xi(t))\nu(x+t),$$

where $\xi(t)$ is a diffusion and ν a deterministic function. Then

$$\mathrm{d}\overline{r}(t) = \left[\frac{\partial f(0,t)}{\partial t} - f(0,t)\frac{\nu'(t)}{\nu(t)} + \int_0^t \beta(u,\xi(u))\nu(u)\,\mathrm{d}u + \overline{r}(t)\frac{\nu'(t)}{\nu(t)}\right]\mathrm{d}t + \beta(t,\xi(t))\nu(t)\,\mathrm{d}W^*(t).$$

Clearly, for $\psi(r,\xi) = r$ and $Y = (\overline{r},\xi)$ we have $\overline{r}(t) = \psi(t,Y(t))$. Unfortunately, the strong ellipticity condition cannot be satisfied. In order to have a non-degenerate diffusion one can replace the term $\int_0^t \beta(u,\xi(u))\nu(u)\,\mathrm{d}u\,\mathrm{d}t$ by its ε-perturbation

$$\int_0^t \beta(u,\xi(u))\nu(u)\,\mathrm{d}u\,\mathrm{d}t + \varepsilon\,\mathrm{d}\tilde{W}(t),$$

where $\tilde{W}$ is an independent Wiener process.

An example of such a model is the Cheyette HJM model with β having the affine structure $\beta(t,\xi) := \zeta_1(t)+\zeta_2(t)\xi$, where ζ_1 and ζ_2 are deterministic functions. Note that in affine models (for the definition see the next section) taking $\varepsilon \to 0$ is an instantaneous operation and does not need separate justification. For more information about such models we refer to Pirjol and Zhu [30].

7. Affine factor models. Our aim here is to present examples of models which admit explicit solutions. As in Section 4, we assume that $\overline{r}(t) = \varphi(t,Y(t))$, where Y is given by (3.1). We focus on affine models, i.e. models with

$$\begin{aligned} B(t,y) &:= B_1(t) + \langle B_2(t), y\rangle, \qquad \Sigma(t,y) := \Sigma(t), \\ \lambda(t,y) &:= \lambda(t), \qquad \varphi(t,y) = \varphi_1(t) + \langle \varphi_2(t), y\rangle, \end{aligned} \tag{7.1}$$

where $B_1(t)$, $B_2(t)$, $\Sigma(t)$, $\lambda(t)$, $\varphi_1(t)$, $\varphi_2(t)$ are deterministic matrix-valued mappings. At the end of the present section we will consider the CIR model, which is not in fact an affine factor model.

Proposition 7.1. *Under the affine specification* (7.1) *the formula for the optimal pair* $(\widehat{\psi}, \widehat{C})$ *reads*

$$\Sigma(t)^\top \int_0^{T^*} \Big[\int_t^{t+x} P_{t,t+x} P_{k,t+x}^{-1} \varphi_2(k)\,\mathrm{d}k \Big]\, \psi_t(\mathrm{d}x) = \frac{\lambda(t)}{1-\alpha} + \Sigma(t)^\top \frac{D_y G(t, Y(t))}{G(t, Y(t)},$$

and $C(t) = G(t, Y(t))^{-1}$, *where*

$$\begin{aligned} G(t,y) &:= a^{\frac{1}{1-\alpha}} \int_t^T \eta(t,u,y)\,\mathrm{d}u + b^{\frac{1}{1-\alpha}} \eta(t,T,y), \\ \eta(t,u,y) &:= \mathrm{e}^{m_{1,t,u} + \langle m_{2,t,u}, y\rangle + \frac{1}{2}\sigma^2_{t,u}}, \end{aligned} \tag{7.2}$$

$$\begin{aligned} m_{2,t,u} &:= \frac{\alpha}{1-\alpha} \int_t^u P_{t,u} P_{k,u}^{-1} \varphi_2(k)\,\mathrm{d}k, \\ m_{1,t,u} + \frac{1}{2}\sigma^2_{t,u} &= \int_t^u f(k,u)\,\mathrm{d}k, \end{aligned} \tag{7.3}$$

$$\begin{aligned} f(t,u) := {}& \frac{1}{2}\langle m_{2,t,u}, \Sigma(t)\Sigma(t)^\top m_{2,t,u}\rangle + \langle B_1(t), m_{2,t,u}\rangle \\ &+ \frac{1}{1-\alpha}\Big[\langle m_{2,t,u}, \Sigma(t)\lambda(t)\rangle + \frac{\alpha}{2(1-\alpha)}\|\lambda(t)\|^2 - \gamma\Big] \end{aligned} \tag{7.4}$$

and $P_{t,k}$ *denotes the time-ordered path exponential of the matrix* $B(t)$.

Proof. Note that in the affine framework, g given by (4.2) is equal to

$$\begin{aligned} g(t,y) &= \frac{1}{1-\alpha}\Big[\alpha\varphi_1(t) + \frac{\alpha}{2(\alpha-1)}\|\lambda(t)\|^2 - \gamma\Big] + \frac{\alpha}{1-\alpha}\langle \varphi_2(t), y\rangle \\ &=: g_0(t) + \langle g_1(t), y\rangle. \end{aligned}$$

Moreover, the process $\widetilde{Y}$ appearing in the Feynman–Kac representation (4.4), (4.5) of the function G is Gaussian. Hence, for any $t \le u$, the random variable $\int_t^u g(k, \widetilde{Y}(k))\,\mathrm{d}k$ has a Gaussian $\mathcal{N}(m_{1,t,u} + \langle m_{2,t,s}, y\rangle, \sigma_{t,u})$ distribution, and consequently we have (7.2).

On the other hand, substituting the above function G into (4.3), we infer that the function η satisfies

$$\begin{aligned} -\frac{\partial \eta}{\partial t} + \frac{1}{2}\operatorname{Tr}(\Sigma\Sigma^\top D^2_{yy}\eta) + \langle B_1 + B_2 y, D_y\eta\rangle \\ + \frac{1}{1-\alpha}\langle \Sigma\lambda, D_y\eta\rangle + g_0 + \langle g_1, y\rangle = 0 \end{aligned}$$

with $\eta(t,t,y)=1$. Consequently,

$$\frac{\partial}{\partial t} m_{2,t,u} = B_2(t) m_{2,t,u} + \frac{\alpha}{1-\alpha}\varphi_2(t), \qquad m_{2,t,t}=0,$$

and therefore

$$m_{2,t,u} = \frac{\alpha}{1-\alpha}\int_t^u P_{t,u}P_{k,u}^{-1}\varphi_2(k)\,\mathrm{d}k,$$

where $P_{t,k}$ denotes the time-ordered path exponential of the matrix $B(t)$. For the function $m_{1,t,s}+\frac{1}{2}\sigma_{t,s}^2$ we have

$$\frac{\partial}{\partial t}\left[m_{1,t,u}+\frac{1}{2}\sigma_{t,u}^2\right] = f(t,u), \qquad m_{1,t,t}+\frac{1}{2}\sigma_{t,t}^2 = 0,$$

where $f(t,u)$ equals

$$\frac{1}{2}\langle m_{2,t,u}, \Sigma(t)\Sigma(t)^\top m_{2,t,u}\rangle + \langle B_1(t), m_{2,t,u}\rangle + \frac{1}{1-\alpha}\left[\langle m_{2,t,u}, \Sigma(t)\lambda(t)\rangle + \frac{\alpha}{2(1-\alpha)}\|\lambda(t)\|^2 - \gamma\right].$$

Hence

$$m_{1,t,u}+\frac{1}{2}\sigma_{t,u}^2 = \int_t^u f(k,u)\,\mathrm{d}k. \ \blacksquare$$

REMARK 7.2. It is worth stressing that in the investment-consumption model under the affine factor assumption, the optimal portfolio weights are not linear combinations of factors. This contrasts with the pure investment problem (see Bielecki and Pliska [6]).

7.1. Short rate affine models. Now, let us consider more specific examples. The model is *short rate affine* if the price $P(t,S)$ at time t of the zero coupon bond with maturity T is

$$P(t,T) = \mathrm{e}^{m(T-t)-n(T-t)\overline{r}(t)}, \tag{7.5}$$

where m and n are deterministic functions. Moreover, it is assumed that the short rate $\overline{r}$ is a diffusion process,

$$\mathrm{d}\overline{r}(t) = B(\overline{r}(t))\,\mathrm{d}t + \Sigma(\overline{r}(t))[\lambda(\overline{r}(t))\,\mathrm{d}t + \mathrm{d}W(t)].$$

In the notation of Section 4, $Y=\overline{r}$. We assume that B, Σ and λ satisfy Hypotheses 1 to 3. Barski and Zabczyk [2] have proved that in affine models, bond prices are local martingales with respect to martingale measures if and only if $B(r)=a+br$ and ($\Sigma(r)=c\sqrt{r}$ or $\Sigma(r)=c$), where $a,b,c\in\mathbb{R}$ are constants. Obviously, (7.5) requires additional assumptions on λ. Under the Musiela parametrization (see Appendix), we have

$$P(t)(x) = \mathrm{e}^{m(x)-n(x)\overline{r}(t)} = F(t,t+x,\overline{r}(t)), \qquad F(t,t+x,r) = \mathrm{e}^{m(x)-n(x)r},$$

and

$$r(t)(x) = -m'(x) + n'(x)\overline{r}(t).$$

Since

$$\frac{\partial}{\partial r}\log F(t, t+x, r) = -n(x)$$

the formula for the optimal investment strategy reads

$$\int_0^{T^*} n(x)\,\widehat{\psi}_t(\mathrm{d}x) = \frac{1}{\alpha-1}\frac{\lambda(\overline{r}(t))}{\Sigma(\overline{r}(t))} - \frac{D_r G(t,\overline{r}(t))}{G(t,\overline{r}(t))}. \tag{7.6}$$

The optimal consumption rate is given by

$$\widehat{C}(t) = G(t,\overline{r}(t))^{-1}. \tag{7.7}$$

7.1.1. *Vasicek model.* First we will consider the *Vasicek model*

$$\mathrm{d}\overline{r}(t) = [\beta - \kappa\overline{r}(t)]\,\mathrm{d}t + \sigma[\lambda\,\mathrm{d}t + \mathrm{d}W(t)], \tag{7.8}$$

where b, κ, λ, and $\sigma > 0$ are constants. Note that under the Vasicek model we have

$$n(x) = \frac{1-\mathrm{e}^{-\kappa x}}{\kappa}, \qquad m(x) = -\beta\int_0^x n(y)\,\mathrm{d}y + \frac{\sigma^2}{2}\int_0^x n(u)^2\,\mathrm{d}u,$$

and

$$\frac{\mathrm{d}U(t)(x)}{U(t)(x)} = \overline{r}(t)\,\mathrm{d}t - \sigma n(x)[\lambda\,\mathrm{d}t + \mathrm{d}W(t)].$$

PROPOSITION 7.3. *In the Vasicek model* (7.8) *the optimal pair* $(\widehat{\psi}, \widehat{C})$ *is given by* (7.6)–(7.7) *where*

$$G(t,r) = a^{\frac{\alpha}{1-\alpha}}\int_t^T \mathrm{e}^{m_{1,t,u}+m_{2,t,u}r+\frac{1}{2}\sigma^2_{t,u}}\,\mathrm{d}u + b^{\frac{\alpha}{1-\alpha}}\,\mathrm{e}^{m_{1,t,T}+m_{2,t,T}r+\frac{1}{2}\sigma^2_{t,T}}, \tag{7.9}$$

and

$$\sigma^2_{t,u} := \overline{\alpha}^2\int_t^u n(k)^2\sigma^2\,\mathrm{d}k,$$

$$m_{1,t,u} := \overline{\alpha}\int_t^u\left[n(k)\left(\beta + \frac{1}{(1-\alpha)}\lambda\sigma\right) + \frac{1}{2(1-\alpha)}\lambda^2 - \gamma\right]\mathrm{d}k,$$

$$m_{2,t,u} := \overline{\alpha}\,n(u-t), \qquad \overline{\alpha} := \alpha/(1-\alpha).$$

Proof. Note that equation (4.3) for G now has the form

$$\frac{\partial G}{\partial t}(t,r) + \tilde{L}G(t,r) + g(r)G(t,r) + a^{\frac{1}{1-\alpha}} = 0, \qquad G(T,y) = b^{\frac{1}{1-\alpha}},$$

where

$$\tilde{L}G(t,r) = \frac{\sigma^2}{2}\frac{\partial^2 G}{\partial r^2}(t,r) + \left[\beta - \kappa r + \frac{\alpha}{1-\alpha}\sigma\lambda\right]\frac{\partial G}{\partial r}(t,r)$$

and

$$g(r) = \frac{1}{1-\alpha}\left[\alpha r + \frac{\alpha}{2(1-\alpha)}\lambda^2 - \gamma\right].$$

Thus

$$G(t,r) = \mathbb{E}\left\{a^{\frac{1}{1-\alpha}}\int_t^T \mathrm{e}^{\int_t^u g(\widetilde{r}(k))\,\mathrm{d}k}\,\mathrm{d}u + b^{\frac{1}{1-\alpha}}\mathrm{e}^{\int_t^T g(\widetilde{r}(k))\,\mathrm{d}k}\right\},$$

where

$$\mathrm{d}\widetilde{r}(k) = \left[\beta - \kappa\widetilde{r}(k) + \frac{1}{1-\alpha}\lambda\sigma\right]\mathrm{d}k + \sigma\,\mathrm{d}W(k), \qquad \widetilde{r}(t) = r.$$

Note that

$$\int_t^u \widetilde{r}(k)\,\mathrm{d}k = n(u-t)r + \int_t^u n(k)\left[\beta + \frac{1}{1-\alpha}\lambda\sigma\right]\mathrm{d}k + \int_t^u n(k)\sigma\,\mathrm{d}W(k).$$

Therefore, the integral $\int_t^u \widetilde{r}(k)\,\mathrm{d}k$ is normally distributed, and in this case G is given by (7.9). ∎

Note that the optimal investment is determined by the condition

$$\int_0^{T^*} \frac{1-\mathrm{e}^{-\kappa x}}{\kappa}\,\widehat{\psi}_t(\mathrm{d}x) = \frac{\lambda}{(\alpha-1)\sigma} - \frac{D_r G(t,\overline{r}(t))}{G(t,\overline{r}(t))}, \qquad \int_0^{T^*} \widehat{\psi}_t(\mathrm{d}x) = 1.$$

In particular, we can take $\widehat{\psi}_t(\mathrm{d}x) = \widehat{\eta}_0(t)\,\delta_0(\mathrm{d}x) + \widehat{\eta}_{\overline{x}}(t)\,\delta_{\overline{x}}(\mathrm{d}x)$, where $\overline{x} \in (0,T^*]$ is an arbitrary fixed time to maturity. The process $(\widehat{\eta}_0(t), \widehat{\eta}_{\overline{x}}(t))$ is determined by

$$\widehat{\eta}_{\overline{x}}(t) = \frac{\kappa}{1-\mathrm{e}^{-\kappa\overline{x}}}\left[\frac{\lambda}{(\alpha-1)\sigma} - \frac{D_r G(t,\overline{r}(t))}{G(t,\overline{r}(t))}\right], \qquad \widehat{\eta}_0(t) = 1 - \widehat{\eta}_{\overline{x}}(t).$$

It should be noted that the solution for the simpler Merton model

$$\mathrm{d}\overline{r}(t) = \beta\,\mathrm{d}t + \sigma(\lambda\,\mathrm{d}t + \mathrm{d}W(t))$$

can be derived by letting $\kappa \to 0$ in the Vasicek model. Thus, in the Merton model, $n(x) = x$ and $m(x) = -\beta x^2 + (\sigma^2/6)x^3$. In particular, the condition for the optimal investment is

$$\int_0^{T^*} x\,\widehat{\psi}_t(\mathrm{d}x) = \frac{\lambda}{(\alpha-1)\sigma} - \frac{D_r G(t,\overline{r}(t))}{G(t,\overline{r}(t))}, \qquad \int_0^{T^*} \widehat{\psi}_t(\mathrm{d}x) = 1.$$

7.1.2. *CIR model.* Another important model worth considering is the Cox–Ingersoll–Ross model. It does not satisfy our Hypotheses 1 and 2. However, we are able to perform some explicit calculations. In fact, we will derive an exact formula for the solution G to (4.3). In this way we obtain a candidate for the value function for the corresponding control problem. We perform calculations only for the case $\alpha < 0$ (see Remark 5.1). For other parameters it might happen that the value function has infinite value (see for example

Korn and Kraft [25, Proposition 3.2]). We start by listing elementary facts about the CIR model:

$$\begin{aligned}
\mathrm{d}\overline{r}(t) &= (\beta - \kappa\overline{r}(t))\,\mathrm{d}t + \sigma\sqrt{\overline{r}(t)}(\lambda\,\mathrm{d}t + \mathrm{d}W(t)),\\
n(x) &= \frac{\sinh\rho x}{\rho\cosh\rho x + \frac{\kappa}{2}\sinh\rho x},\\
m(x) &= \frac{2\beta}{\sigma}\log\bigg(\frac{\mathrm{e}^{\kappa x/2}}{\rho\cosh\rho x + \frac{\kappa}{2}\sinh\rho x}\bigg),\\
\rho &= \frac{1}{2}(\kappa^2 + 2\sigma^2)^{1/2},\\
\frac{\mathrm{d}U(t)(x)}{U(t)(x)} &= \overline{r}(t)\,\mathrm{d}t - \sigma\sqrt{\overline{r}(t)}\,n(x)(\lambda\,\mathrm{d}t + \mathrm{d}W(t)).
\end{aligned}$$

Usually it is assumed that $2\beta \geq \sigma^2$. To obtain a closed form solution we assume here $\lambda(r) := \lambda\sqrt{r}$.

Note that

$$\mathrm{d}\widetilde{r}(k) = [\beta - \widetilde{\kappa}\widetilde{r}(k)]\,\mathrm{d}k + \sigma\sqrt{\widetilde{r}(k)}\,\mathrm{d}W(k), \qquad \widetilde{r}(t) = r, \qquad \widetilde{\kappa} = \kappa - \frac{\lambda\sigma}{1-\alpha}$$

and

$$\mathbb{E}\,\mathrm{e}^{\int_t^u \frac{\alpha}{1-\alpha}[1+\frac{\lambda^2}{2(1-\alpha)}]\widetilde{r}(k)\,\mathrm{d}k} = \mathrm{e}^{\widetilde{m}(u-t)-\widetilde{n}(u-t)r},$$

where

$$\begin{aligned}
\widetilde{n}(x) &= \frac{|\alpha|}{1-\alpha}\bigg[1 + \frac{\lambda^2}{2(1-\alpha)}\bigg]\frac{\sinh\widetilde{\gamma}x}{\widetilde{\gamma}\cosh\widetilde{\gamma}x + \frac{\widetilde{\kappa}}{2}\sinh\widetilde{\gamma}x},\\
\widetilde{m}(x) &= \sqrt{\frac{|\alpha|}{(1-\alpha)}\bigg[1 + \frac{\lambda^2}{2(1-\alpha)}\bigg]}\,\frac{2\beta}{\sigma}\log\bigg(\frac{\mathrm{e}^{\widetilde{\kappa}x/2}}{\widetilde{\gamma}\cosh\widetilde{\gamma}x + \frac{\widetilde{\kappa}}{2}\sinh\widetilde{\gamma}x}\bigg),\\
\widetilde{\gamma} &= \frac{1}{2}\bigg(\widetilde{\kappa}^2 + 2\frac{|\alpha|}{1-\alpha}\sigma^2\bigg)^{1/2}.
\end{aligned}$$

So the solution to (4.3) corresponding to the CIR model is given by

$$G(t,r) = a^{\frac{\alpha}{1-\alpha}}\int_t^T \mathrm{e}^{m_{1,t,u}+m_{2,t,u}r}\,\mathrm{d}u + b^{\frac{\alpha}{1-\alpha}}\,\mathrm{e}^{m_{1,t,T}+m_{2,t,T}r}, \tag{7.10}$$

where

$$m_{1,t,u} := \widetilde{m}(u-t) - \frac{\gamma}{1-\alpha}(u-t), \qquad m_{2,t,u} := \widetilde{n}(u-t).$$

Summing up, we get

PROPOSITION 7.4. *In the CIR model a candidate for the optimal pair $(\widehat{\psi}, \widehat{C})$ is given by*

$$\int_0^{T^*} \widetilde{n}(x)\,\widehat{\psi}_t(\mathrm{d}x) = \frac{1}{\alpha-1}\,\frac{\lambda(\overline{r}(t))}{\Sigma(\overline{r}(t))} - \frac{D_r G(t,\overline{r}(t))}{G(t,\overline{r}(t))}, \qquad \widehat{C}(t) = G(t,\overline{r}(t))^{-1},$$

where $G(t,r)$ *is given by* (7.10).

7.1.3. *Multidimensional model.* It is time to present a multidimensional model. Here we focus on the G2++ model, important for applications. Let $\overline{r}(t) = Y_1(t) + Y_2(t)$, where

$$\begin{aligned}
\mathrm{d}Y_1(t) &= -\kappa_1 Y_1(t)\,\mathrm{d}t + \sigma_1[\lambda_1\,\mathrm{d}t + \mathrm{d}W_1(t)],\\
\mathrm{d}Y_2(t) &= -\kappa_2 Y_2(t)\,\mathrm{d}t + \sigma_2\big[\rho[\lambda_1\,\mathrm{d}t + \mathrm{d}W_1(t)] + \sqrt{1-\rho^2}\,[\lambda_2\,\mathrm{d}t + \mathrm{d}W_2(t)]\big],
\end{aligned}$$

and W_1,W_2 are independent Brownian motions. In other words, the short rate is the sum of two correlated Vasicek (or Ornstein–Uhlenbeck) processes.

Proposition 7.5. *In the* G2++ *model the optimal pair* $(\widehat{\psi},\widehat{C})$ *is given by*

$$-\int_0^{T^*} (n_1(x), n_2(x))\,\widehat{\psi}_t(\mathrm{d}x) = \frac{(\lambda_1,\lambda_2)\Sigma^{-1}}{1-\alpha} + \frac{D_y G(t,Y(t))}{G(t,Y(t))},$$
$$\widehat{C}(t) = G(t,Y(t))^{-1},$$

where

$$\begin{aligned}
G(t,y) &= a^{\frac{1}{1-\alpha}} \int_t^T \eta(t,u,y)\,\mathrm{d}u + b^{\frac{1}{1-\alpha}}\eta(t,T,y),\\
\eta(t,u,y) &= \mathrm{e}^{m_{1,t,u} + \langle m_{2,t,u}, y\rangle + \frac{1}{2}\sigma^2_{t,u}},\\
\sigma^2_{t,u} &:= \left[\frac{\alpha}{1-\alpha}\right]^2 \Big[\int_t^u (n_1(k)\sigma_1 + n_2(k)\sigma_2\rho)^2 + n_2(k)^2\sigma_2^2(1-\rho^2)\Big]\,\mathrm{d}k,\\
m_{1,t,u} &:= \frac{\alpha}{1-\alpha}\int_t^u \left[n_1(k)\beta_1 + n_2(k)\beta_2 + \frac{\alpha}{2(1-\alpha)}\|\lambda\|^2 + \varphi - \gamma\right]\mathrm{d}k,\\
m_{2,t,u} &:= \left(\frac{\alpha}{1-\alpha} n_1(u-t), \frac{\alpha}{1-\alpha} n_2(u-t)\right),\\
\Sigma &:= \begin{bmatrix} \sigma_1 & 0\\ \rho\sigma_2 & \sigma_2\sqrt{1-\rho^2}\end{bmatrix},\\
n_1(x) &:= \frac{1-\mathrm{e}^{-\kappa_1 x}}{\kappa_1}, \qquad n_2(x) := \frac{1-\mathrm{e}^{-\kappa_2 x}}{\kappa_2}.
\end{aligned}$$

Proof. Here the formula for the function G is

$$G(t,y) = \mathbb{E}\bigg\{\int_t^T a^{\frac{1}{1-\alpha}}\mathrm{e}^{\int_t^u g(\widetilde{Y}_1(k)+\widetilde{Y}_2(k))\,\mathrm{d}k}\,\mathrm{d}u + b^{\frac{1}{1-\alpha}}\mathrm{e}^{\int_t^T g(\widetilde{Y}_1(k)+\widetilde{Y}_2(k))\,\mathrm{d}k}\bigg\},$$

where

$$g(z)=\frac{1}{1-\alpha}\left[\alpha z+\frac{\alpha}{2(1-\alpha)}\|(\lambda_1,\lambda_2)\|^2-\gamma\right]$$

and

$$\begin{aligned}
\mathrm{d}\widetilde{Y}_1(u)&=[\beta_1-\kappa_1\widetilde{Y}_1(u)]\,\mathrm{d}u+\sigma_1\,\mathrm{d}W_1(u),\\
\widetilde{Y}_1(t)&=y_1,\\
\mathrm{d}\widetilde{Y}_2(u)&=[\beta_2-\kappa_2\widetilde{Y}_2(u)]\,\mathrm{d}t+\sigma_2(\rho\mathrm{d}W_1(u)+\sqrt{1-\rho^2}\,\mathrm{d}W_2(u)),\\
\widetilde{Y}_2(t)&=y_2,\\
\beta_1&:=\frac{1}{1-\alpha}\lambda_1\sigma_1,\qquad \beta_2:=\frac{1}{1-\alpha}(\lambda_1\sigma_2\rho+\lambda_2\sigma_2\sqrt{1-\rho^2}).
\end{aligned}$$

Taking advantage of the Vasicek model we obtain

$$\begin{aligned}
\int_t^u(\widetilde{Y}_1(k)+\widetilde{Y}_2(k))\,\mathrm{d}k&=n_1(u-t)y_1+n_2(u-t)y_2+\int_t^u n_1(k)\beta_1\,\mathrm{d}k\\
&\quad+\int_t^u n_2(k)\beta_2\,\mathrm{d}k+\int_t^u(n_1(k)\sigma_1+n_2(k)\sigma_2\rho)\,\mathrm{d}W_1(k)\\
&\quad+\int_t^u n_2(k)\sigma_2\sqrt{1-\rho^2}\,\mathrm{d}W_2(k).
\end{aligned}$$

This implies the desired identities. ∎

In particular, if we take

$$\widehat{\psi}_t(\mathrm{d}x)=\widehat{\eta}_0(t)\,\delta_0(\mathrm{d}x)+\widehat{\eta}_{x_1}(t)\,\delta_{x_1}(\mathrm{d}x)+\widehat{\eta}_{x_2}(t)\,\delta_{x_2}(\mathrm{d}x),$$

where $x_1,x_2\in(0,T^*]$, then the process $(\widehat{\eta}_0(t),\widehat{\eta}_{x_1}(t)),\widehat{\eta}_{x_2}(t))$ is determined by

$$\begin{aligned}
(\widehat{\eta}_{x_1}(t),\widehat{\eta}_{x_2}(t))&=\frac{(\lambda_1,\lambda_2)\Sigma^{-1}M(x_1,x_2)^{-1}}{\alpha-1}-\frac{D_yG(t,Y(t))M(x_1,x_2)^{-1}}{G(t,Y(t))},\\
\widehat{\eta}_0(t)&=1-\widehat{\eta}_{x_1}(t)-\widehat{\eta}_{x_2}(t),
\end{aligned}$$

where

$$M(x_1,x_2):=\begin{bmatrix}n_1(x_1)&n_1(x_2)\\ n_2(x_1)&n_2(x_2)\end{bmatrix}.$$

8. Infinite horizon problem. Recall that in the infinite horizon case the reward functional is given by (3.8), and the short rate has the form $\overline{r}(t)=\varphi(Y(t))$ where Y solves (3.1). We assume that Hypotheses 1 to 3 are fulfilled. Moreover, in this section, φ as well as the coefficients B, Σ and λ appearing in (3.1) do not depend on the time variable.

Let

$$V(z,y) = \sup_{\psi,C} J(z,y,\psi,C)$$

be the value function of the investor.

We start by providing a simple example of a model that satisfies our Hypotheses 1 to 3, but produces an infinite value function V.

EXAMPLE 8.1. Consider the Merton model

$$\mathrm{d}\overline{r}(t) = \beta\,\mathrm{d}t + \sigma(\lambda\,\mathrm{d}t + \mathrm{d}W(t)).$$

Choose $\psi = \delta_0$, which corresponds to the investment in the bank account only, and fix the consumption at a constant level $c > 0$. Then $C(t) \equiv c$ and

$$\mathrm{d}z^{\psi,C}(t) = [\overline{r}(t) - c]z^{\psi,C}(t)\,\mathrm{d}t.$$

Thus

$$z^{\psi,C}(t) = z\,\mathrm{e}^{\int_0^t [\overline{r}(u)-c]\,\mathrm{d}u},$$

and

$$J(z,r,\psi,C) = \frac{cz}{\alpha}\,\mathbb{E}\int_0^{+\infty} \mathrm{e}^{\int_0^t [\alpha\overline{r}(u)-\alpha c-\gamma]\,\mathrm{d}u}\,\mathrm{d}t.$$

Note that

$$\int_0^t \overline{r}(u)\,\mathrm{d}u = \frac{(\beta+\lambda)t^2}{2} + \sigma\int_0^t W(u)\,\mathrm{d}u.$$

Since $\int_0^t W(u)\,\mathrm{d}u$ has the normal distribution with variance $t^3/3$ we have $J(z,r,\psi,C) = +\infty$ (see Synowiec [33]).

The HJB approach gives the following candidate for the value function V, optimal consumption $\widehat{C}$ and investment strategy $\widehat{\psi}$: we can expect that

$$V(z,y) = G(y)^{\frac{1}{1-\alpha}} z^{\alpha}, \tag{8.1}$$

$$\widehat{C}(t) = G(Y(t))^{-1}, \tag{8.2}$$

and $\widehat{\psi} \in \mathcal{M}_{+\infty}$ is such that

$$A(\widehat{\psi}_t) = \frac{\lambda(Y(t))}{1-\alpha} + \varSigma(Y(t))^{\top}\frac{D_yG(Y(t))}{G(Y(t))}, \tag{8.3}$$

where G solves the elliptic equation

$$LG(y) + \frac{\alpha}{1-\alpha}\langle \varSigma(y)\lambda(y), D_yG(y)\rangle + g(y)G(y) + 1 = 0. \tag{8.4}$$

Recall that L given by (4.1) is the generator of the diffusion defined by (3.1) and g is given by (4.2). Finally, we can expect that G has the Feynman–Kac representation

$$G(y) = \mathbb{E}\int_0^{+\infty} \mathrm{e}^{\int_0^u g(\widetilde{Y}(k))\,\mathrm{d}k}\,\mathrm{d}u, \tag{8.5}$$

where $\widetilde{Y}$ solves the stochastic differential equation

$$\mathrm{d}\widetilde{Y}(t) = \left[B(\widetilde{Y}(t)) + \frac{1}{1-\alpha}\Sigma(\widetilde{Y}(t))\lambda(\widetilde{Y}(t))\right]\mathrm{d}t + \Sigma(\widetilde{Y}(t))\,\mathrm{d}W(t)$$

with $\widetilde{Y}(0) = y$. Taking into account Example 8.1, we see that a special care has to be taken to ensure the existence of a solution to (8.5) or to the convergence of the integral in (8.5). Moreover, at the end we will have to verify the assumption of the verification theorem.

Below we present a general theorem which ensures that convergence. Let $h\colon \mathbb{R}^n \to \mathbb{R}$ be a Lipschitz continuous function. Let us consider the elliptic equation

$$LG(y) + h(y)G + 1 = 0, \qquad y \in \mathbb{R}^n. \tag{8.6}$$

THEOREM 8.2. *In addition to Hypotheses* 1 *to* 3, *suppose that there exists a constant* $L_2 > 0$ *and a function* $\kappa\colon [0,+\infty)\times\mathbb{N}\to\mathbb{R}$ *decreasing in the first argument such that*

$$\langle B(x) - B(y), x - y\rangle + \tfrac{1}{2}\|\Sigma(x) - \Sigma(y)\|^2 \le -L_2\|x-y\|^2,$$

and

$$\mathbb{E}\,\mathrm{e}^{\int_0^t 2h(Y(u))\,\mathrm{d}u} \le \kappa(t,n), \qquad y \in B(0,n), \qquad \int_0^{+\infty} \sqrt{t\kappa(t,n)}\,\mathrm{d}t < +\infty. \tag{8.7}$$

Then

$$G(y) := \mathbb{E}\int_0^{+\infty} \mathrm{e}^{\int_0^t h(Y(u))\,\mathrm{d}u}\,\mathrm{d}t$$

is a classical, $C^2(\mathbb{R}^n)$ *solution to* (8.6).

Proof. Consider the parabolic problem

$$\frac{\partial G}{\partial t}(t,y) - LG(t,y) - h(y)G(t,y) - 1 = 0, \qquad G(0,y) = 0.$$

We have

$$G(t,y) = \mathbb{E}\int_0^t \mathrm{e}^{\int_0^s h(Y(k))\,\mathrm{d}k}\,\mathrm{d}s, \qquad Y(0) = y.$$

Note that

$$\frac{\partial G}{\partial t}(t,y) = \mathbb{E}\,\mathrm{e}^{\int_0^t h(Y(k))\,\mathrm{d}k} \le \left[\mathbb{E}\,\mathrm{e}^{\int_0^t 2h(Y(k))\,\mathrm{d}k}\right]^{1/2}.$$

Therefore, by (8.7), $\frac{\partial G}{\partial t}$ converges to 0 as $t \to +\infty$, uniformly on each ball $B(0,n)$.

Secondly, we need to estimate the Lipschitz constant of G in y. Let $Y(\cdot;y)$ be the solution to SDE (3.1) with initial condition $Y(0;y) = y$. By standard SDE estimates there exists a constant $M > 0$ such that for all $t > 0$ and $y_1, y_2 \in \mathbb{R}^n$,

$$\mathbb{E}\,\|Y(t;y_1) - Y(t;y_2)\|^2 \le M\mathrm{e}^{-2L_2 t}\|y_1 - y_2\|^2.$$

Because the function h is Lipschitz continuous there exists a constant $N > 0$ such that

$$\begin{aligned}
&|G(t,y_1) - G(t,y_2)| \\
&\quad \le N\,\mathbb{E}\int_0^t \mathrm{e}^{\max\{\int_0^s h(Y(k;y_1))\,\mathrm{d}k, \int_0^s h(Y(k;y_2))\,\mathrm{d}k\}} \int_0^s \|Y(k;y_1) - Y(k;y_2)\|\,\mathrm{d}k\,\mathrm{d}s \\
&\quad \le N\int_0^t \Big[\mathbb{E}\int_0^s \|Y(k;y_1) - Y(k;y_2)\|^2\,\mathrm{d}k\Big]^{1/2} \\
&\qquad \times \big[\mathbb{E}\,s\mathrm{e}^{2\max\{\int_0^s h(Y(k;y_1))\,\mathrm{d}k, \int_0^s h(Y(k;y_2))\,\mathrm{d}k\}}\big]^{1/2}\,\mathrm{d}s.
\end{aligned}$$

Letting $y_1 \to y_2$ and using the dominated convergence theorem we get

$$\|D_y G(t,y)\| \le N_1 \int_0^t \Big[\int_0^s \mathrm{e}^{-2L_2 k}\,\mathrm{d}k\Big]^{1/2} \big[\mathbb{E}\,s\mathrm{e}^{\int_0^s 2h(Y(k;y))\,\mathrm{d}k}\big]^{1/2}\,\mathrm{d}s.$$

Finally, for $y \in B(0,n)$ we get

$$\|D_y G(t,y)\| \le \frac{N_1}{\sqrt{2L_2}} \int_0^{+\infty} \sqrt{s\kappa(s,n)}\,\mathrm{d}s.$$

Almost the same estimates can be made for the Lipschitz constant of $\frac{\partial G}{\partial t}$. In fact,

$$\left|\frac{\partial G}{\partial t}(t,y_1) - \frac{\partial G}{\partial t}(t,y_2)\right| = \big|\mathrm{e}^{\int_0^t 2h(Y(k;y_1))\,\mathrm{d}k} - \mathrm{e}^{\int_0^t 2h(Y(k;y_2))\,\mathrm{d}k}\big|.$$

And by repetitive arguments we arrive at the estimate

$$\left|\frac{\partial G}{\partial t}(t,y_1) - \frac{\partial G}{\partial t}(t,y_2)\right| \le \widetilde{N}\sqrt{\kappa(t,n)}\,\|y_1 - y_2\|, \qquad y_1, y_2 \in B(0,n).$$

Now we may use the Schauder estimates (see e.g. Gilbarg–Trudinger [16, Theorem 6.2]) to prove that there exists a sequence $(t_n, n \in \mathbb{N})$ such that $G(y,t_n)$ converges uniformly on each ball to the function $G \in C^2(\mathbb{R}^n)$ satisfying (8.6). ■

REMARK 8.3. Condition (8.7) is not in analytic form but it can be easily verified for example in the affine model framework.

In the case of φ unbounded the use of the verification theorem needs a justification. To do this consider an arbitrary sequence $(t_n, n \in \mathbb{N})$ of finite time investment horizons and the corresponding sequence of value functions

$$V_{t_n}(z,y) := \sup_{\psi,C} \mathbb{E}\, \frac{1}{\alpha} \int_0^{t_n} \mathrm{e}^{-\gamma t} (C(t) z^{\psi,C}(t))^{\alpha}\, \mathrm{d}t.$$

Let $((\widehat{\psi}_n, \widehat{C}_n), n \in \mathbb{N})$ be the corresponding sequence of optimal pairs of controls.

Theorem 8.4 (Verification theorem). *Assume Hypotheses* 1 *to* 4. *Additionally assume that there exists a sequence* $(t_n, n \in \mathbb{N})$ *with* $t_n \to \infty$ *such that* $V_{t_n}(z,y) \to V(z,y)$, *where* V *is given by* (8.1) *and* G *is a* $C^2(\mathbb{R}^n)$ *classical solution to* (8.4). *Then any pair* $(\widehat{\psi}, \widehat{C})$ *satisfying* (8.2), (8.3) *and*

$$\begin{aligned} &\mathbb{E} \sup_{0\le k\le t} [(z_k^{\widehat{\psi},\widehat{c}})^{\alpha} G^{\frac{1}{1-\alpha}}(Y_k)] < +\infty, \qquad \forall t \ge 0, \\ &\lim_{t\to\infty} \mathbb{E}\, \mathrm{e}^{-\gamma t} V(z^{\widehat{\psi},\widehat{C}}(t), Y(t)) = 0, \end{aligned} \tag{8.8}$$

is an optimal solution for the infinite horizon optimization problem.

Proof. Suppose $V_{t_n}(z,y)$ converges to $V(z,y)$. Choose any admissible strategy $(\psi, C) \in \mathcal{A}_{+\infty}$. Note that

$$V_{t_n}(z,y) = \mathbb{E}\, \frac{1}{\alpha} \int_0^{t_n} \mathrm{e}^{-\gamma s} (\widehat{C}_n(s) z^{\widehat{\psi}_n,\widehat{C}_n}(s))^{\alpha}\, \mathrm{d}s \ge \mathbb{E}\, \frac{1}{\alpha} \int_0^{t_n} \mathrm{e}^{-\gamma s} (C(s) z^{\psi,C}(s))^{\alpha}\, \mathrm{d}s.$$

This ensures that

$$V(z,y) = \sup_{\psi,C} \mathbb{E}\, \frac{1}{\alpha} \int_0^{+\infty} \mathrm{e}^{-\gamma s} (C(s) z^{\psi,C}(s))^{\alpha}\, \mathrm{d}s. \tag{8.9}$$

Now, we need only prove that the supremum in (8.9) is attained at $(\widehat{\psi}, \widehat{C})$. Let us apply the Itô formula to obtain the dynamics of $\mathrm{e}^{-\gamma t} V(z^{\widehat{\psi},\widehat{C}}(t), Y(t))$. We obtain

$$\begin{aligned} &\mathbb{E}\, \mathrm{e}^{-\gamma t\wedge\tau_n} V(z^{\widehat{\psi},\widehat{C}}(t\wedge\tau_n), Y(t\wedge\tau_n)) \\ &\qquad\qquad = V(z,y) - \mathbb{E}\, \frac{1}{\alpha} \int_0^{t\wedge\tau_n} e^{-\gamma s} (\widehat{C}(s) z^{\widehat{\psi},\widehat{C}}(s))^{\alpha}\, \mathrm{d}s, \end{aligned}$$

where $\{\tau_n\}_{n\in\mathbb{N}}$ is a localizing sequence of stopping times. We can now let $n \to +\infty$ and use the first condition of (8.8) to apply dominated convergence on the left hand side. On the right hand side we can use the monotone convergence theorem. Altogether, we have

$$\mathbb{E}\,\mathrm{e}^{-\gamma t}V(z^{\widehat{\psi},\widehat{C}}(t),Y(t))=V(z,y)-\mathbb{E}\,\frac{1}{\alpha}\int_0^t e^{-\gamma s}(\widehat{C}(s)z^{\widehat{\psi},\widehat{C}}(s))^{\alpha}\,\mathrm{d}s.$$

Consequently, by applying the second condition of (8.8), we obtain the desired formula

$$V(z,y)=\mathbb{E}\,\frac{1}{\alpha}\int_0^{+\infty} e^{-\gamma s}(\widehat{C}(s)z^{\widehat{\psi},\widehat{C}}(s))^{\alpha}\,\mathrm{d}s. \ \blacksquare$$

EXAMPLE 8.5 (Vasicek model). In the Vasicek model (see Section 7.1.1), $G(t,r)$ is given by (7.9) and consequently

$$G(r)=\int_0^{+\infty}\mathrm{e}^{m_{1,0,s}+m_{2,0,s}r+\frac{1}{2}\sigma_{0,s}^2}\,\mathrm{d}s, \tag{8.10}$$

where $\sigma_{0,s}$, $m_{1,0,s}$, and $m_{1,0,s}$ were defined in Proposition 7.3. To ensure convergence of the integral in (8.10) we have to assume

$$\int_0^{+\infty}\mathrm{e}^{m_{1,0,s}+\frac{1}{2}\sigma_{0,s}^2}\,\mathrm{d}s<+\infty.$$

It is not difficult to find a sufficient condition for the coefficients of the model to ensure that convergence; we leave this to the reader. However, it should be noted that the coefficient $m_{2,0,s}$ is uniformly bounded and therefore the fraction $\frac{D_rG}{G}$ is uniformly bounded as well.

EXAMPLE 8.6 (CIR model). In the CIR model (see Section 7.1.2), $G(t,r)$ is given by (7.10) and consequently $G(r)=\int_0^{+\infty}\mathrm{e}^{m_{1,0,s}+m_{2,0,s}r}\,\mathrm{d}s$. Obviously, we should require that $\int_0^{+\infty}\mathrm{e}^{m_{1,0,s}}\,\mathrm{d}s<+\infty$. Note that under this assumption the quotient $\frac{D_rG}{G}$ is uniformly bounded.

Appendix. Short introduction to the HJM model. Let us denote by $P(t,S)$ the price at time t of a bond paying 1 at time S. Assume that the *forward rates* $f(t,S)=-\frac{\partial}{\partial S}\log P(t,S)$, $0\le t\le S$, are given by the Itô equation

$$\mathrm{d}f(t,S)=\mu(t,S)\,\mathrm{d}t+\langle\xi(t,S),\mathrm{d}W(t)\rangle,$$

where $W=(W_1,\dots,W_m)$ is an m-dimensional Wiener process defined on a filtered probability space $(\Omega,\mathfrak{F},(\mathfrak{F}_t),\mathbb{P})$, μ and ξ are $\mathbb{R}$- and $\mathbb{R}^m$-valued processes which may depend on the forward rate f, and $\langle\cdot,\cdot\rangle$ denotes the Euclidean scalar product. We denote by $\|\cdot\|$ the corresponding norm. Clearly $P(t,S)=\mathrm{e}^{-\int_t^S f(t,u)\,\mathrm{d}u}$. The so-called *short rate* process $\overline{r}(t):=f(t,t)$ defines the bank rate at time t.

Let $T \in (0,+\infty)$ be a finite time horizon. It is well-known (see e.g. Barski and Zabczyk [1] or the original Heath, Jarrow and Morton paper [22]) that the model $P(t,S)$, where $t \in [0,T]$ and $t \le S < +\infty$, is free of arbitrage if and only if there is an adapted process λ such that

$$\mathbb{P}\Big(\int_0^T \|\lambda(u)\|^2 \,\mathrm{d}u < +\infty\Big) = 1, \qquad \mathbb{E}\,\mathcal{E}(\lambda) = 1,$$

where

$$\mathcal{E}\,(\lambda) := \mathrm{e}^{-\int_0^T \langle \lambda(u), \mathrm{d}W(u)\rangle - \frac{1}{2}\int_0^T \|\lambda(u)\|^2\,\mathrm{d}u},$$

and the following *HJM condition* is satisfied:

$$\mu(t,S) = \Big\langle \xi(t,S), \int_t^S \xi(t,u)\,\mathrm{d}u + \lambda(t)\Big\rangle, \qquad \forall 0 \le t \le T,\ \forall t \le S.$$

Recall that $\mathrm{d}\mathbb{P}^* = \mathcal{E}(\lambda)\,\mathrm{d}\mathbb{P}$ is the *martingale measure*; the discounted prices $P(t,S)\mathrm{e}^{-\int_0^t \overline{r}(s)\,\mathrm{d}s}$, $S \le T$, are local martingales with respect to $\mathbb{P}^*$. Moreover, $W^*(t) = W(t) + \int_0^t \lambda(u)\,\mathrm{d}u$ is a Wiener process with respect to $\mathbb{P}^*$.

For our purposes it is convenient to rewrite the prices and forward rates in the so-called *Musiela parametrization*

$$P(t)(x) := P(t,t+x), \qquad r(t)(x) := f(t,t+x), \qquad x \ge 0.$$

Then

$$P(t)(x) = \mathrm{e}^{-\int_0^x r(t)(u)\,\mathrm{d}u}, \qquad r(t)(x) = -\frac{\partial}{\partial x}\log P(t)(x),$$

the short rate is given by $\overline{r}(t) := r(t)(0)$ and

$$r(t)(x) = r(0)(t+x) + \int_0^t b(s)(x+t-u)\,\mathrm{d}u + \int_0^t \langle \sigma(s)(x+t-u), \mathrm{d}W(u)\rangle,$$

where

$$b(t)(x) := \mu(t,t+x), \qquad \sigma(t)(x) := \xi(t,t+x), \qquad x \ge 0.$$

Note that the HJM condition has the form

$$\begin{aligned} b(t)(x) = \mu(t,t+x) &= \Big\langle \xi(t,t+x), \int_t^{t+x} \xi(t,u)\,\mathrm{d}u + \lambda(t)\Big\rangle \\ &= \Big\langle \sigma(t)(x), \int_0^x \sigma(t)(y)\,\mathrm{d}y + \lambda(t)\Big\rangle = \langle \sigma(t)(x), \tilde{\sigma}(t)(x) + \lambda(t)\rangle, \end{aligned}$$

where

$$\tilde{\sigma}(t)(x) := \int_0^x \sigma(t)(u)\,\mathrm{d}u.$$

In general b and σ may depend on r.

Informally, $S(t)\psi(x) = \psi(x+t)$ is the semigroup generated by the operator $\frac{\partial}{\partial x}$. Thus r is the so-called mild solution to the stochastic partial differential equation

$$\mathrm{d}r = \left(\frac{\partial r}{\partial x} + b\right)\mathrm{d}t + \langle \sigma, \mathrm{d}W\rangle.$$

We can now compute the stochastic derivative of

$$P(t)(x) = \mathrm{e}^{-\int_0^x r(t)(u)\,\mathrm{d}u}, \qquad t \geq 0.$$

We have

$$\mathrm{d}P(t)(x) = P(t)(x)\Big[-\mathrm{d}\int_0^x r(t)(u)\,\mathrm{d}u + \frac{1}{2}\Big\|\int_0^x \sigma(t)(u)\,\mathrm{d}u\Big\|^2\,\mathrm{d}t\Big].$$

Since, by the HJM condition,

$$\begin{aligned}
-\int_0^x & b(t)(u)\,\mathrm{d}u + \frac{1}{2}\Big\|\int_0^x \sigma(t)(u)\,\mathrm{d}u\Big\|^2 \\
&= -\int_0^x \Big\langle \sigma(t)(u), \int_0^u \sigma(t)(y)\,\mathrm{d}y + \lambda(t)\Big\rangle\,\mathrm{d}u + \frac{1}{2}\Big\|\int_0^x \sigma(t)(u)\,\mathrm{d}u\Big\|^2 \\
&= -\Big\langle \lambda(t), \int_0^x \sigma(t)(u)\,\mathrm{d}u\Big\rangle,
\end{aligned}$$

we eventually have

$$\begin{aligned}
\frac{\mathrm{d}P(t)(x)}{P(t)(x)} &= -\int_0^x \frac{\partial r}{\partial u}(t)(u)\,\mathrm{d}u\,\mathrm{d}t - \Big\langle \int_0^x \sigma(t)(u)\,\mathrm{d}u, \lambda(t)\,\mathrm{d}t + \mathrm{d}W(t)\Big\rangle \\
&= [-r(t)(x) + \overline{r}(t)]\,\mathrm{d}t - \langle \tilde{\sigma}(t)(x), \lambda(t)\,\mathrm{d}t + \mathrm{d}W(t)\rangle.
\end{aligned}$$

The instrument $P(t)(x)$, $t \geq 0$, is called a *sliding bond*. After discounting it by a bank account we get

$$\frac{\mathrm{d}\bar{P}(t)(x)}{\bar{P}(t)(x)} = -r(t)(x)\,\mathrm{d}t - \langle \tilde{\sigma}(t)(x), \lambda(t)\,\mathrm{d}t + \mathrm{d}W(t)\rangle.$$

The discounted sliding bonds are not tradable instruments. To work with a portfolio process we use *rolling bonds* (see (1.2)).

Acknowledgements. The authors would like to thank the anonymous referees for careful reading of the paper and for providing many useful remarks and comments which have led to a substantially improved and more complete presentation.

The work was supported by Polish National Science Center grant 2017/25/B/ST1/02584.

References

[1] M. Barski and J. Zabczyk, *Mathematics of the Bond Market (a Lévy Processes Approach)*, Cambridge Univ. Press, 2020.

[2] M. Barski and J. Zabczyk, *On CIR equations with general factors*, SIAM J. Financial Math. 11 (2020), 131–147.

[3] P. Battochio and F. Menoncin, *Optimal pension management in a stochastic framework*, Insurance Math. Econom. 34 (2004), 79–95.

[4] I. Beyna, *Interest Rate Derivatives: Valuation, Calibration and Sensitivity Analysis*, Springer, 2013.

[5] T. Bielecki, S. Pliska and J. Yong, *Optimal investment decisions for a portfolio with a rolling horizon bond and a discount bond*, Int. J. Theor. Appl. Finance 8 (2005), 871–913.

[6] T. Bielecki and S. Pliska, *Risk-sensitive ICAPM with application to fixed-income management*, IEEE Trans. Autom. Control 49 (2004), 420–432.

[7] T. Björk, G. Di Masi, Y. Kabanov and W. Rungaldier, *Towards a general theory of bond markets*, Finance Stoch. 1 (1997), 141–174.

[8] M. J. Brennan and Y. Xia, *Stochastic interest rates and the bond-stock mix*, Rev. Financ. 4 (2000), 197–210.

[9] H. Chang and K. Chang, *Optimal investment consumption strategy in the Vasicek model: HARA utility and Legandre transform*, Insurence Math. Econom. 72 (2017), 215–227.

[10] J. C. Cox and C. F. Huang, *Optimal consumption and portfolio policies when asset prices follow a diffusion process*, J. Econom. Theory 49 (1989), 33–83.

[11] G. Deelstra, M. Grasselli and P. F. Koehl, *Optimal investment strategies in a CIR framework*, J. Appl. Probab. 37 (2000), 936–946.

[12] J. Detemple and M. Rindisbacher, *Closed-form solutions for optimal portfolio selection with stochastic interest rate and investment constraints*, Math. Finance 15 (2005), 539–568.

[13] D. Filipović, *Term-Structure Models: a Graduate Course*, Springer, Berlin, 2009.

[14] W. Fleming and H. M. Soner, *Controlled Markov Processes and Viscosity Solutions*, 2nd ed., Springer, New York, 2006.

[15] A. Friedman, *Stochastic Differential Equations and Applications*, Vol. 1, Academic Press, New York, 1975.

[16] D. Gilbarg and N. S. Trudinger, *Elliptic Partial Differential Equations of Second Order*, Springer, 2015.

[17] G. Guan and Z. Liang, *Mean-variance efficiency of DC pension plan under stochastic interest rate and mean-reverting returns*, Insurance Math. Econom. 61 (2015), 99–109.

[18] G. Guan and Z. Liang, *Robust optimal reinsurance and investment strategies for an AAI with multiple risks*, Insurance Math. Econom. 89 (2019), 63–78.

[19] P. Guasoni and G. Wang, *Consumption and investment with interest rate risk*, J. Math. Anal. Appl. 476 (2019), 215–239.

[20] P. Guasoni and G. Wang, *Consumption in incomplete markets*, Finance Stoch. 24 (2020), 383–422.

[21] H. Hata, H. Nagai and S. J. Sheu, *An optimal consumption problem for general factor models*, SIAM J. Control Optim. 56 (2018), 3149–3183.

[22] D. Heath, R. Jarrow and A. Morton, *Bond pricing and the term structure of interest rates: A new methodology for contingent claim valuation*, Econometrica 60 (1992), 77–105.

[23] I. Karatzas, J. P. Lehoczky and S. E. Shreve, *Optimal portfolio and consumption for a "small investor" on a finite horizon*, SIAM J. Control Optim. 25 (1987), 1557–1586.

[24] R. Korn and H. Kraft, *A stochastic control approach to portfolio problems with stochastic interest rates*, SIAM J. Control Optim. 40 (2001), 1250–1269.

[25] R. Korn and H. Kraft, *On the stability of continuous-time portfolio problems with stochastic opportunity set*, Math. Finance 14 (2004), 403–414.

[26] C. Munk and C. Sørensen, *Optimal consumption and investment strategies with stochastic interest rates*, J. Banking Finance 28 (2004), 1987–2013.

[27] H. Nagai, *H-J-B equations of optimal consumption-investment and verification theorems*, Appl. Math. Optim. 71 (2015), 279–311.

[28] B. Øksendal and A. Sulem, *Applied Stochastic Control of Jump Diffusions*, Springer, 2007.

[29] A. Palczewski, *Risk minimizing strategies for a portfolio of interest-rate securities*, in: Banach Center Publ. 83, Inst. Math., Polish Acad. Sci., 2008, 195–212.

[30] D. Pirjol and L. Zhu, *Explosion in the quasi-Gaussian HJM model*, Finance Stoch. 22 (2018), 643–666.

[31] N. Ringer and M. Tehranchi, *Optimal portfolio choice in the bond market*, Finance Stoch. 10 (2006), 553–573.

[32] M. Rutkowski, *Self-financing trading strategies for sliding, rolling-horizon, and consol bonds*, Math. Finance 9 (1999), 361–385.

[33] D. Synowiec, *Optimal consumption problem in a diffusion short-rate model*, arXiv: 0910.0378 (2009).

[34] J. Trybuła, *Optimal consumption problem in the Vasicek model*, Opuscula Math. 35 (2015), 547–560.

[35] J. Wachter, *Portfolio and consumption decisions under mean-reverting returns: An exact solution for complete markets*, J. Financial Quant. Anal. 37 (2002), 63–91.

[36] D. Zawisza, *Existence results for Isaacs equations with local conditions and related semilinear Cauchy problems*, Ann. Polon. Math. 121 (2018), 175–196.

[37] D. Zawisza, *On the parabolic equation for portfolio problems*, in: Banach Center Publ. 122, Inst. Math., Polish Acad. Sci., 2020, 287–302.

[38] L. Zhang, D. Li and Y. Lai, *Equilibrium investment strategy for a defined contribution pension plan under stochastic interest rate and stochastic volatility*, J. Comput. Appl. Math. 368 (2020), 1125–1136.

Szymon Peszat, Dariusz Zawisza
Faculty of Mathematics and Computer Science
Jagiellonian University
Łojasiewicza 6
30-348 Kraków, Poland
E-mail: szymon.peszat@im.uj.edu.pl
dariusz.zawisza@im.uj.edu.pl