Convergence Rates of Turnpike Theorems for Portfolio Choice in Stochastic Factor Models

Hiroki Yamamichi[*]

Abstract

Turnpike theorems state that if an investor's utility is asymptotically equivalent to a power utility, then the optimal investment strategy converges to the CRRA strategy as the investment horizon tends to infinity. This paper aims to derive the convergence rates of the turnpike theorem for optimal feedback functions in stochastic factor models. In these models, optimal feedback functions can be decomposed into two terms: myopic portfolios and excess hedging demands. We obtain convergence rates for myopic portfolios in nonlinear stochastic factor models and for excess hedging demands in quadratic term structure models, where the interest rate is a quadratic function of a multivariate Ornstein–Uhlenbeck process. We show that the convergence rates are determined by (i) the decay speed of the price of a zero-coupon bond and (ii) how quickly the investor's utility becomes power-like at high levels of wealth. As an application, we consider optimal collective investment problems and show that sharing rules for terminal wealth affect convergence rates.
Keywords: portfolio choice, turnpike property, convergence rate, stochastic opportunity sets, collective utility function.

1 Introduction

Since the seminal works of Merton [30, 31], optimal investment problems for continuous-time models have been developed in various directions. In particular, stochastic factor models have been used to capture *stochastic investment opportunity sets*, such as the predictability of stock returns, stochastic volatility, and stochastic interest rates. For an overview of optimal investment problems with stochastic factor models, we refer the reader to the review paper [45]. In these models, optimal investment strategies can be decomposed into two terms, namely, *myopic portfolios* and *excess hedging demands*. This means that at first investors choose the myopic portfolios as if the investment opportunity sets are constant, then they adjust their portfolios by adding the excess hedging demands to adapt to future changes in the investment environment. Although computing these terms typically requires analyzing fully nonlinear Hamilton–Jacobi–Bellman (HJB) equations, the computations become tractable for homothetic utilities, such as exponential, power, and log utilities. As a result, explicit formulas for optimal investment strategies are typically available for these special utilities. In contrast, it is challenging to analyze optimal investment strategies for generic utilities, and, to the best of our knowledge, only a few works have calculated the optimal strategies, including Detemple and Rindisbacher [6], Fukaya [10], Lakner [26], Ocone and Karatzas [36], and Putschögl and Sass [37].

Turnpike theorems fill this gap between power utilities and general utilities. Informally, turnpike theorems state that if a utility function is similar to a power utility at large wealth levels, then both the optimal wealth process and the optimal investment strategy converge to those for a power utility as the investment horizon tends to infinity. This paper aims to derive the convergence rates of turnpike theorems for both myopic portfolios and excess hedging demands in stochastic factor models.

Turnpike theorems originate in the classic work of von Neumann [35] in economic growth theory. In the context of optimal portfolios, Mossin [32] first proves portfolio turnpikes in discrete-time settings under the assumption that a utility function U has affine risk tolerance, that is, $-\dfrac{U'(x)}{U''(x)} = ax + b$. Mossin's results are extended by Leland [26], Ross [41], and Hakansson [14] to include general utility functions. Huberman and Ross [16] derive a necessary and sufficient condition for the turnpike property.

---
[*]Graduate School of Engineering Science, The University of Osaka, 1-3, Machikaneyama, Toyonaka, Osaka, 560-8531, Japan, Email: `yamamichi@sigmath.es.osaka-u.ac.jp`

Cox and Huang [5] prove the first turnpike theorem in continuous-time settings using martingale methods under the assumption that there exist constants $A_1, A_2, b, z^* > 0$ such that

$$\left| (U')^{-1}(z) - A_1 z^{-\frac{1}{b}} \right| \leq A_2 z^{-a}, \quad z \in (0, z^*] \tag{1}$$

holds for some $a \in [0, 1/b)$. Jin [19] extends their results to include consumption. Huang and Zariphopoulou [17] show that the condition

$$\lim_{x \nearrow \infty} \frac{U'(x)}{x^{\gamma-1}} = K$$

for some $K > 0$, which is weaker than the condition (1), is sufficient for the turnpike property, using viscosity solutions for the associated HJB equations. Dybvig et al. [9] prove portfolio turnpikes for complete markets in the Brownian filtration without assuming stationary investment opportunity sets. Guasoni et al. [13] consider general incomplete market models that include one-dimensional stochastic factor models, and they prove three types of turnpikes (abstract, classic, and explicit).

Although it is important to know how fast turnpike theorems hold in practice, the above works do not derive convergence rates. However, Bian and Zheng [3] first estimate the convergence rate of the turnpike property under the Black–Scholes model. Beyond the classical expected utility framework, Geng and Zariphopoulou [12] recently studied turnpike-type limiting properties and the convergence rate for the forward relative risk tolerance function under time-monotone forward performance criteria in Ito diffusion markets. Since Geng and Zariphopoulou [12] focus on the subclass of time-monotone forward utilities, there is still no result on convergence rates for excess hedging demands. In the present paper, we aim to derive convergence rates for both myopic portfolios and excess hedging demands in stochastic factor models.

Our contributions are fourfold. First, we derive convergence rates of turnpike theorems for myopic portfolios with general stochastic factor models in complete markets (Theorem 2.4), which extends the results of Bian and Zheng [3] from the Black–Scholes model to more general settings. In particular, we find that the convergence rate is determined by the price of a zero-coupon bond and the rate at which the investor's utility becomes power-like at high levels of wealth. As made precise in Remark 2.7, the rate is typically exponential. Moreover, we prove uniform convergence in the wealth variable for portfolio proportions (Theorem 2.5), which has not yet been documented. Furthermore, we also derive the convergence rate for the optimal portfolio processes (Theorem 2.6).

Second, we also derive convergence rates of turnpike theorems for excess hedging demands with quadratic term structure models (Theorem 2.7), where the instantaneous rate is a quadratic function of the stochastic factor process. To the best of our knowledge, no previous studies have derived the convergence rates for excess hedging demands. We find that the convergence rates are the same as those of myopic portfolios and uniform convergence in wealth holds (Theorem 2.8).

Third, by applying our main results, we analyze the turnpike properties of the optimal strategies for optimal collective investment problems. In these problems, there are n investors who delegate their portfolio management to a fund manager. The fund manager invests on their behalf to optimize the expected social utility, constructed from the individual utilities and a sharing rule, according to which the fund manager allocates the terminal wealth to individuals. We consider two sharing rules: a Pareto-optimal sharing rule and a linear sharing rule. We find that these sharing rules affect convergence rates (Theorems 2.10 and 2.12). In particular, we show that the convergence rate for a linear sharing rule is faster than for a Pareto-optimal sharing rule when the least risk-averse investor among n investors is no less risk-averse than the log investor.

Finally, we methodologically provide a probabilistic approach based on martingale duality methods and Malliavin calculus, in contrast to the PDE techniques used in prior work by Bian and Zheng [3]. When applying Malliavin calculus to optimal investment problems, previous research papers such as [26, 37] often assume conditions that depend on the investment horizon. In the present paper, we assume conditions (Assumptions 2.1–2.3) that are independent of the investment horizon, allowing us to apply Malliavin calculus techniques when the investment horizon tends to infinity.

The rest of this paper is organized as follows. The main results are discussed in Sect. 2, which consists of four subsections. In Sect. 2.1, we provide stochastic flow representations of optimal feedback strategies for general utilities. Sect. 2.2 estimates the convergence rate of the turnpike theorem for myopic portfolios with general factor models. Sect. 2.3 estimates the rate for excess hedging demands

with quadratic term structure models. In Sect. 2.4, we offer applications of our main results to optimal collective investment problems. Sect. 3 contains proofs of our main results. The appendix contains short reviews of Malliavin calculus, option pricing theory in stochastic factor models, the relationship between stochastic control methods and martingale duality methods, and the matrix Riccati equation.

2 Main results

2.1 Stochastic flow representation of optimal feedback functions

Let $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\in[0,T]})$ be a filtered probability space endowed with $(\mathcal{F}_t)_{t\in[0,T]}$, the augmentation of the natural filtration generated by the n-dimensional Brownian motion $W := (W^1, \ldots, W^n)^\top$. We consider a financial market with one riskless bond and n risky assets. The price processes of the riskless bond S^0 and n risky assets $S = (S^1, \ldots, S^n)^\top$ are modeled as

$$
\begin{aligned}
dS_t^0 &= S_t^0 r(Y_t)dt, & S_0^0 &= 1, \\
dS_t &= \mathrm{diag}(S_t)\{\mu(Y_t)dt + \sigma(Y_t)dW_t\}, & S_0 &= s_0 \in \mathbb{R}_{++}^n, \\
dY_t &= b(Y_t)dt + a(Y_t)dW_t, & Y_0 &= y \in \mathbb{R}^m,
\end{aligned}
\tag{2}
$$

where $r : \mathbb{R}^m \to \mathbb{R}, \mu : \mathbb{R}^m \to \mathbb{R}^n, \sigma : \mathbb{R}^m \to \mathbb{R}^{n\times n}, b : \mathbb{R}^m \to \mathbb{R}^m, a : \mathbb{R}^m \to \mathbb{R}^{m\times n}$ satisfy Assumption 2.1 below, $\mathrm{diag}(x)$ denotes the diagonal $n \times n$ matrix whose (i,i)-th element is component x_i of $x = (x_1, \ldots, x_n)^\top \in \mathbb{R}^n$, and the m-dimensional process $Y = (Y^1, \ldots, Y^m)^\top$ is referred to as the stochastic factor processes, which affect the coefficients of the asset price processes. We denote the market price of risk as $\theta(y) := \sigma(y)^{-1}\{\mu(y) - r(y)\mathbf{1}\}$, where $\mathbf{1} = (1, \ldots, 1)^\top \in \mathbb{R}^n$. Furthermore, we let $\beta = (\beta_t)_{t\in[0,T]}$ be a discounted price process, $Z = (Z_t)_{t\in[0,T]}$ be a likelihood ratio process, and $H = (H_t)_{t\in[0,T]}$ be a state price density process as follows:

$$
\begin{aligned}
\beta_t &:= \frac{1}{S_t^0} = \exp\left(-\int_0^t r(Y_s)ds\right), \\
Z_t &:= \exp\left\{-\int_0^t \theta^\top(Y_s)dW_s - \frac{1}{2}\int_0^t |\theta(Y_s)|^2 ds\right\}, \\
H_t &:= \beta_t Z_t.
\end{aligned}
$$

We assume some conditions on these coefficient functions.

Assumption 2.1. (i) All functions $r, \mu, \sigma, b, a, \theta$ on $\mathbb{R}^m$ are continuous functions.

(ii) $\sigma(y)$ is invertible for all $y \in \mathbb{R}^m$.

(iii) θ has linear growth.

(iv) b, a are Lipschitz continuous and a is bounded.

Under Assumption 2.1, a local martingale Z is a martingale. For the proof, see [27, Section 6.2]. Therefore, we can define a unique equivalent martingale measure $\mathbb{Q}$ and n-dimensional $\mathbb{Q}$-Brownian motion $W^{\mathbb{Q}}$ as follows:

$$
\begin{aligned}
d\mathbb{Q}|_{\mathcal{F}_t} &:= Z_t d\mathbb{P}|_{\mathcal{F}_t}, \\
W_t^{\mathbb{Q}} &:= W_t + \int_0^t \theta(Y_s)ds.
\end{aligned}
$$

$\pi = (\pi_t)_{t\in[0,T]}$ is a portfolio process if π is an n-dimensional $(\mathcal{F}_t)_t$-progressively measurable process and satisfies

$$
\int_0^T |\pi_t|^2 dt < \infty \quad P - a.s.
$$

For each initial wealth $x \geq 0$ and portfolio process π, we define the corresponding wealth process $X^{x,\pi} = (X_t^{x,\pi})_{t\in[0,T]}$ as the solution to the following SDE:

$$
dX_t^{x,\pi} = \left[r(Y_t)X_t^{x,\pi} + \pi_t^\top(\mu(Y_t) - r(Y_t)\mathbf{1})\right]dt + \pi_t^\top\sigma(Y_t)dW_t, \quad X_0 = x.
$$

3

Given $x \geq 0$, we say that a portfolio process π is admissible at x if the corresponding wealth process $X^{x,\pi}$ satisfies

$$X_t^{x,\pi} \geq 0, \quad t \in [0,T]$$

almost surely.

An investor's risk preference is represented by a utility function U.

Definition 2.1. We call $U : (0, \infty) \to \mathbb{R}$ a utility function if U is strictly increasing, strictly concave, and twice continuously differentiable on $(0, \infty)$ and satisfies the Inada conditions

$$\lim_{x \searrow 0} U'(x) = \infty, \quad \lim_{x \nearrow \infty} U'(x) = 0.$$

Let $I := (U')^{-1} : (0, \infty) \to (0, \infty)$ be the inverse marginal utility U'. By the definition of U, I is strictly decreasing and continuously differentiable on $(0, \infty)$ and satisfies

$$\lim_{z \searrow 0} I(z) = \infty, \quad \lim_{z \nearrow \infty} I(z) = 0.$$

The investor in this paper desires to maximize their expected utility and find an optimal portfolio process $\hat{\pi}$. This problem is formulated as follows.

Problem 2.2. Find an optimal $\pi \in \mathcal{A}(x)$ for the problem

$$\sup_{\pi \in \mathcal{A}(x)} \mathbb{E}\left[U(X_T^{x,\pi})\right]$$

of maximizing expected utility from terminal wealth, where

$$\mathcal{A}(x) := \left\{ \pi ; \pi \text{ is admissible at } x, \mathbb{E}\left[U(X_T^{x,\pi})^-\right] < \infty \right\}.$$

To use the martingale method, we assume the following growth conditions.

Assumption 2.2. (i) There exist $r_0 \in \mathbb{R}$, $r_1 \in \mathbb{R}^m$ such that for every $y \in \mathbb{R}^m$,

$$r(y) \geq r_0 + r_1^\top y.$$

(ii) There exist $\kappa > 0$, $\rho \in (0, 1]$ such that for every $z > 0$,

$$I(z) \leq \kappa(1 + z^{-\rho}).$$

Under Assumptions 2.1 and 2.2, we can use the martingale method as in [20, Theorem 3.7.6]. For the proof, see Sect. 3.1.

Theorem 2.1. *Under Assumptions 2.1 and 2.2, for each $x > 0$, there exists $\hat{\lambda} > 0$ such that $x = \mathbb{E}[H_T I(\hat{\lambda} H_T)]$. The optimal terminal wealth ξ and the optimal wealth process $\hat{X} = (\hat{X}_t)_{t \in [0,T]}$ are*

$$\xi = I(\hat{\lambda} H_T), \quad \hat{X}_t = \frac{1}{H_t} \mathbb{E}_t\left[H_T I(\hat{\lambda} H_T)\right].$$

Moreover, the optimal portfolio process $\hat{\pi} = (\hat{\pi}_t)_{t \in [0,T]}$ is given by

$$\hat{\pi}_t = (\sigma^\top(Y_t))^{-1}\left(\frac{\psi_t}{H_t} + X_t \theta(Y_t)\right),$$

where ψ is the integrand in the stochastic integral representation $M_t = x + \int_0^t (\psi_u)^\top dW_u$ of the martingale $\left(\mathbb{E}_t[H_T I(\hat{\lambda} H_T)]\right)_{t \in [0,T]}$.

To derive explicit formulas for ψ, we assume regularity and growth conditions for the market coefficients, r, θ, b, a, and the derivative of I, which enable us to apply the results of the Malliavin calculus.

4

Assumption 2.3. (i) r, θ are continuously differentiable and of polynomial growth, and their Jacobian matrices $Dr, D\theta$ are also of polynomial growth.

(ii) b, a are continuously differentiable, and the Jacobian matrices $Db, Da_{\cdot l} : \mathbb{R}^m \to \mathbb{R}^{m \times m}$ are bounded, where $Db = \left(\dfrac{\partial b_i}{\partial x_j} \right)_{\substack{1 \le i \le m \\ 1 \le j \le m}}$, $Da_{\cdot l} = \left(\dfrac{a_{i,l}}{\partial x_j} \right)_{\substack{1 \le i \le m \\ 1 \le j \le m}}$, $(l = 1, \dots, n)$.

(iii) There exists $K > 0$ such that for any $y, z_1, \dots, z_n \in \mathbb{R}^m$,

$$y^\top \left(b(y) - a(y)\theta(y) \right) + \sum_{j=1}^n z_j^\top \left(Db(y) - \sum_{l=1}^n Da_{\cdot l}(y)\theta^l(y) \right) z_j \le K \left(1 + |y|^2 + \sum_{j=1}^n |z_j|^2 \right).$$

(iv) $|zI'(z)| \le \kappa(1 + z^{-\rho})$ for some $\kappa > 0$, $\rho \in (0, 1]$.

For Theorem 2.2, we prepare some notation. Let a pair (Y, Z) of $\mathbb{R}^m$-valued stochastic process Y and $\mathbb{R}^{m \times m}$-valued stochastic process Z be the solution to the following system of SDEs:

$$dY_s = b(Y_s)ds + a(Y_s)dW_s, \qquad\qquad Y_t = y,$$
$$dZ_s = Db(Y_s)Z_s ds + \sum_{j=1}^n Da_{\cdot j}(Y_s)Z_s dW_s^j, \qquad\qquad Z_t = I.$$

Then a pair (Y, Z) is a Markov process and $(Y^{(t,y)}, Z^{(t,y)})$ denotes the solution to the above system of SDEs when (Y, Z) starts from $(y, I) \in \mathbb{R}^m \times \mathbb{R}^{m \times m}$ at time 0. Note that $Z^{(t,y)}$ always starts from the identity matrix $I \in \mathbb{R}^{m \times m}$ and $Y_s^{(0,y)} = Y_s^{(t, Y_t^{(0,y)})}$ for $s \in [t, T]$. Because $Z^{(t,y)}$ can be thought of as the derivative of $Y^{(t,y)}$ with respect to the initial value y, we use the notation $\nabla_y Y^{(t,y)} := Z^{(t,y)}$ instead of $Z^{(t,y)}$. Furthermore, let $H^{(t,y)} = \left(H_s^{(t,y)} \right)_{s \in [t,T]}$ be a state price density process that starts at time t and is given by

$$H_s^{(t,y)} := \exp \left(- \int_t^s r(Y_u^{(t,y)})du - \int_t^s \theta^\top(Y_u^{(t,y)})dW_u - \frac{1}{2} \int_t^s |\theta(Y_u^{(t,y)})|^2 du \right).$$

Let $\nabla_y H^{(t,y)} = \left(\nabla_y H_s^{(t,y)} \right)_{s \in [t,T]}$, $L^{(t,y)} = \left(L_s^{(t,y)} \right)_{s \in [t,T]}$ be $\mathbb{R}^m$-valued stochastic processes given by

$$L_s^{(t,y)} := \int_t^s (Dr(Y_u^{(t,y)})\nabla_y Y_u^{(t,y)})^\top du + \int_t^s \left(D\theta(Y_u^{(t,y)})\nabla_y Y_u^{(t,y)} \right)^\top dW_u$$
$$+ \int_t^s \left(D\theta(Y_u^{(t,y)})\nabla_y Y_u^{(t,y)} \right)^\top \theta(Y_u^{(t,y)})du,$$
$$\nabla_y H_s^{(t,y)} := -H_s^{(t,y)} L_s^{(t,y)}.$$

We can also consider $\nabla_y H^{(t,y)}$ as the derivative of $H^{(t,y)}$ with respect to y. When Y starts from y at time 0, we drop the superscripts $(0, y)$.

Theorem 2.2. *Under Assumptions 2.1–2.3, the optimal portfolio process $\hat{\pi}$ in Theorem 2.1 can be represented by*

$$\hat{\pi}_t = -(\sigma^\top(Y_t))^{-1}\theta(Y_t)\frac{1}{H_t}\mathbb{E}_t \left[H_T \cdot \hat{\lambda} H_T I'(\hat{\lambda} H_T) \right]$$
$$+ (\sigma^\top(Y_t))^{-1}a^\top(Y_t)\mathbb{E}_t \left[\nabla_y H_T^{(t,Y_t)} \left(I(\lambda \hat{H}_T) + \hat{\lambda} H_T I'(\hat{\lambda} H_T) \right) \right].$$

Remark 2.1. (i) We can find similar results in [6, 10, 24, 37]. Here, we state the differences between our arguments and those of the previous papers. Firstly, [6] considers the same market model as ours and assumes that $H_T I(zH_T) \in \mathbb{D}_{2,1}$, which is difficult to check in general. In the present paper, we check that $H_T I(zH_T) \in \mathbb{D}_{1,1}$, which is enough to use Clark's formula (Proposition A.1), with the growth conditions for an inverse marginal utility I, its derivative I', and market coefficients (Assumptions 2.1–2.3). Secondly, [24, 37] check the conditions that $H_T I(zH_T) \in \mathbb{D}_{1,1}$ with

5

linear Gaussian dynamics for the drift process μ_t under partial information. Their assumptions for market coefficients depend on the investment horizon T, but our assumptions do not depend on T, which is more useful for proving the turnpike theorem. Lastly, [10] uses stochastic flow techniques instead of Malliavin calculus. Because the assumptions in [10] also depend on T and are difficult to check for our model, we do not use the results of [10].

(ii) We can represent the optimal portfolio process $\hat{\pi}$ by *the Arrow–Pratt measure of absolute risk tolerance* $ART_U(x) := -\dfrac{U'(x)}{U''(x)}$ as in [6]:

$$\hat{\pi}_t = (\sigma^\top(Y_t))^{-1}\theta(Y_t)\frac{1}{H_t}\mathbb{E}_t\left[H_T \cdot ART_U(\hat{X}_T)\right]$$
$$+ (\sigma^\top(Y_t))^{-1}a^\top(Y_t)\mathbb{E}_t\left[\nabla_y H_T^{(t,Y_t)}\left(\hat{X}_T - ART_U(\hat{X}_T)\right)\right].$$

Moreover, the optimal portfolio $\hat{\pi}$ can be divided into two components, namely, the myopic portfolio $\hat{\pi}^M$ and the excess hedging demand $\hat{\pi}^H$:

$$\hat{\pi}_t = \hat{\pi}_t^M + \hat{\pi}_t^H,$$
$$\hat{\pi}_t^M = (\sigma^\top(Y_t))^{-1}\theta(Y_t)\frac{1}{H_t}\mathbb{E}_t\left[H_T \cdot ART_U(\hat{X}_T)\right],$$
$$\hat{\pi}_t^H = (\sigma^\top(Y_t))^{-1}a^\top(Y_t)\mathbb{E}_t\left[\nabla_y H_T^{(t,Y_t)}\left(\hat{X}_T - ART_U(\hat{X}_T)\right)\right].$$

By the Markov property of the market model, the optimal portfolio process $\hat{\pi}$ obtained in Theorem 2.2 is given by a feedback form.

Proposition 2.3. *Under Assumptions 2.1–2.3, the optimal portfolio process $\hat{\pi}$ in Theorem 2.1 can be represented by a feedback form*

$$\hat{\pi}_t = \hat{\pi}(T - t, \hat{X}_t, Y_t), \quad t \in [0, T],$$

where $\hat{\pi} : (0, \infty) \times (0, \infty) \times \mathbb{R}^m \to \mathbb{R}^n$ is defined by

$$\hat{\pi}(\tau, x, y) := \hat{\pi}^M(\tau, x, y) + \hat{\pi}^H(\tau, x, y),$$
$$\hat{\pi}^M(\tau, x, y) := -(\sigma^\top(y))^{-1}\theta(y)\mathbb{E}^y\left[H_\tau \cdot \hat{\lambda}H_\tau I'(\hat{\lambda}H_\tau)\right],$$
$$\hat{\pi}^H(\tau, x, y) := (\sigma^\top(y))^{-1}a^\top(y)\mathbb{E}^y\left[\nabla_y H_\tau\left(I(\hat{\lambda}H_\tau) + \hat{\lambda}H_\tau I'(\hat{\lambda}H_\tau)\right)\right],$$

and $\hat{\lambda} = \hat{\lambda}(\tau, x, y)$ is defined by an equality $x = \mathbb{E}^y[H_\tau I(\lambda H_\tau)]$.

2.2 Turnpike theorem for myopic portfolios under general stochastic factor models

In this subsection, we consider two investors with utility functions $U_i : (0, \infty) \to (-\infty, \infty)$, $(i = 1, 2)$, and we fix an initial wealth $x > 0$ for both investors. We denote the corresponding optimal terminal wealth ξ, optimal wealth process $\hat{X}$, and optimal feedback function $\hat{\pi}$ for the i-th investor by $\xi^{i,T}, \hat{X}^{i,T}$, and $\hat{\pi}^i$, respectively. In this subsection, we consider the turnpike theorem for myopic portfolios under general stochastic factor models (2) introduced in the previous subsection. That is, we show that for each x and y,
$$|\hat{\pi}^{1,M}(T, x, y) - \hat{\pi}^{2,M}(T, x, y)| \to 0, \quad (T \nearrow \infty),$$
and we derive its convergence rate in terms of $\mathbb{E}[H_T]$.

Assumption 2.4. (i) Let $p \in (-\infty, 0]$. For $p < 0$, we set

$$U_2(x) := \frac{x^p}{p}, \quad x \in (0, \infty),$$

for $p = 0$,

$$U_2(x) := \log x, \quad x \in (0, \infty).$$

(ii) Let $q := \dfrac{p}{p-1} \in [0,1)$. There exist constants $K \in [0,\infty), \alpha \in (q-1,0]$ such that

$$|I_1(z) - I_2(z)| \le K(1 + z^\alpha), \quad z \in (0,\infty), \tag{3}$$
$$|zI_1'(z) - zI_2'(z)| \le K(1 + z^\alpha), \quad z \in (0,\infty). \tag{4}$$

(iii)
$$\mathbb{E}[H_T] \searrow 0, \quad (T \nearrow \infty).$$

Remark 2.2. Because the von Neumann–Morgenstern (vN-M) utility function U_1 for the first investor is uniquely determined up to positive affine transformations [11, Theorem 2.21], any conditions for vN-M utility function U_1 must be invariant for positive affine transformations. In this case, we have to assume a generalized version of Assumption 2.4(ii) as follows.

Assumption. There exist constants $C > 0, K \ge 0, \alpha \in (q-1,0]$ such that

$$
\begin{aligned}
|I_1(z) - CI_2(z)| &\le K(1 + z^\alpha), \quad z \in (0,\infty), \\
|zI_1'(z) - CzI_2'(z)| &\le K(1 + z^\alpha), \quad z \in (0,\infty).
\end{aligned}
\tag{5}
$$

If U_1 satisfies (5), then $\tilde{U}_1 := aU_1 + b$ also satisfies (5) for any $a > 0$, $b \in \mathbb{R}$, which means that (5) is invariant for positive affine transformations. In particular, if U_1 satisfies (5), then $\tilde{U}_1 := aU_1$, $a := C^{\frac{1}{q-1}}$ satisfies (5) for $C = 1$ and the optimal feedback function given by Proposition 2.3 is invariant, which means that without loss of generality, we can assume $C = 1$ in (5), that is, (ii) in Assumption 2.4.

Remark 2.3. If a function U is twice differentiable and satisfies (3) in Assumption 2.4(ii) for $p \in (-\infty, 1)$, then U' is regularly varying with an exponent $p - 1$ and

$$\lim_{x \nearrow \infty} RRA_U(x) = \lim_{x \nearrow \infty} \frac{-U''(x)x}{U'(x)} = 1 - p.$$

For a proof, see [5, Proposition 2]. Furthermore, by the identity $I_2(z) = z^{q-1}$, the ratio of the inverse marginal utilities, $\dfrac{I_1}{I_2}$, and the ratio of the derivatives, $\dfrac{I_1'}{I_2'}$, satisfy

$$
\left| \frac{I_1(z)}{I_2(z)} - 1 \right| \le K \left(z^{-(q-1)} + z^{\alpha - (q-1)} \right),
$$
$$
\left| \frac{I_1'(z)}{I_2'(z)} - 1 \right| \le \frac{K}{1-q} \left(z^{-(q-1)} + z^{\alpha - (q-1)} \right).
$$

These inequalities imply that these ratios converge to 1, and the speed of convergence is determined by $\alpha - (q-1)$.

Remark 2.4. Although the inequality (3) can be derived from (4), we assume both inequalities for simplicity.

Remark 2.5. When showing turnpike theorems, Assumption 2.4(iii) is the usual one in previous works. Indeed, Dybvig et al. [9] assume the same condition. Furthermore, Cox and Huang [5], Jin [19], Huang and Zariphopoulou[17], and Bian and Zheng [3, 4] consider the Black–Scholes model and assume that the interest rate r is strictly positive, which is equivalent to $\mathbb{E}[H_T] \searrow 0$ in the model. Because

$$\mathbb{E}[H_T] = \mathbb{E}^{\mathbb{Q}} \left[\exp \left(- \int_0^T r(Y_t) dt \right) \right]$$

is the price of a zero-coupon bond with maturity T at $t = 0$, Assumption 2.4(iii) implies that the interest rate $r(Y_t)$ is positive in the long run.

The following theorem is one of our main results.

Theorem 2.4. *Under Assumptions 2.1–2.4, there exists an $M = M(x, y) \in (-\infty, x]$ that is independent of T such that*

$$|\hat{\pi}^{1,M}(T, x, y) - \hat{\pi}^{2,M}(T, x, y)| \leq K(2-q)|\sigma^\top(y)^{-1}\theta(y)| \left(\mathbb{E}[H_T] + (x-M)^{\frac{\alpha}{q-1}} \mathbb{E}[H_T]^{1-\frac{\alpha}{q-1}} \right)$$

$$= O \left(\mathbb{E}[H_T]^{1-\frac{\alpha}{q-1}} \right), \quad (T \nearrow \infty).$$

Remark 2.6. Here we want to emphasize that the convergence rate of the turnpike theorem in stochastic factor models is determined by two components: (i) the speed of market growth, $\mathbb{E}[H_T]$, and (ii) the similarity between utilities captured by α and q. In addition, if the interest rate is a positive constant, $r(Y_t) = r > 0$, then the convergence rate is $e^{-r\left(1-\frac{\alpha}{q-1}\right)T}$, which is the same rate as in Bian and Zheng [3]. Therefore, the rate $\mathbb{E}[H_T]^{1-\frac{\alpha}{q-1}}$ derived in the present paper is a natural extension of the rate in [3].

Remark 2.7. As studied by Qin and Linetsky [38], the decay speed of $\mathbb{E}[H_T]$ is exponential in general. Indeed, (ii) of Theorem 3.2 in [38] shows that under a general semimartingale model satisfying some assumptions,

$$\lim_{T \nearrow \infty} \frac{-\log \mathbb{E}[H_T]}{T} = \lambda$$

holds for some $\lambda \in \mathbb{R}$. When all uncertainty is generated by a time-homogeneous Markov process X, $e^{\lambda t}$ is an eigenvalue of the pricing operator $\mathcal{P}_t f(x) := \mathbb{E}[H_t f(X_t)|X_0 = x]$. For details, see [38, 39, 40].

Although Theorem 2.4 seems to imply that the convergence is not uniform in x, we can prove uniform convergence in x for portfolio proportions $\dfrac{\hat{\pi}^{i,M}(T, x, y)}{x}$.

Theorem 2.5. *Under Assumptions 2.1–2.4, for any $\epsilon > 0$,*

$$\sup_{x > \epsilon} \left| \frac{\hat{\pi}^{1,M}(T, x, y)}{x} - \frac{\hat{\pi}^{2,M}(T, x, y)}{x} \right| = O \left(\mathbb{E}[H_T]^{1-\frac{\alpha}{q-1}} \right), \quad (T \nearrow \infty).$$

Remark 2.8. We cannot prove uniform convergence in y because $y \mapsto \theta(y)$ has linear growth and is generally unbounded.

We can also show that the time-0 value of the difference between the optimal wealth processes at time t converges to 0 uniformly in t and that the convergence rate is the same as in the above theorems.

Theorem 2.6. *Under Assumptions 2.1–2.4,*

$$\sup_{t \in [0,T]} \mathbb{E}[H_t|\hat{X}_t^{1,T} - X_t^{2,T}|] = O \left(\mathbb{E}[H_T]^{1-\frac{\alpha}{q-1}} \right), \quad (T \nearrow \infty). \tag{6}$$

Remark 2.9. The convergence $\lim_{T \nearrow \infty} \mathbb{E}[H_t|\hat{X}_t^{1,T} - X_t^{2,T}|] = 0$ is already proved in complete markets by Dybvig et al. [9]. Here, our focus is on the convergence rate.

2.3 Turnpike theorem for excess hedging demands under quadratic term structure models

In this subsection, we consider quadratic term structure models studied by [1, 25]. The market model consists of a riskless bond S^0, n risky assets $S = (S^1, \ldots, S^n)^\top$, and an m-dimensional factor process $Y = (Y^1, \ldots, Y^m)^\top$ that affects the risk-free interest rate $r(Y_t)$ of S^0 and the mean return rate $\mu(Y_t)$ of S:

$$\begin{aligned} dS_t^0 &= S_t^0 r(Y_t)dt, & S_0^0 &= 1, \\ dS_t &= \text{diag}(S_t) \left\{ \mu(Y_t)dt + \Sigma dW_t \right\}, & S_0 &\in \mathbb{R}_{++}^n, \\ dY_t &= (b + BY_t)dt + \Lambda dW_t, & Y_0 &= y \in \mathbb{R}^m, \end{aligned} \tag{7}$$

where $r : \mathbb{R}^m \to \mathbb{R}$, $\mu : \mathbb{R}^m \to \mathbb{R}^n$, $\Sigma \in \mathbb{R}^{n \times n}$, $b \in \mathbb{R}^m$, $B \in \mathbb{R}^{m \times m}, \Lambda \in \mathbb{R}^{m \times n}$. Furthermore, we assume that the risk-free interest rate $r(Y_t)$ of S^0 is a quadratic Gaussian process and the market price

of risk $\theta(Y_t) := \Sigma^{-1}\left(\mu(Y_t) - \mathbf{1}r(Y_t)\right)$ is a linear Gaussian process as follows:

$$
\begin{aligned}
r(y) &:= r_0 + r_1^\top y + \frac{1}{2}y^\top R_2 y, \\
\theta(y) &:= a + Ay, \\
\mu(y) &:= \Sigma\theta(y) + \mathbf{1}r(y),
\end{aligned}
\tag{8}
$$

where $r_0 \in \mathbb{R}$, $r_1 \in \mathbb{R}^m$, $R_2 \in \mathbb{R}^{m\times m}$, $a \in \mathbb{R}^n$, $A \in \mathbb{R}^{n\times m}$, and $\mathbf{1} = (1,\ldots,1)^\top \in \mathbb{R}^n$. We denote the totality of $m \times m$, real, symmetric matrices by $\mathbb{S}^m$ and $\mathbb{S}^m_+ := \{M \in \mathbb{S}^m;\ M \geq 0\}$, $\mathbb{S}^m_{++} := \{M \in \mathbb{S}^m;\ M > 0\}$. We assume the following conditions.

Assumption 2.5. (i) Σ is invertible.

(ii) $R_2 \in \mathbb{S}^m_+$.

(iii) $R_2 = 0$ or $\left(\gamma(1-\gamma)A^\top A + \gamma R_2\right) \in \mathbb{S}^m_{++}$ for $\gamma \in \{q, 1+\alpha\}$.

(iv) B is stable; that is, all its eigenvalues have negative real parts.

Under Assumptions 2.4 and 2.5, the quadratic term structure model given by (7) and (8) satisfies Assumptions 2.1–2.3. Therefore, all main results in Sects. 2.1 and 2.2 are valid.

Remark 2.10. Assumption 2.5(iv) implies that Y is a multivariate Ornstein–Uhlenbeck process. In particular, the model includes well-known short-rate models, such as the Vasicek model and special versions of the CIR model; for details, see [1].

Remark 2.11. We restrict our analysis to the quadratic term structure model given by (7) and (8) for the following reason. As Proposition 3.11 says, to estimate the rates for excess hedging demands, we have to compute the asymptotic behavior of the stochastic factor process Y under myopic probability measures $\mathbb{Q}_T^\gamma$, $\gamma \in [0,1]$, defined by

$$
d\mathbb{Q}_T^\gamma := \frac{H_T^\gamma}{\mathbb{E}[H_T^\gamma]}d\mathbb{P}.
$$

When $\gamma = 1$, the myopic probability measure $\mathbb{Q}_T^1$ is a T-forward measure. When $\gamma \in (0,1)$, the measures are given by the optimal wealth processes $\hat{X}_T^T$ for CRRA investors:

$$
d\mathbb{Q}_T^\gamma = \frac{U(\hat{X}_T^T)}{\mathbb{E}\left[U(\hat{X}_T^T)\right]}d\mathbb{P},
$$

where, $U(x) := \dfrac{x^p}{p}$, $p := \dfrac{\gamma}{\gamma-1}$. As Guasoni et al. [13] say, by using results on CRRA utility maximization problems, we can represent the probability density processes of $\mathbb{Q}_T^\gamma$ as stochastic exponential martingales in terms of the optimal portfolio processes. Therefore, in the quadratic term structure model given by (7) and (8), we can analyze the asymptotic behavior of Y under $\mathbb{Q}_T^\gamma$ by using the asymptotic behavior of the solutions to Riccati differential equations. Because the optimal portfolios in nonlinear stochastic factor models (2) are given by solutions of semilinear PDEs (see Nagai [33]), our analysis will require more-advanced techniques, and further research for nonlinear stochastic factor models will be addressed in future work.

Under these assumptions, we can also derive the convergence rates for excess hedging demands.

Theorem 2.7. *Under Assumptions 2.4 and 2.5,*

$$
|\hat{\pi}^{1,H}(T,x,y) - \hat{\pi}^{2,H}(T,x,y)| = O\left(\mathbb{E}[H_T]^{1-\frac{\alpha}{q-1}}\right), \quad (T \nearrow \infty).
$$

By combining the above theorem with Theorem 2.4, we obtain the convergence rates for the optimal feedback functions $\hat{\pi}^1 = \hat{\pi}^{1,M} + \hat{\pi}^{1,H}$.

Corollary 2.7.1. *Under Assumptions 2.4 and 2.5,*

$$
|\hat{\pi}^1(T,x,y) - \hat{\pi}^2(T,x,y)| = O\left(\mathbb{E}[H_T]^{1-\frac{\alpha}{q-1}}\right), \quad (T \nearrow \infty).
$$

Moreover, by considering portfolio proportions rather than dollar amounts, we can prove the uniform turnpike theorem for excess hedging demands and obtain its convergence rate.

Theorem 2.8. *Under Assumptions 2.4 and 2.5, for any $\epsilon > 0$,*

$$\sup_{x > \epsilon} \left| \frac{\hat{\pi}^{1,H}(T,x,y)}{x} - \frac{\hat{\pi}^{2,H}(T,x,y)}{x} \right| = O\left(\mathbb{E}[H_T]^{1-\frac{\alpha}{q-1}} \right), \quad (T \nearrow \infty).$$

Corollary 2.8.1. *Under Assumptions 2.4 and 2.5, for any $\epsilon > 0$,*

$$\sup_{x > \epsilon} \left| \frac{\hat{\pi}^{1}(T,x,y)}{x} - \frac{\hat{\pi}^{2}(T,x,y)}{x} \right| = O\left(\mathbb{E}[H_T]^{1-\frac{\alpha}{q-1}} \right), \quad (T \nearrow \infty).$$

2.4   Applications: optimal collective investment problems

In this subsection, we offer applications of our main results to optimal collective investment problems. For detailed descriptions of the problems, see [2, 18], for example. We consider the quadratic term structure model, given by (7) and (8), that satisfies Assumptions 2.4(iii) and 2.5 as described in Sect. 2.3. We assume there are n investors with CRRA utility functions $U_i(x) := \dfrac{x^{p_i}}{p_i}$, $(i = 1, \ldots, n)$, $-\infty < p_1 < \ldots p_n \leq 0$, where the relative risk aversion level of the i-th investor is given by $\gamma_i := 1 - p_i$. At the beginning of the investment period $(t = 0)$, investors delegate their investment management to a fund manager. At the end of the investment period $(t = T)$, a fund manager distributes the aggregate terminal wealth among n investors according to a sharing rule. In this paper, we consider two well-known sharing rules: a Pareto optimal sharing rule and a linear sharing rule.

2.4.1   Pareto optimal sharing rule

We assume that a fund manager chooses a Pareto optimal distribution of the terminal wealth, which is represented by

$$\tilde{U}(x) := \max \left\{ \sum_{i=1}^{n} \beta_i U_i(x_i) \,\middle|\, x_i \in \mathbb{R}, \; \sum_{i=1}^{n} x_i = x \right\},$$

where β_i is the weight granted to the i-th investor and satisfies $\beta_i > 0$, $(i = 1, \ldots, n)$, $\sum_{i=1}^{n} \beta_i = 1$.

[2] uses this utility function to analyze collective investment problems. $\tilde{U}$ also appears as a utility function for a representative agent in the context of market equilibrium [20, Chapter 4]. Because the operations of sup-convolution and addition are dual to each other, the inverse marginal utility of a fund manager is given by the sum of those of individuals.

Lemma 2.9. *$\tilde{U}$ is a utility function in Definition 2.1 and the inverse marginal utility $\tilde{I} := (\tilde{U})^{-1}$ is given by*

$$\tilde{I}(z) = \sum_{i=1}^{n} I_i\left(\frac{z}{\beta_i} \right), \quad I_i(z) = z^{q_i - 1}, \quad q_i := \frac{p_i}{p_i - 1}.$$

Proof. See [2, Appendix A]. □

By using our main results in Sect. 2.3, we can show that the optimal feedback function of a fund manager converges to that of the most risk-seeking investor. The convergence rate is determined by the price of a zero-coupon bond and the relative difference between the relative risk aversion levels of the least risk-averse investor $(i = n)$ and the second least risk-averse investor $(i = n - 1)$.

Theorem 2.10. *Let $\hat{\pi}^P$ be the optimal feedback function of a fund manager with the Pareto optimal sharing rule and $\hat{\pi}^i$ be that of the i-th investor. Then*

$$\left| \hat{\pi}^P(T,x,y) - \hat{\pi}^n(T,x,y) \right| = O\left(\mathbb{E}[H_T]^{\frac{\gamma_{n-1} - \gamma_n}{\gamma_{n-1}}} \right), \quad (T \nearrow \infty).$$

Proof. It is enough to check that the inequalities (5) hold for $I_1(z) = \tilde{I}(z)$, $I_2(z) = z^{q_n-1}$, $\alpha = q_{n-1} - 1$, $C = \left(\dfrac{1}{\beta_n}\right)^{q_n-1}$. Because $0 \leq q_n < q_{n-1} < \cdots < q_1 < 1$, there exists some $K > 0$ such that for all $z > 0$,

$$\left| \tilde{I}(z) - \left(\frac{z}{\beta_n}\right)^{q_n-1} \right| = \sum_{i=1}^{n-1} \left(\frac{z}{\beta_i}\right)^{q_i-1}$$
$$\leq K(1 + z^{q_{n-1}-1})$$

and

$$\left| z\tilde{I}'(z) - (q_n - 1)\left(\frac{z}{\beta_n}\right)^{q_n-1} \right| = \sum_{i=1}^{n-1}(1 - q_i)\left(\frac{z}{\beta_i}\right)^{q_i-1}$$
$$\leq K(1 + z^{q_{n-1}-1})$$

hold. Moreover, the identity

$$\frac{\alpha}{q-1} = \frac{q_{n-1} - 1}{q_n - 1} = \frac{1 - p_n}{1 - p_{n-1}} = \frac{\gamma_n}{\gamma_{n-1}}$$

gives the convergence rate. $\qquad\square$

2.4.2 Linear sharing rule

Here, we assume that a fund manager allocates the terminal wealth according to a linear sharing rule, where the i-th investor receives a fixed proportion α_i of the terminal wealth. This linear sharing rule is represented by

$$U(x) := \sum_{i=1}^{n} \beta_i U_i(\alpha_i x),$$

where α_i, $\beta_i > 0$, $(i = 1, \ldots, n)$, $\sum_{i=1}^{n} \alpha_i = \sum_{i=1}^{n} \beta_i = 1$. Although U seems simpler than $\tilde{U}$, we cannot derive the inverse marginal utility of U in general. Therefore, to estimate the differences between inverse marginal utilities and their derivatives, we perform complex calculations; for the proof, see Sect. 3.4.

Proposition 2.11. *Let*

$$U_1(x) := \sum_{i=1}^{n} \beta_i \frac{(\alpha_i x)^{p_i}}{p_i} = \sum_{i=1}^{n} w_i \frac{x^{p_i}}{p_i},$$
$$U_2(x) := w_n \frac{x^{p_n}}{p_n},$$

where $w_i := \beta_i \alpha_i^{p_i} > 0$. For any nonnegative $\beta \in (1 + p_{n-1} - p_n, 1)$, there exists $K > 0$ such that for all $z > 0$,

$$|I_1(z) - I_2(z)| \leq K(1 + z^{\beta(q_n-1)}), \tag{9}$$
$$|zI_1'(z) - zI_2'(z)| \leq K(1 + z^{\beta(q_n-1)}) \tag{10}$$

hold.

Theorem 2.12. *Let $\hat{\pi}^L$ be the optimal feedback function of a fund manager with the linear sharing rule and $\hat{\pi}^i$ be that of the i-th investor. Then for any nonnegative $\beta \in (1 + p_{n-1} - p_n, 1)$,*

$$\left| \hat{\pi}^L(T, x, y) - \hat{\pi}^n(T, x, y) \right| = O\left(\mathbb{E}[H_T]^{1-\beta}\right), \quad (T \nearrow \infty)$$

holds.

11

Proof. By Proposition 2.11 and the main results in Sects. 2.2 and 2.3, the convergence rate is given by

$$\mathbb{E}[H_T]^{1-\frac{\beta(q_n-1)}{q_n-1}} = \mathbb{E}[H_T]^{1-\beta}.$$

$\square$

Remark 2.12. If $1+p_{n-1}-p_n \geq 0$, which is equivalent to $\gamma_{n-1} \leq \gamma_n+1$ in terms of relative risk aversion levels γ_i, then the convergence rate is given by $O(\mathbb{E}[H_T]^{\gamma_{n-1}-\gamma_n-\epsilon})$ for any $\epsilon > 0$. If $1+p_{n-1}-p_n < 0$, which is equivalent to $\gamma_{n-1} > \gamma_n+1$, then the convergence rate is given by $O(\mathbb{E}[H_T])$. These facts mean that for a fund manager with a linear sharing rule, the convergence rate is determined by $\gamma_{n-1} - \gamma_n$, the absolute difference between the relative risk aversion levels of the least risk-averse investor and the second least risk-averse investor. On the other hand, for a fund manager with a Pareto optimal sharing rule, the convergence rate is determined by $\dfrac{\gamma_{n-1} - \gamma_n}{\gamma_{n-1}}$, the relative difference between the relative risk aversion levels for the same investors. In particular, because we consider the case $\gamma_{n-1} > 1$, the convergence rate under a linear sharing rule is faster than that under a Pareto optimal sharing rule.

3   Proofs for main results

3.1   Proofs for Sect. 2.1

The following lemma is used in the proof of Theorem 2.1 and Lemma 3.6, which shows $H_T \in \mathbb{D}_{1,1}$.

Lemma 3.1. *Under Assumptions 2.1 and 2.2, for any $T > 0$, $\lambda \geq 0$, $y \in \mathbb{R}^m$,*

$$\mathbb{E}_y^{\mathbb{Q}}\left[\exp\left(-\lambda \int_0^T r(Y_t)dt\right)\right] < \infty.$$

Proof. We fix $T > 0$, $y \in \mathbb{R}^m$ and for any $\lambda \geq 0$,

$$\begin{aligned}
\mathbb{E}_y^{\mathbb{Q}}\left[\exp\left(-\lambda \int_0^T r(Y_t)dt\right)\right] &\leq \mathbb{E}_y^{\mathbb{Q}}\left[\frac{1}{T}\int_0^T \exp\left(-\lambda T r(Y_t)\right)dt\right] \\
&\leq \sup_{t\in[0,T]} \mathbb{E}_y^{\mathbb{Q}}\left[\exp\left(-\lambda T r(Y_t)\right)\right] \\
&\leq \sup_{t\in[0,T]} \mathbb{E}_y^{\mathbb{Q}}\left[\exp\left(-\lambda T\left(r_0 + r_1^\top Y_t\right)\right)\right] \\
&= \exp\left(-\lambda T r_0\right) \sup_{t\in[0,T]} \mathbb{E}_y^{\mathbb{Q}}\left[\exp\left(-\lambda T r_1^\top Y_t\right)\right]
\end{aligned}$$

holds, where the first inequality follows from Jensen's inequality, and the third inequality follows from Assumption 2.2(i). Therefore, it suffices to prove that for any $\lambda \geq 0$,

$$\sup_{t\in[0,T]} \mathbb{E}_y^{\mathbb{Q}}\left[\exp\left(\lambda|Y_t|\right)\right]$$

holds. Because Y satisfies the following SDE under $\mathbb{Q}$,

$$Y_t = y + \int_0^t \tilde{b}(Y_s)ds + \int_0^t a(Y_s)dW_s^{\mathbb{Q}},$$

and $\tilde{b}(y) = b(y) - a(y)\theta(y)$ is of linear growth, $|Y_t|$ satisfies

$$|Y_t| \leq |y| + KT + \sup_{t\in[0,T]} |M_t| + \int_0^t K|Y_s|ds,$$

where K is a some constant and $M_t := \int_0^t a(Y_s)dW_s^{\mathbb{Q}}$. By Gronwall's inequality,

$$|Y_t| \leq \left(|y| + KT + \sup_{t\in[0,T]} |M_t|\right)\exp(KT)$$

12

holds, which leads to

$$\exp\left(\lambda|Y_t|\right) \leq \exp\left\{\lambda\left(|y| + KT + \sup_{t\in[0,T]}|M_t|\right)\exp(KT)\right\}$$

$$= C_1\exp\left(C_2\sup_{t\in[0,T]}|M_t|\right) \quad (C_1 := \exp\{\lambda(|y| + KT)\exp(KT)\}, \quad C_2 := \lambda\exp(KT))$$

$$\leq C_1\exp\left(C_2\sum_{i=1}^{m}\sup_{t\in[0,T]}|M_t^i|\right)$$

$$\leq \sum_{i=1}^{m}C_1\exp\left(mC_2\sup_{t\in[0,T]}|M_t^i|\right).$$

As a result, it suffices to show that for any $\lambda \geq 0$, $i = 1, \ldots, m$,

$$\mathbb{E}_y^{\mathbb{Q}}\left[\exp\left(\lambda\sup_{t\in[0,T]}|M_t^i|\right)\right] < \infty$$

holds. The Dambis–Dubins–Schwarz theorem implies that there exists a Brownian motion β^i such that $M_t^i = \beta_{\langle M\rangle_t}^i$, and because a is bounded [Assumption 2.1(iv)], there exists a constant L such that $\langle M\rangle_T \leq LT$ holds. Therefore,

$$\mathbb{E}_y^{\mathbb{Q}}\left[\exp\left(\lambda\sup_{t\in[0,T]}|M_t^i|\right)\right] = \mathbb{E}_y^{\mathbb{Q}}\left[\exp\left(\lambda\sup_{t\in[0,T]}|\beta_{\langle M\rangle_t}^i|\right)\right]$$

$$\leq \mathbb{E}_y^{\mathbb{Q}}\left[\exp\left(\lambda\sup_{t\in[0,LT]}|\beta_t^i|\right)\right]$$

$$< \infty,$$

which completes the proof. □

Proof of Theorem 2.1. By [20, Theorem 3.7.6], it suffices to check that

$$\mathbb{E}[H_T] < \infty, \quad \mathbb{E}[H_T I(zH_T)]$$

for any $z > 0$. By Lemma 3.1,

$$\mathbb{E}[H_T] = \mathbb{E}^{\mathbb{Q}}\left[\exp\left(-\int_0^T r(Y_t)dt\right)\right] < \infty.$$

By Assumption 2.2(ii),

$$H_T I(zH_T) \leq \kappa\left(H_T + z^{-\rho}H_T^{1-\rho}\right)$$

holds, and by Jensen's inequality,

$$\mathbb{E}[H_T I(zH_T)] \leq \kappa\left(\mathbb{E}[H_T] + z^{-\rho}\mathbb{E}\left[H_T^{1-\rho}\right]\right)$$

$$\leq \kappa\left(\mathbb{E}[H_T] + z^{-\rho}\mathbb{E}[H_T]^{1-\rho}\right)$$

$$< \infty$$

holds, which completes the proof. □

Throughout this subsection, we assume Assumptions 2.1–2.3. From Proposition A.5 and Assumption 2.3(ii), the following holds.

Lemma 3.2. $Y = (Y^1, \ldots, Y^m)$ *satisfies the following.*

(i) $Y_s^k \in \bigcap_{p \geq 1} \mathbb{D}_{p,1}, \quad k = 1, \ldots, m, \ s \in [0, T].$

(ii) $D_t Y_s$ satisfies

$$D_t Y_s = a(Y_t) + \int_t^s Db(Y_u) D_t Y_u du + \sum_{l=1}^n \int_t^s Da_{\cdot l}(Y_u) D_t Y_u dW_u^l$$

for $t \in [0, s]$ and $D_t Y_s = 0$ for $t \in (s, T]$.

(iii) For $j = 1, \ldots, n, \ p \in [1, \infty),$

$$\sup_{r \in [0,T]} \mathbb{E} \left[\sup_{s \in [0,T]} \left| D_r^j Y_s^k \right|^p \right] < \infty.$$

(iv) $D_t Y_s = \nabla_y Y_s (\nabla_y Y_t)^{-1} a(Y_t)$ for $t \in [0, s]$, where $\nabla_y Y$ is an $\mathbb{R}^{m \times m}$-valued stochastic process satisfying

$$\nabla_y Y_s = I + \int_0^s Db(Y_u) \nabla_y Y_u du + \sum_{j=1}^n \int_t^s Da_{\cdot j}(Y_u) \nabla_y Y_u dW_u^j$$

for $s \in [0, T]$ and $I \in \mathbb{R}^{m \times m}$ is the identity matrix.

Lemma 3.3. *For $s, t \in [0, T]$,*

(i) $D_s(r(Y_t)) = Dr(Y_t) D_s Y_t,$

(ii) $D_s(\theta(Y_t)) = D\theta(Y_t) D_s Y_t,$

(iii) $D_s \left(\frac{1}{2} |\theta(Y_t)|^2 \right) = \theta^\top(Y_t) D\theta(Y_t) D_s Y_t,$

(iv) $D_s \int_0^T r(Y_t) dt = \int_s^T Dr(Y_t) D_s Y_t dt,$

(v) $D_s \left(\frac{1}{2} \int_0^T |\theta(Y_t)|^2 dt \right) = \int_s^T \theta^\top(Y_t) D\theta(Y_t) D_s Y_t dt,$

(vi) $D_s \int_0^T \theta^\top(Y_t) dW_t = \theta^\top(Y_s) + \left(\int_s^T (D\theta(Y_t) D_s Y_t)^\top dW_t \right)^\top.$

Proof. (i), (ii), and (iii) follow from Proposition A.2 with the remark below it, Lemma 3.2, and Assumption 2.3(i) that $r, \theta, Dr, D\theta$ are of polynomial growth. (iv) and (v) follow from Proposition A.3, and (vi) follows from Proposition A.4. □

The following lemma implies finiteness for p-th moments of $\sup_{t \in [0,T]} |Y_t|$ and $|D_t Y_s|$ under the equivalent martingale measure $\mathbb{Q}$. In this lemma, we use (iii) of Assumption 2.3, which is a monotone condition for SDE (12) that $(Y_s, D_t^1 Y_s, \ldots D_t^n Y_s)_{s \in [t,T]}$ satisfies under $\mathbb{Q}$.

Lemma 3.4. *For any $p \in [1, \infty)$,*

(i) $\mathbb{E}^{\mathbb{Q}}[\sup_{s \in [0,T]} |Y_s|^p] < \infty,$

(ii) $\sup_{s,t \in [0,T]} \mathbb{E}^{\mathbb{Q}}[|D_t Y_s|^p] < \infty.$

Proof. (i) $Y = (Y_t)_{t \in [0,T]}$ satisfies

$$\begin{aligned}
dY_s &= b(Y_s) ds + a(Y_s) dW_s \\
&= \{b(Y_s) - a(Y_s)\theta(Y_s)\} ds + a(Y_s) dW_s^{\mathbb{Q}}.
\end{aligned} \tag{11}$$

Because b and θ have linear growth and a is bounded, SDE (11) satisfies a linear growth condition and (i) follows from Theorem 4.4 in [29, Chapter 2].

(ii) Let $t \in [0, T]$. $\mathbb{R}^m$-valued stochastic processes $(D_t^j Y_s)_{s \in [t,T]}$, $j = 1, \ldots, n$ satisfy

$$D_t^j Y_s = a_{\cdot j}(Y_t) + \int_t^s Db(Y_u)D_t^j Y_u du + \sum_{l=1}^n \int_t^s Da_{\cdot l}(Y_u)D_t^j Y_u dW_u^l$$

$$= a_{\cdot j}(Y_t) + \int_t^s \left(Db(Y_u) - \sum_{l=1}^n Da_{\cdot l}(Y_u)\theta^l(Y_u) \right) D_t^j Y_u du + \sum_{l=1}^n \int_t^s Da_{\cdot l}(Y_u)D_t^j Y_u dW_u^{\mathbb{Q},l}.$$

Set $\mathbb{Y} = (Y_s, D_t^1 Y_s, \ldots D_t^n Y_s)_{s \in [t,T]}^\top$. Then, $\mathbb{R}^{m(n+1)}$-valued stochastic process $\mathbb{Y}$ satisfies

$$d\mathbb{Y}_s = B(\mathbb{Y}_s)ds + \sum_{l=1}^n A^l(\mathbb{Y}_s)dW_s^{\mathbb{Q},l}, \tag{12}$$

with $\mathbb{Y}_t = (Y_t, a_{\cdot 1}(Y_t), \ldots, a_{\cdot n}(Y_t))^\top$, where $B, A^l : \mathbb{R}^{m(n+1)} \to \mathbb{R}^{m(n+1)}$, $l = 1, \ldots, n$ are given by

$$B(y, z_1, \ldots, z_n) = \begin{pmatrix} b(y) - a(y)\theta(y) \\ \left(Db(y) - \sum_{l=1}^n Da_{\cdot l}(y)\theta^l(y) \right) z_1 \\ \vdots \\ \left(Db(y) - \sum_{l=1}^n Da_{\cdot l}(y)\theta^l(y) \right) z_n \end{pmatrix}, \quad A^l(y, z_1, \ldots, z_n) = \begin{pmatrix} a_{\cdot l}(y) \\ Da_{\cdot l}(y)z_1 \\ \vdots \\ Da_{\cdot l}(y)z_n \end{pmatrix}$$

for $y, z_1 \ldots, z_n \in \mathbb{R}^m$. By Assumption 2.3(iii), SDE (12) satisfies a monotone condition, which leads to

$$\mathbb{E}^{\mathbb{Q}}[|D_s Y_t|^p] \le C(1 + \mathbb{E}^{\mathbb{Q}}[|Y_t|^p])$$

from Theorem 4.1 in [29, Chapter 2], where $C > 0$ does not depend on s, t. Using (i) in this lemma, we complete the proof. $\qquad\square$

Lemma 3.5. *Let* $\tilde{L} = (\tilde{L}_s)_{s \in [0,T]}$ *be given by*

$$\tilde{L}_s := \int_s^T Dr(Y_t)D_s Y_t dt + \theta^\top(Y_s) + \left(\int_s^T (D\theta(Y_t)D_s Y_t)^\top dW_t \right)^\top + \int_s^T \theta^\top(Y_t)D\theta(Y_t)D_s Y_t dt, \quad s \in [0,T].$$

Then

$$\sup_{s \in [0,T]} \mathbb{E}^{\mathbb{Q}}[|\tilde{L}_s|^2] < \infty.$$

Proof. Using the assumption that $Dr, \theta, D\theta$ are of polynomial growth and Lemma 3.4, we can easily prove this lemma. $\qquad\square$

Lemma 3.6. (i) $H_T \in \mathbb{D}_{1,1}$ *and*

$$D_s H_T = -H_T \left(\int_s^T Dr(Y_t)D_s Y_t dt + \theta^\top(Y_s) + \left(\int_s^T (D\theta(Y_t)D_s Y_t)^\top dW_t \right)^\top + \int_s^T \theta^\top(Y_t)D\theta(Y_t)D_s Y_t dt \right)$$

$$= -H_T \tilde{L}_s.$$

(ii) Let $z > 0$. $H_T I(zH_T) \in \mathbb{D}_{1,1}$ *and*

$$D_s(H_T I(zH_T)) = -H_T \tilde{L}_s(I(zH_T) + zH_T I'(zH_T)).$$

Proof. (i) From Lemmas 3.1 and 3.5,

$$\mathbb{E}\left[\left(\int_0^T |H_T \tilde{L}_s|^2 ds\right)^{\frac{1}{2}}\right] = \mathbb{E}^{\mathbb{Q}}\left[\beta_T \left(\int_0^T |\tilde{L}_s|^2 ds\right)^{\frac{1}{2}}\right]$$

$$\leq \mathbb{E}^{\mathbb{Q}}\left[\beta_T^2\right]^{\frac{1}{2}} \left(\int_0^T \mathbb{E}^{\mathbb{Q}}\left[|\tilde{L}_s|^2\right] ds\right)^{\frac{1}{2}}$$

$$\leq \mathbb{E}^{\mathbb{Q}}\left[\beta_T^2\right]^{\frac{1}{2}} \left(T \cdot \sup_{s \in [0,T]} \mathbb{E}^{\mathbb{Q}}\left[|\tilde{L}_s|^2\right]\right)^{\frac{1}{2}}$$

$$< \infty$$

holds. Therefore, we can use the chain rule (Proposition A.2), which proves (i).

(ii) By Assumptions 2.2(ii) and 2.3(iv), there exists $\kappa > 0$ such that

$$I(z) + zI'(z) \leq \kappa(1 + z^{-1})$$

for any $z > 0$. As a result,

$$|H_T \tilde{L}_s(I(zH_T) + zH_T I'(zH_T))| \leq \kappa H_T |\tilde{L}_s| + \frac{\kappa}{z}|\tilde{L}_s|,$$

and

$$\mathbb{E}\left[\left(\int_0^T |H_T \tilde{L}_s|^2 ds\right)^{\frac{1}{2}}\right] < \infty, \quad \mathbb{E}\left[\left(\int_0^T |\tilde{L}_s|^2 ds\right)^{\frac{1}{2}}\right] < \infty$$

holds. Therefore, we can use the chain rule again and complete the proof. □

Let a pair (Y, Z) of $\mathbb{R}^m$-valued stochastic process Y and $\mathbb{R}^{m \times m}$-valued stochastic process Z be the solution to the following system of SDEs:

$$dY_s = b(Y_s)ds + a(Y_s)dW_s, \qquad\qquad Y_t = y,$$

$$dZ_s = Db(Y_s)Z_s ds + \sum_{j=1}^n Da_{\cdot j}(Y_s)Z_s dW_s^j, \qquad\qquad Z_t = I.$$

Then (Y, Z) is a Markov process and $(Y^{(t,y)}, Z^{(t,y)})$ denotes the solution to the above system of SDEs when (Y, Z) starts from $(y, I) \in \mathbb{R}^m \times \mathbb{R}^{m \times m}$. Note that $Z^{(t,y)}$ always starts from the identity matrix $I \in \mathbb{R}^{m \times m}$ and $Y_s^{(0,y)} = Y_s^{(t, Y_t^{(0,y)})}$ for $s \in [t, T]$. Because $Z^{(t,y)}$ can be thought of as the derivative of $Y^{(t,y)}$ with respect to the initial value y, we use the notation $\nabla_y Y^{(t,y)} := Z^{(t,y)}$ instead of $Z^{(t,y)}$. Using these notations and Lemma 3.2,

$$D_t Y_s = \nabla_y Y_s^{(t, Y_t)} a(Y_t), \quad s \in [t, T],$$

where $Y_t = Y_t^{(0,y)}$. Furthermore, let $H^{(t,y)} = \left(H_s^{(t,y)}\right)_{s \in [t,T]}$ be

$$H_s^{(t,y)} := \exp\left(-\int_t^s r(Y_u^{(t,y)})du - \int_t^s \theta^\top(Y_u^{(t,y)})dW_u - \frac{1}{2}\int_t^s |\theta(Y_u^{(t,y)})|^2 du\right),$$

and let $\nabla_y H^{(t,y)} = \left(\nabla_y H_s^{(t,y)}\right)_{s \in [t,T]}$, $L^{(t,y)} = \left(L_s^{(t,y)}\right)_{s \in [t,T]}$ be an $\mathbb{R}^m$-valued stochastic process given by

$$L_s^{(t,y)} := \int_t^s (Dr(Y_u^{(t,y)})\nabla_y Y_u^{(t,y)})^\top du + \int_t^s \left(D\theta(Y_u^{(t,y)})\nabla_y Y_u^{(t,y)}\right)^\top dW_u + \int_t^s \left(D\theta(Y_u^{(t,y)})\nabla_y Y_u^{(t,y)}\right)^\top \theta(Y_u^{(t,y)})du,$$

$$\nabla_y H_s^{(t,y)} := -H_s^{(t,y)} L_s^{(t,y)}.$$

16

From these notations,

$$D_t H_T = -H_T \tilde{L}_t$$

$$= -H_T \left(\int_t^T Dr(Y_s) D_t Y_s ds + \theta^\top(Y_t) + \left(\int_t^T (D\theta(Y_s) D_t Y_s)^\top dW_s \right)^\top + \int_s^T \theta^\top(Y_s) D\theta(Y_s) D_t Y_s ds \right)$$

$$= -H_T \theta^\top(Y_t) - H_t^{(0,y)} H_T^{(t,Y_t)} \left(\int_t^T Dr(Y_s) D_t Y_s ds + \left(\int_t^T (D\theta(Y_s) D_t Y_s)^\top dW_s \right)^\top + \int_s^T \theta^\top(Y_s) D\theta(Y_s) D_t Y_s ds \right)$$

$$= -H_T \theta^\top(Y_t) - H_t^{(0,y)} H_T^{(t,Y_t)} \left(\int_t^T Dr(Y_s) \nabla_y Y_s^{(t,Y_t)} ds + \left(\int_t^T (D\theta(Y_s) \nabla_y Y_s^{(t,Y_t)})^\top dW_s \right)^\top \right.$$

$$\left. + \int_s^T \theta^\top(Y_s) D\theta(Y_s) \nabla_y Y_s^{(t,Y_t)} ds \right) a(Y_t)$$

$$= -H_T \theta^\top(Y_t) + H_t^{(0,y)} \left(\nabla_y H_T^{(t,Y_t)} \right)^\top a(Y_t).$$

Proof of Theorem 2.2. By Clark's formula (Proposition A.1), ψ in Theorem 2.1 is given by

$$\psi_t = \mathbb{E}_t \left[D_t \left(H_T I(\hat{\lambda} H_T) \right)^\top \right]$$

$$= -\mathbb{E}_t \left[\left(H_T \theta(Y_t) + H_t a^\top(Y_t) \nabla_y H_T^{(t,Y_t)} \right) \left(I(\hat{\lambda} H_T) + \hat{\lambda} H_T I'(\hat{\lambda} H_T) \right) \right]$$

$$= -H_t \hat{X}_t \theta(Y_t) - \theta(Y_t) \mathbb{E}_t \left[H_T \cdot \hat{\lambda} H_T I'(\hat{\lambda} H_T) \right] + H_t a^\top(Y_t) \mathbb{E}_t \left[\nabla_y H_T^{(t,Y_t)} \left(I(\hat{\lambda} H_T) + \hat{\lambda} H_T I'(\hat{\lambda} H_T) \right) \right].$$

As a result,

$$\hat{\pi}_t = \sigma^\top(Y_t)^{-1} \left(\frac{\psi_t}{H_t} + \hat{X}_t \theta(Y_t) \right)$$

$$= -\sigma^\top(Y_t)^{-1} \theta(Y_t) \frac{1}{H_t} \mathbb{E}_t \left[H_T \cdot \hat{\lambda} H_T I'(\hat{\lambda} H_T) \right] + \sigma^\top(Y_t)^{-1} a^\top(Y_t) \mathbb{E}_t \left[\nabla_y H_T^{(t,Y_t)} \left(I(\hat{\lambda} H_T) + \hat{\lambda} H_T I'(\hat{\lambda} H_T) \right) \right].$$

$$\square$$

Proof of Proposition 2.3. Let $\mathcal{X} : [0,\infty) \times (0,\infty) \times \mathbb{R}^m \to (0,\infty)$ be

$$\mathcal{X}(t,z,y) := \mathbb{E}^y \left[H_t I(z H_t) \right],$$

where $\mathbb{E}^y[\cdot]$ stands for the expectation when Y starts from $y \in \mathbb{R}^m$ at time 0. By the Markov property of Y, $\mathcal{X}(T-t,z,Y_t) = \mathbb{E}_t \left[\frac{H_T}{H_t} I \left(z \frac{H_T}{H_t} \right) \right]$. Furthermore, because H_t is $\mathcal{F}_t$-measurable,

$$\mathcal{X}(T-t, \hat{\lambda} H_t, Y_t) = \frac{1}{H_t} \mathbb{E}_t \left[H_T I \left(\hat{\lambda} H_t \cdot \frac{H_T}{H_t} \right) \right] = \hat{X}_t.$$

Therefore, if we let $F(t,\cdot,y) := (\mathcal{X}(t,\cdot,y))^{-1}$, then

$$F(T-t, \hat{X}_t, Y_t) = \hat{\lambda} H_t \tag{13}$$

holds. Let $\hat{\pi}^M : [0,\infty) \times (0,\infty) \times \mathbb{R}^m \to \mathbb{R}^n$ be

$$\hat{\pi}^M(t,x,y) := -(\sigma^\top(y))^{-1} \theta(y) F(t,x,y) \mathbb{E}^y \left[(H_t)^2 I' (F(t,x,y) H_t) \right].$$

By the Markov property of Y_t,

$$\hat{\pi}^M(T-t,x,Y_t) = -(\sigma^\top(Y_t))^{-1} \theta(Y_t) F(T-t,x,Y_t) \mathbb{E}_t \left[\left(\frac{H_T}{H_t} \right)^2 I' \left(F(T-t,x,Y_t) \frac{H_T}{H_t} \right) \right]$$

17

holds. Moreover, by the identity (13) and the $\mathcal{F}_t$-measurability of $\hat{X}_t$,

$$
\begin{aligned}
\hat{\pi}^M(T-t,\hat{X}_t,Y_t) &= -(\sigma^\top(Y_t))^{-1}\theta(Y_t)F(T-t,\hat{X}_t,Y_t)\mathbb{E}_t\left[\left(\frac{H_T}{H_t}\right)^2 I'\left(F(T-t,\hat{X}_t,Y_t)\frac{H_T}{H_t}\right)\right]\\
&= -(\sigma^\top(Y_t))^{-1}\theta(Y_t)\frac{1}{H_t}\mathbb{E}_t\left[H_T\cdot\hat{\lambda}H_T I'\left(\hat{\lambda}H_T\right)\right]\\
&= \hat{\pi}_t^M.
\end{aligned}
$$

In particular, if we set $T=\tau$, $t=0$, then

$$
\hat{\pi}^M(\tau,x,y) = \hat{\pi}_0^M = -(\sigma^\top(y))^{-1}\theta(y)\mathbb{E}^y\left[H_\tau\cdot\hat{\lambda}H_\tau I'(\hat{\lambda}H_\tau)\right],
$$

where $\hat{\lambda}=\hat{\lambda}(\tau,x,y)$ is defined by an equality $x=\mathbb{E}^y[H_\tau I(\lambda H_\tau)]$. As a result, we have proved that the myopic portfolio $\left(\hat{\pi}_t^M\right)_{t\in[0,T]}$ can be represented by a feedback form. By the same argument, we can also prove that the excess hedging demand $\left(\hat{\pi}_t^H\right)_{t\in[0,T]}$ has a feedback form, which completes the proof. $\qquad\square$

3.2 Proofs for Sect. 2.2

In this subsection, we assume Assumptions 2.1–2.4 and prove the main results in Sect. 2.2. Lemma 3.7 and Proposition 3.8 are useful to estimate the rate of the turnpike theorem for myopic portfolios (Theorem 2.4) and optimal wealth processes (Theorem 2.6). Lemmas 3.9 and 3.10 are used to prove the uniform turnpike theorem for optimal portfolio proportions (Theorem 2.5).

Lemma 3.7.
$$
\mathbb{E}[H_T^{1+\alpha}] \le \mathbb{E}\left[H_T^q\right]^{\frac{\alpha}{q-1}}\mathbb{E}\left[H_T\right]^{1-\frac{\alpha}{q-1}}.
$$

Proof. The case $\alpha=0$ is trivial. When $\alpha\in(q-1,0)$, using Hölder's inequality leads to

$$
\begin{aligned}
\mathbb{E}[H_T^{1+\alpha}] &= \mathbb{E}\left[H_T^{\frac{q\alpha}{q-1}}H_T^{\frac{q-1-\alpha}{q-1}}\right]\\
&\le \mathbb{E}\left[H_T^{\frac{q\alpha}{q-1}\cdot\frac{q-1}{\alpha}}\right]^{\frac{\alpha}{q-1}}\mathbb{E}\left[H_T^{\frac{q-1-\alpha}{q-1}\cdot\frac{q-1}{q-1-\alpha}}\right]^{\frac{q-1-\alpha}{q-1}}\\
&= \mathbb{E}\left[H_T^q\right]^{\frac{\alpha}{q-1}}\mathbb{E}\left[H_T\right]^{1-\frac{\alpha}{q-1}}.
\end{aligned}
$$

$\qquad\square$

Proposition 3.8. *Let* $d(z):=I_1(z)-I_2(z)$. *Then, there exists an* $M=M(x,y)\in(-\infty,x]$, *which is independent of* T, *such that*

$$
\begin{aligned}
\left|\mathbb{E}[H_T d(\hat{\lambda}^{1,T}H_T)]\right| &\le K\left(\mathbb{E}[H_T]+(\hat{\lambda}^{1,T})^\alpha\mathbb{E}[H_T^{1+\alpha}]\right)\\
&\le K\left(\mathbb{E}[H_T]+(x-M)^{\frac{\alpha}{q-1}}\mathbb{E}[H_T]^{1-\frac{\alpha}{q-1}}\right), \quad T>0.
\end{aligned}
$$

In particular,
$$
\left|\mathbb{E}[H_T d(\hat{\lambda}^{1,T}H_T)]\right| = O\left(\mathbb{E}[H_T]^{1-\frac{\alpha}{q-1}}\right), \quad (T\nearrow\infty).
$$

Proof. Note that

$$
\mathbb{E}[H_T d(\hat{\lambda}^{1,T}H_T)] = x - \mathbb{E}[H_T I_2(\hat{\lambda}^{1,T}H_T)] = x-(\hat{\lambda}^{1,T})^{q-1}\mathbb{E}[H_T^q] \le x. \tag{14}
$$

We define $M=M(x,y):=\inf_{T>0}\mathbb{E}[H_T d(\hat{\lambda}^{1,T}H_T)]\in[-\infty,x]$. By Assumption 2.4(ii), (14), and

18

Lemma 3.7,

$$\left| \mathbb{E}[H_T d(\hat{\lambda}^{1,T} H_T)] \right| \leq K \left(\mathbb{E}[H_T] + (\hat{\lambda}^{1,T})^\alpha \mathbb{E}[H_T^{1+\alpha}] \right)$$

$$= K \left\{ \mathbb{E}[H_T] + \left(\frac{x - \mathbb{E}[H_T d(\hat{\lambda}^{1,T} H_T)]}{\mathbb{E}[H_T^q]} \right)^{\frac{\alpha}{q-1}} \mathbb{E}[H_T^{1+\alpha}] \right\}$$

$$\leq K \left\{ \mathbb{E}[H_T] + \left(x - \mathbb{E}[H_T d(\hat{\lambda}^{1,T} H_T)] \right)^{\frac{\alpha}{q-1}} \mathbb{E}[H_T]^{1-\frac{\alpha}{q-1}} \right\}$$

$$\leq K \left\{ \mathbb{E}[H_T] + (x - M)^{\frac{\alpha}{q-1}} \mathbb{E}[H_T]^{1-\frac{\alpha}{q-1}} \right\}$$

holds. Lastly, we prove that $M > -\infty$. Because $\mathbb{E}[H_T] \searrow 0$, there exists some constant C such that $\mathbb{E}[H_T] \leq C \mathbb{E}[H_T]^{1-\frac{\alpha}{q-1}}$. As a result,

$$\left| \mathbb{E}[H_T d(\hat{\lambda}^{1,T} H_T)] \right| \leq K \left\{ C + \left(x - \mathbb{E}[H_T d(\hat{\lambda}^{1,T} H_T)] \right)^{\frac{\alpha}{q-1}} \right\} \mathbb{E}[H_T]^{1-\frac{\alpha}{q-1}} \tag{15}$$

holds. Dividing both sides of the inequality (15) by $|\mathbb{E}[H_T d(\hat{\lambda}^{1,T})]|$, we have

$$1 \leq K \left\{ \frac{C}{\left| \mathbb{E}[H_T d(\hat{\lambda}^{1,T} H_T)] \right|} + \left| \frac{x - \mathbb{E}[H_T d(\hat{\lambda}^{1,T} H_T)]}{\mathbb{E}[H_T d(\hat{\lambda}^{1,T} H_T)]} \right|^{\frac{\alpha}{q-1}} \cdot \frac{1}{\left| \mathbb{E}[H_T d(\hat{\lambda}^{1,T} H_T)] \right|^{1-\frac{\alpha}{q-1}}} \right\} \mathbb{E}[H_T]^{1-\frac{\alpha}{q-1}}. \tag{16}$$

If $M = -\infty$, there exists a sequence $(T_n)_{n \geq 1}$ such that $\mathbb{E}\left[H_{T_n} d(\hat{\lambda}^{1,T_n} H_{T_n}) \right] \searrow -\infty$ and the inequality (16) leads to $1 \leq 0$. Therefore, $M > -\infty$ and the proof is completed. $\qquad\square$

Proof of Theorem 2.4. Let $J_i(z) := z I_i'(z)$, $(i = 1, 2)$.

$$\mathbb{E}\left[H_T \cdot J_1\left(\hat{\lambda}^{1,T} H_T \right) \right] - \mathbb{E}\left[H_T \cdot J_2\left(\hat{\lambda}^{2,T} H_T \right) \right] = \mathbb{E}\left[H_T \left\{ J_1\left(\hat{\lambda}^{1,T} H_T \right) - J_2\left(\hat{\lambda}^{1,T} H_T \right) \right\} \right]$$

$$+ \mathbb{E}\left[H_T \left\{ J_2\left(\hat{\lambda}^{1,T} H_T \right) - J_2\left(\hat{\lambda}^{2,T} H_T \right) \right\} \right]$$

$$=: \text{(I)} + \text{(II)}.$$

Because $J_2(z) = (q-1)z^{q-1}$,

$$\text{(II)} = (q-1) \left\{ \left(\hat{\lambda}^{1,T} \right)^{q-1} - \left(\hat{\lambda}^{2,T} \right)^{q-1} \right\} \mathbb{E}[H_T^q].$$

By the definition of $\hat{\lambda}^{i,T}$,

$$x = \mathbb{E}[H_T I_1(\hat{\lambda}^{1,T} H_T)] = \mathbb{E}[H_T d(\hat{\lambda}^{1,T} H_T)] + (\hat{\lambda}^{1,T})^{q-1} \mathbb{E}[H_T^q],$$

$$x = (\hat{\lambda}^{2,T})^{q-1} \mathbb{E}[H_T^q]$$

hold, and it leads to

$$\text{(II)} = (1 - q) \mathbb{E}[H_T d(\hat{\lambda}^{1,T} H_T)].$$

As a result, by Assumption 2.4(ii) and Proposition 3.8, we obtain

$$|\hat{\pi}^{1,M}(T, x, y) - \hat{\pi}^{2,M}(T, x, y)| \leq K(2 - q) \left| (\sigma^\top(y))^{-1} \theta(y) \right| \left(\mathbb{E}\left[H_T \right] + (\hat{\lambda}^{1,T})^\alpha \mathbb{E}\left[H_T^{\alpha+1} \right] \right)$$

$$\leq K(2 - q) \left| (\sigma^\top(y))^{-1} \theta(y) \right| \left(\mathbb{E}[H_T] + (x - M)^{\frac{\alpha}{q-1}} \mathbb{E}[H_T]^{1-\frac{\alpha}{q-1}} \right).$$

$$\qquad\square$$

Lemma 3.9. *For constants $C > 0$, $\gamma \in [0, 1)$, we define a function f by*

$$f(x, z) := \frac{|z|}{C + (x - z)^\gamma}, \quad x > 0, \ z < -x.$$

Then f satisfies

$$z \in [-x, x) \Rightarrow f(x, z) \geq \frac{|z|}{C + (2x)^\gamma},$$

$$z \in (-\infty, -x) \Rightarrow f(x, z) \geq \frac{|z|^{1-\gamma}}{Cx^{-\gamma} + 2^\gamma}.$$

Proof. When $z \in [-x, x)$,

$$\begin{aligned}
f(x, z) &= \frac{|z|}{C + (x - z)^\gamma} \\
&\geq \frac{|z|}{C + (x - (-x))^\gamma} \\
&= \frac{|z|}{C + (2x)^\gamma}.
\end{aligned}$$

When $z \in (-\infty, -x)$,

$$\begin{aligned}
f(x, z) &= \frac{|z|}{C + (x + |z|)^\gamma} \\
&= \frac{|z|^{1-\gamma}}{C|z|^{-\gamma} + \left(\frac{x}{|z|} + 1\right)^\gamma} \\
&\geq \frac{|z|^{1-\gamma}}{Cx^{-\gamma} + 2^\gamma}.
\end{aligned}$$

$\square$

Lemma 3.10. *For any $\epsilon > 0$,*

$$\sup_{x > \epsilon} \left(\frac{x - M(x, y)}{x}\right) < \infty.$$

Proof. Let $G(T, x, y) := \mathbb{E}[H_T d(\hat{\lambda}^{1,T} H_T)]$. Because

$$\sup_{x > \epsilon} \left(\frac{x - M(x, y)}{x}\right) = 1 - \inf_{x > \epsilon, T > 0} \frac{G(T, x, y)}{x},$$

it suffices to prove

$$\inf_{x > \epsilon, T > 0} \frac{G(T, x, y)}{x} > -\infty.$$

From (15),

$$f(x, G) = \frac{|G|}{C + (x - G)^{\frac{\alpha}{q-1}}} \leq KE[H_T]^{1 - \frac{\alpha}{q-1}},$$

where we define f as in Lemma 3.9 with $\gamma = \dfrac{\alpha}{q-1}$. Lemma 3.9 implies that when $\dfrac{G}{x} < -1$,

$$\frac{|G|^{1-\gamma}}{Cx^{-\gamma} + 2^\gamma} \leq f(x, G) \leq KE[H_T]^{1-\gamma},$$

which means that

$$\begin{aligned}
\frac{G}{x} &\geq -\frac{1}{x} \left\{K \left(Cx^{-\gamma} + 2^\gamma\right)\right\}^{\frac{1}{1-\gamma}} \mathbb{E}[H_T] \\
&= -\left\{K \left(Cx^{-1} + 2^\gamma x^{\gamma-1}\right)\right\}^{\frac{1}{1-\gamma}} \mathbb{E}[H_T] \\
&> -\infty
\end{aligned}$$

20

Therefore, we obtain

$$\inf_{x>\epsilon,T>0} \frac{G(T,x,y)}{x} \geq -\sup_{x>\epsilon}\left\{K\left(Cx^{-1}+2^\gamma x^{\gamma-1}\right)\right\}^{\frac{1}{1-\gamma}} \cdot \sup_{T>0}\mathbb{E}[H_T]$$

$$= -\left\{K\left(C\epsilon^{-1}+2^\gamma\epsilon^{\gamma-1}\right)\right\}^{\frac{1}{1-\gamma}} \cdot \sup_{T>0}\mathbb{E}[H_T]$$

$$> -\infty.$$

□

Proof of Theorem 2.5. From Theorem 2.4 and Lemma 3.10,

$$\sup_{x>\epsilon}\left|\frac{\hat{\pi}^{1,M}(T,x,y)}{x}-\frac{\hat{\pi}^{2,M}(T,x,y)}{x}\right| \leq K(2-q)|\sigma^\top(y)^{-1}\theta(y)|\left(\frac{\mathbb{E}[H_T]}{\epsilon}+\sup_{x>\epsilon}\left(\frac{x-M}{x}\right)^{\frac{\alpha}{q-1}}\left(\frac{\mathbb{E}[H_T]}{\epsilon}\right)^{1-\frac{\alpha}{q-1}}\right)$$

$$= O\left(\mathbb{E}[H_T]^{1-\frac{\alpha}{q-1}}\right).$$

□

Proof of Theorem 2.6. Because $\left(H_t|\hat{X}_t^{1,T}-\hat{X}_t^{2,T}|\right)_t$ is a submartingale,

$$\mathbb{E}[H_t|\hat{X}_t^{1,T}-\hat{X}_t^{2,T}|] \leq \mathbb{E}[H_T|\hat{X}_T^{1,T}-\hat{X}_T^{2,T}|]$$

holds. By the identity $\hat{X}_T^{i,T}=I_i\left(\hat{\lambda}^{i,T}H_T\right)$ and $\mathbb{E}[H_Td(\hat{\lambda}^{1,T}H_T)]=\left\{\left(\hat{\lambda}^{1,T}\right)^{q-1}-\left(\hat{\lambda}^{2,T}\right)^{q-1}\right\}\mathbb{E}[H_T^q]$, it follows that

$$\mathbb{E}\left[H_T|\hat{X}_T^{1,T}-\hat{X}_T^{2,T}|\right] \leq \mathbb{E}\left[H_T\left|I_1\left(\hat{\lambda}^{1,T}H_T\right)-I_2\left(\hat{\lambda}^{1,T}H_T\right)\right|\right]+\mathbb{E}\left[H_T\left|I_2\left(\hat{\lambda}^{1,T}H_T\right)-I_2\left(\hat{\lambda}^{2,T}H_T\right)\right|\right]$$

$$= \mathbb{E}\left[H_T\left|d\left(\hat{\lambda}^{1,T}H_T\right)\right|\right]+\left|\left(\hat{\lambda}^{1,T}\right)^{q-1}-\left(\hat{\lambda}^{2,T}\right)^{q-1}\right|\cdot\mathbb{E}\left[H_T^q\right]$$

$$\leq 2\mathbb{E}\left[H_T\left|d\left(\hat{\lambda}^{1,T}H_T\right)\right|\right].$$

By combining the above inequalities with Proposition 3.8, we get

$$\sup_{t\in[0,T]}\mathbb{E}[H_t|\hat{X}_t^{1,T}-\hat{X}_t^{2,T}|] \leq 2\mathbb{E}\left[H_T\left|d\left(\hat{\lambda}^{1,T}H_T\right)\right|\right]$$

$$\leq 2K\left(\mathbb{E}[H_T]+(\hat{\lambda}^{1,T})^\alpha\mathbb{E}[H_T^{1+\alpha}]\right)$$

$$\leq 2K\left(\mathbb{E}[H_T]+(x-M)^{\frac{\alpha}{q-1}}\mathbb{E}[H_T]^{1-\frac{\alpha}{q-1}}\right)$$

$$= O\left(\mathbb{E}[H_T]^{1-\frac{\alpha}{q-1}}\right), \quad (T\nearrow\infty).$$

□

3.3 Proofs for Sect. 2.3

For excess hedging demands $\hat{\pi}^{i,H}$, the following estimate holds under Assumptions 2.1–2.4.

Proposition 3.11. *We assume Assumptions 2.1–2.4. Let* $M_T^\gamma := \dfrac{H_T^\gamma}{\mathbb{E}[H_T^\gamma]}$, $d\mathbb{Q}_T^\gamma := M_T^\gamma d\mathbb{P}$, $\gamma\in[0,1]$, *and* $M\in(-\infty,x]$ *be as in Proposition 3.8. Then*

$$|\hat{\pi}^{1,H}(T,x,y)-\hat{\pi}^{2,H}(T,x,y)| \leq \left|(\sigma^{-1}(y))^\top a^\top(y)\right|\left\{q\left|\mathbb{E}[H_Td(\hat{\lambda}^{1,T}H_T)]\right|\cdot\mathbb{E}^{\mathbb{Q}_T^q}[|L_T|]\right.$$

$$\left.+2K\left(\mathbb{E}[H_T]\cdot\mathbb{E}^{\mathbb{Q}_T^1}[|L_T|]+(x-M)^{\frac{\alpha}{q-1}}\cdot\mathbb{E}[H_T]^{1-\frac{\alpha}{q-1}}\cdot\mathbb{E}^{\mathbb{Q}_T^{1+\alpha}}[|L_T|]\right)\right\},$$

where

$$L_T := \int_0^T (Dr(Y_u)\nabla_y Y_u)^\top du + \int_0^T (D\theta(Y_u)\nabla_y Y_u)^\top dW_u + \int_0^T (D\theta(Y_u)\nabla_y Y_u)^\top \theta(Y_u)du. \quad (17)$$

21

Proof. Recall that the excess hedging demand is given by

$$\hat{\pi}^{i,H}(T,x,y) = (\sigma^\top(y))^{-1} a^\top(y) \mathbb{E}\left[\nabla_y H_T F_i\left(\hat{\lambda}^{i,T} H_T\right)\right],$$

where $F_i(z) := I_i(z) + z I_i'(z)$ and $\nabla_y H_T = -H_T L_T$.

$$\mathbb{E}\left[\nabla_y H_T \left(F_1(\hat{\lambda}^{1,T} H_T) - F_2(\hat{\lambda}^{2,T} H_T)\right)\right] = \mathbb{E}\left[\nabla_y H_T \left(F_1(\hat{\lambda}^{1,T} H_T) - F_2(\hat{\lambda}^{1,T} H_T)\right)\right]$$
$$+ \mathbb{E}\left[\nabla_y H_T \left(F_2(\hat{\lambda}^{1,T} H_T) - F_2(\hat{\lambda}^{2,T} H_T)\right)\right]$$
$$=: (I) + (II).$$

First, we evaluate (I). Assumption 2.4(ii) implies $|F_1(z) - F_2(z)| \le 2K(1+z^\alpha)$ and Proposition 3.8 implies $(\hat{\lambda}^{1,T})^\alpha \mathbb{E}[H_T^{1+\alpha}] \le (x-M)^{\frac{\alpha}{q-1}} \mathbb{E}[H_T]^{1-\frac{\alpha}{q-1}}$. Using these inequalities and myopic probability measures $d\mathbb{Q}_T^\gamma := \dfrac{H_T}{\mathbb{E}[H_T]} d\mathbb{P}$, it follows that

$$
\begin{aligned}
|(I)| &\le \mathbb{E}\left[H_T |L_T| \left|F_1(\hat{\lambda}^{1,T} H_T) - F_2(\hat{\lambda}^{1,T} H_T)\right|\right] \\
&\le 2K\left(\mathbb{E}[|L_T| H_T] + \left(\hat{\lambda}^{1,T}\right)^\alpha \mathbb{E}[|L_T| H_T^{\alpha+1}]\right) \\
&\le 2K\left(\mathbb{E}[H_T] \mathbb{E}^{\mathbb{Q}_T^1}[|L_T|] + (x-M)^{\frac{\alpha}{q-1}} \mathbb{E}[H_T]^{1-\frac{\alpha}{q-1}} \mathbb{E}^{\mathbb{Q}_T^{1+\alpha}}[|L_T|]\right).
\end{aligned}
\tag{18}
$$

Next, we evaluate (II). Because $F_2(z) = q z^{q-1}$,

$$(II) = q\left\{\left(\hat{\lambda}^{1,T}\right)^{q-1} - \left(\hat{\lambda}^{2,T}\right)^{q-1}\right\} \mathbb{E}\left[\nabla_y H_T H_T^{q-1}\right].$$

Because $\hat{\lambda}^{i,T}$ $(i=1,2)$ are determined by the identity

$$x = \mathbb{E}\left[H_T I_1(\hat{\lambda}^{1,T} H_T)\right] = \left(\hat{\lambda}^{1,T}\right)^{q-1} \mathbb{E}[H_T^q] + \mathbb{E}\left[H_T d(\hat{\lambda}^{1,T} H_T)\right],$$
$$x = \mathbb{E}\left[H_T I_2(\hat{\lambda}^{2,T} H_T)\right] = \left(\hat{\lambda}^{2,T}\right)^{q-1} \mathbb{E}[H_T^q],$$

it follows that

$$\left(\hat{\lambda}^{1,T}\right)^{q-1} - \left(\hat{\lambda}^{2,T}\right)^{q-1} = -\frac{\mathbb{E}\left[H_T d(\hat{\lambda}^{1,T} H_T)\right]}{\mathbb{E}[H_T^q]},$$

which leads to

$$
\begin{aligned}
(II) &= -q\mathbb{E}\left[H_T d(\hat{\lambda}^{1,T} H_T)\right] \frac{\mathbb{E}\left[\nabla_y H_T H_T^{q-1}\right]}{\mathbb{E}[H_T^q]} \\
&= q\mathbb{E}\left[H_T d(\hat{\lambda}^{1,T} H_T)\right] \mathbb{E}^{\mathbb{Q}_T^q}[L_T].
\end{aligned}
\tag{19}
$$

Therefore, the statement follows from (18) and (19). $\qquad\square$

To estimate the convergence rate of the turnpike theorem for excess hedging demands in stochastic factor models (2), the above proposition implies that it suffices to derive $O\left(\mathbb{E}^{\mathbb{Q}_T^\gamma}[|L_T|]\right)$ for $\gamma \in [0,1]$, where $\mathbb{Q}_T^\gamma$ denotes the myopic probability measures used in [13] and L_T is given by (17). Because the martingale density processes of myopic probabilities $\mathbb{Q}_T^\gamma$ can be computed by the optimal portfolios for CRRA investors, which can be represented by solutions to semilinear PDEs, the estimation of $O\left(\mathbb{E}^{\mathbb{Q}_T^\gamma}[|L_T|]\right)$ requires the asymptotic behavior of the solutions to semilinear PDEs. Here, we restrict our models to the quadratic term structure model given by (7) and (8) and use the asymptotic properties of the solutions to the Riccati differential equations. Further research for general stochastic factor models (2) is postponed to future work.

Remark 3.1. When $\gamma = 0$, $\mathbb{Q}_T^\gamma = \mathbb{P}$. When $\gamma = 1$, $d\mathbb{Q}_T^\gamma = \dfrac{H_T}{\mathbb{E}[H_T]}d\mathbb{P}$ is a T-forward measure under which the price process of a zero-coupon bond is chosen as numéraire. When $\gamma \in (0,1)$, as Proposition 3.13 will say, $d\mathbb{Q}_T^\gamma = H_T \cdot \hat{X}_T^T d\mathbb{P}$, where $\hat{X}^T$ is the optimal wealth process for a CRRA investor with a utility function $x \mapsto \dfrac{x^p}{p}$, $p = \dfrac{\gamma}{\gamma - 1}$ and initial unit wealth. Therefore, in the case of $\gamma \in (0,1)$, we choose $\hat{X}^T$ as numéraire under the myopic probability $\mathbb{Q}_T^\gamma$.

3.3.1 Martingale density processes of the myopic probabilities

In Sects. 3.3.1 and 3.3.2, we consider the quadratic term structure model given by (7) and (8) and assume Assumptions 2.4 and 2.5. In this subsection, we first consider the pricing and hedging problem for a T-bond (Proposition 3.12) and utility maximization problems for CRRA investors (Proposition 3.13). Using these results, we compute martingale density processes of myopic probabilities $\mathbb{Q}_T^\gamma$ (Proposition 3.14).

Proposition 3.12. *The price of a T-bond at time t is given by*

$$F(t,y) = \mathbb{E}^{\mathbb{Q}}\left[\exp\left(-\int_t^T r(Y_v)dv\right)\bigg| Y_t = y\right] = \exp\left(-\alpha(t;T) - \beta(t;T)^\top y - \frac{1}{2}y^\top C(t;T)y\right), \quad (20)$$

where $\alpha(\cdot;T) : [0,T] \to \mathbb{R}$, $\beta(\cdot;T) : [0,T] \to \mathbb{R}^m$, and $C(\cdot;T) : [0,T] \to \mathbb{S}_+^m$ solve the following system of ordinary differential equations:

$$\dot{C}(t) - C(t)^\top \Lambda\Lambda^\top C(t) + \tilde{B}^\top C(t) + C(t)\tilde{B} + R_2 = 0,$$
$$C(T) = 0, \tag{21}$$

$$\dot{\beta}(t) + (\tilde{B} - \Lambda\Lambda^\top C(t))\beta(t) + C^\top(t)\tilde{b} + r_1 = 0,$$
$$\beta(T) = 0, \tag{22}$$

$$\dot{\alpha}(t) + \frac{1}{2}\mathrm{Tr}(\Lambda\Lambda^\top C(t)) - \frac{1}{2}\beta^\top(t)\Lambda\Lambda^\top\beta(t) + \tilde{b}\beta(t) + r_0 = 0,$$
$$\alpha(T) = 0, \tag{23}$$

where $\tilde{B} := B - \Lambda A$, $\tilde{b} := b - \Lambda a$. Furthermore, the portfolio proportion process $\hat{\pi}$ that hedges a T-bond and the corresponding wealth process $X^{x,\hat{\pi}}$ with initial wealth $x = F(0,y)$ are given by

$$\hat{\pi}_t = -(\Lambda\Sigma^{-1})^\top \left(\beta(t;T) + C(t;T)Y_t\right),$$
$$X_t^{x,\hat{\pi}} = F(t,Y_t).$$

Proof. This proposition follows from well-known arguments in option pricing theory, which for convenience we include in Appendix B. By Theorem B.1, it suffices to check that the pricing PDE (24) for a T-bond has a unique solution F and $(H_t X_t^{x,\pi})$ is $\mathbb{P}$-martingale, where x and π are given in terms of the solution F. First, the pricing PDE is given by

$$\partial_t F + D_y F^\top (b - \Lambda a + (B - \Lambda A)y) + \frac{1}{2}\mathrm{Tr}\left[\Lambda\Lambda^\top D_{yy}F\right] - r(y)F = 0,$$
$$F(T,y) = 1. \tag{24}$$

If F has the form

$$F(t,y) = \exp\left(-\alpha(t;T) - \beta(t;T)^\top y - \frac{1}{2}y^\top C(t;T)y\right),$$

then $\alpha(\cdot;T) : [0,T] \to \mathbb{R}$, $\beta(\cdot;T) : [0,T] \to \mathbb{R}^m$, and $C(\cdot;T) : [0,T] \to \mathbb{S}_+^m$ are solutions to the ODEs (21), (22), and (23). By Theorem D.1, the Riccati equation (21) has a unique solution, and thus the linear ODEs (22) and (23) have unique solutions. Therefore, Theorem B.1(i) implies that the replicating cost x, the hedging portfolio proportion process $\hat{\pi} = (\hat{\pi}_t)_{t\in[0,T]}$, and the corresponding wealth process $X^{x,\hat{\pi}}$ for a T-bond are given by

$$x = F(0,y), \quad \hat{\pi}_t = -(\Lambda\Sigma^{-1})^\top \left(\beta(t;T) + C(t;T)Y_t\right), \quad X_t^{x,\hat{\pi}} = F(t,Y_t).$$

By Ito's formula,

$$H_t X_t^{x,\hat{\pi}} = x + \int_0^t H_s X_s^{x,\hat{\pi}} \left(\Sigma^\top \hat{\pi}_s - \theta(Y_s) \right)^\top dW_s$$

$$= x + \int_0^t H_s X_s^{x,\hat{\pi}} \left\{ -(\Lambda^\top C(s;T) + A)Y_s - \Lambda^\top \beta(s;T) - a \right\}^\top dW_s,$$

which means that

$$H_t X_t^{x,\hat{\pi}} = x \cdot \mathcal{E} \left(\int_0^\cdot \left\{ -(\Lambda^\top C(s;T) + A)Y_s - \Lambda^\top \beta(s;T) - a \right\}^\top dW_s \right)_t.$$

By the same argument as in [27, Section 6.2], $(H_t X_t^{x,\hat{\pi}})$ is a $\mathbb{P}$-martingale and thus Theorem B.1(ii) implies (20). $\qquad\square$

Next, we recall results on a utility maximization problem for CRRA utility, $x \mapsto \dfrac{x^p}{p}$, $p < 0$.

$$V(t,x,y) := \sup_{\pi \in \mathcal{A}(x)} \mathbb{E} \left[\frac{(X_T^\pi)^p}{p} \,\middle|\, X_t^\pi = x, Y_t = y \right], \quad (t,x,y) \in [0,T] \times (0,\infty) \times \mathbb{R}^m.$$

$$V(T,x,y) := \frac{x^p}{p}, \qquad\qquad\qquad\qquad (x,y) \in (0,\infty) \times \mathbb{R}^m. \tag{25}$$

Proposition 3.13. *The value function V is given by*

$$V(t,x,y) = \frac{x^p}{p} \exp\left(-\frac{1}{2} y^\top P(t;T) y - q(t;T)^\top y - k(t;T) \right), \quad (t,x,y) \in [0,T] \times (0,\infty) \times \mathbb{R}^m,$$

where $P(\cdot;T) : [0,T] \to \mathbb{S}^m, q(\cdot;T) : [0,T] \to \mathbb{R}^m, k(\cdot;T) : [0,T] \to \mathbb{R}$ satisfy the following system of ODEs:

$$\dot{P}(t) - P(t)^\top K_0 P(t) + K_1^\top P(t) + P(t)K_1 + \frac{p}{p-1} A^\top A - p R_2 = 0,$$

$$P(T) = 0, \tag{26}$$

$$\dot{q}(t) + \{K_1 - K_0 P(t)\}^\top q(t) + P(t)b + \frac{p}{p-1} \left(A - \Lambda^\top P(t) \right)^\top a - p r_1 = 0,$$

$$q(T) = 0, \tag{27}$$

$$\dot{k}(t) + \frac{1}{2} \text{Tr}(\Lambda \Lambda^\top P(t)) + \frac{1}{2(p-1)} q(t)^\top \Lambda \Lambda^\top q(t) + \left\{ b - \frac{p}{p-1} \Lambda a \right\}^\top q(t)$$

$$+ \frac{p}{2(p-1)} ||a||^2 - p r_0 = 0, \tag{28}$$

$$k(T) = 0,$$

where $K_0 := \dfrac{1}{1-p} \Lambda \Lambda^\top$, $K_1 := B - \dfrac{p}{p-1} \Lambda A$. Moreover, the optimal portfolio proportion process $\hat{\pi}_t$ and the optimal terminal wealth $X_T^{x,\hat{\pi}}$ for the problem (25) are given by

$$\hat{\pi}_t = \frac{1}{1-p} (\Sigma^\top)^{-1} \theta(Y_t) - \frac{1}{1-p} (\Sigma^\top)^{-1} \Lambda^\top \left(P(t;T) Y_t + q(t;T) \right), \tag{29}$$

$$X_T^{x,\hat{\pi}} = x \frac{H_T^{q-1}}{\mathbb{E}[H_T^q]}. \tag{30}$$

Proof. By Theorem D.1, the Riccati equation (26) and linear ODEs (27) and (28) have unique solutions. Therefore

$$V(t,x,y) := \frac{x^p}{p} \exp\left(-\frac{1}{2} y^\top P(t;T) y - q(t;T)^\top y - k(t;T) \right), \quad (t,x,y) \in [0,T] \times (0,\infty) \times \mathbb{R}^m$$

is the solution to the HJB equation,

$$V_t + xr(y)V_x + (b + By)^\top D_y V + \frac{1}{2}\mathrm{Tr}(\Lambda\Lambda^\top D_{yy}^2 V) - \frac{1}{2V_{xx}}\left|\theta(y)V_x + \Lambda^\top D_y V_x\right|^2 = 0,$$

$$V(T, x, y) = \frac{x^p}{p},$$

and the candidate for the optimal portfolio proportion process $\hat{\pi}$ is given by

$$\hat{\pi}_t = \frac{1}{1-p}(\Sigma^\top)^{-1}\theta(Y_t) - \frac{1}{1-p}(\Sigma^\top)^{-1}\Lambda^\top\left(P(t;T)Y_t + q(t;T)\right).$$

Here, we do not use standard verification arguments. Instead, we directly show that the terminal wealth obtained by the candidate $\hat{\pi}$ matches that of the martingale methods, that is, we establish the identity (30). To do so, Theorem C.1 implies that it suffices to check that $\left(H_t X_t^{x,\hat{\pi}}\right)_{t\in[0,T]}$ is a $\mathbb{P}$-martingale. Because

$$
\begin{aligned}
H_t X_t^{x,\hat{\pi}} &= x \cdot \mathcal{E}\left(\int_0^\cdot \left(\Sigma^\top\hat{\pi}_s - \theta(Y_s)\right)^\top dW_s\right) \\
&= x \cdot \mathcal{E}\left(\int_0^\cdot \left(-\frac{p}{p-1}\theta(Y_s) + \frac{1}{p-1}\Lambda^\top\left(P(s;T)Y_s + q(s;T)\right)\right)^\top dW_s\right) \\
&= x \cdot \mathcal{E}\left(\int_0^\cdot \left\{\left(-\frac{p}{p-1}A + \frac{1}{p-1}\Lambda^\top P(s;T)\right)Y_s - \frac{p}{p-1}a + \frac{1}{p-1}\Lambda^\top q(s;T)\right\}^\top dW_s\right)_t,
\end{aligned}
$$

$$(31)$$

the same argument as in [27, Section 6.2] implies that $\left(H_t X_t^{x,\hat{\pi}}\right)_{t\in[0,T]}$ is a $\mathbb{P}$-martingale, which completes the proof. $\square$

Remark 3.2. If $p < 0$, $P(t;T) : [0,T] \to \mathbb{S}_+^m$ always exists. If $0 < p < 1$, the solution P may blow up at a finite time. See [15] for details.

Proposition 3.14. *Let $\gamma \in [0,1]$. We denote by P^γ, q^γ the solutions to the system of ODEs (26), (27) for $p = \dfrac{\gamma}{\gamma-1}$, $\gamma \in (0,1)$. Let $C^\gamma(\cdot;T) : [0,T] \to \mathbb{S}_+^m$, $\beta^\gamma(\cdot;T) : [0,T] \to \mathbb{R}^m$ be given by*

$$
C^\gamma(t;T) := \begin{cases} 0, & t \in [0,T], \ \gamma = 0, \\ (1-\gamma)P^\gamma(t;T), & t \in [0,T], \ \gamma \in (0,1), \\ C(t;T), & t \in [0,T], \ \gamma = 1, \end{cases}
$$

$$
\beta^\gamma(t;T) := \begin{cases} 0, & t \in [0,T], \ \gamma = 0, \\ (1-\gamma)q^\gamma(t;T), & t \in [0,T], \ \gamma \in (0,1), \\ \beta(t;T), & t \in [0,T], \ \gamma = 1. \end{cases}
$$

Then, the martingale density processes of myopic probability measures $\mathbb{Q}_T^\gamma$ are given by

$$M_T^\gamma = \mathcal{E}\left(\int_0^\cdot \left[-\left\{\gamma A + \Lambda^\top C^\gamma(t;T)\right\}Y_t - \left\{\gamma a + \Lambda^\top\beta^\gamma(t;T)\right\}\right]^\top dW_t\right)_T.$$

Hence, $W^{\mathbb{Q}_T^\gamma} = \left(W_t^{\mathbb{Q}_T^\gamma}\right)_{t\in[0,T]}$, given by

$$W_t^{\mathbb{Q}_T^\gamma} := W_t + \left\{\left(\gamma A + \Lambda^\top C^\gamma(t;T)\right)Y_t + \left(\gamma a + \Lambda^\top\beta^\gamma(t;T)\right)\right\}dt, \tag{32}$$

is an n-dimensional Brownian motion under the myopic probability $\mathbb{Q}_T^\gamma$. Moreover, $Y = (Y_t)_{t\in[0,T]}$ satisfies

$$dY_t = \left\{b - \gamma\Lambda a - \Lambda\Lambda^\top\beta^\gamma(t;T) + \left(B - \gamma\Lambda A - \Lambda\Lambda^\top C^\gamma(t;T)\right)Y_t\right\}dt + \Lambda dW_t^{\mathbb{Q}_T^\gamma}, \quad Y_0 = y. \tag{33}$$

Proof. For $\gamma = 0$, $\mathbb{Q}_T^\gamma = \mathbb{P}$ and the statements follow immediately. For $\gamma = 1$, by Proposition 3.12 and its proof, the state price density process H admits the following stochastic exponential representation:

$$\frac{H_T}{\mathbb{E}[H_T]} = \mathcal{E}\left(\int_0^\cdot \left\{-(\Lambda^\top C(s;T) + A)Y_s - \Lambda^\top \beta(s;T) - a\right\}^\top dW_s\right)_T.$$

From Girsanov's theorem, $W^{\mathbb{Q}_T^\gamma}$ given by (32) is a Brownian motion under $\mathbb{Q}_T^\gamma$ and (33) follows. Therefore, the statements hold for $\gamma = 1$. Next, we consider the case of $\gamma \in (0,1)$. By Proposition 3.13 and (31) for $p = \dfrac{\gamma}{\gamma - 1}$, we have

$$\begin{aligned}
M_T^\gamma &= H_T \cdot \frac{H_T^{\gamma-1}}{\mathbb{E}[H_T^\gamma]} \\
&= \frac{1}{x} \cdot H_T X_T^{x,\hat{\pi}} \\
&= \mathcal{E}\left(\int_0^\cdot \left[\{-\gamma A + (\gamma-1)\Lambda^\top P^\gamma(s;T)\} Y_s - \{\gamma a + (1-\gamma)\Lambda^\top q^\gamma(s;T)\}\right]^\top dW_s\right)_T \\
&= \mathcal{E}\left(\int_0^\cdot \left[-\{\gamma A + \Lambda^\top C^\gamma(t;T)\} Y_t - \{\gamma a + \Lambda^\top \beta^\gamma(t;T)\}\right]^\top dW_t\right)_T.
\end{aligned}$$

By Girsanov's theorem again, the statements follow for $\gamma \in (0,1)$. $\qquad\square$

Remark 3.3. By (26) and (27), for $\gamma \in (0,1)$, $C^\gamma(\cdot;T)$ and $\beta^\gamma(\cdot;T)$ satisfy

$$\dot{C}^\gamma(t) - C^\gamma(t)^\top \Lambda\Lambda^\top C^\gamma(t) + K_1^\top C^\gamma(t) + C^\gamma(t)K_1 + \gamma(1-\gamma)A^\top A + \gamma R_2 = 0,$$
$$C^\gamma(T) = 0, \tag{34}$$

$$\dot{\beta}^\gamma(t) + \{K_1 - \Lambda\Lambda^\top C^\gamma(t)\}^\top \beta^\gamma(t) + C^\gamma(t)\{b - \gamma\Lambda a\} + \gamma(1-\gamma)A^\top a + \gamma r_1 = 0,$$
$$\beta^\gamma(T) = 0, \tag{35}$$

where $K_1 := B - \gamma\Lambda A$. Compared with (21) and (22), $C^\gamma(\cdot;T)$ and $\beta^\gamma(\cdot;T)$ seem to converge to $C^1(\cdot;T) := C(\cdot;T)$ and $\beta^1(\cdot;T) := \beta(\cdot;T)$ in some sense as $\gamma \nearrow 1$. However, we do not prove it because it is not needed for our main results.

3.3.2 Proofs of main results in Sect. 2.3

In this subsection, we first consider the asymptotic behavior of $C^\gamma(t;T)$ and $\beta^\gamma(t;T)$ when $T \nearrow \infty$ (Proposition 3.15), which affects the asymptotic moments of Y_t under myopic probabilities $\mathbb{Q}_\mathbb{T}^\gamma$. Using this proposition, we show

$$\sup_{t \in [0,T]} \mathbb{E}^{\mathbb{Q}_T^\gamma}[|Y_t|^2] = O(1), \qquad \mathbb{E}^{\mathbb{Q}_T^\gamma}[|L_T|] = O(1), \quad (T \nearrow \infty)$$

in Propositions 3.16 and 3.17. Combining these estimates with Proposition 3.11, we can derive the rate of the turnpike theorem for excess hedging demands (Theorem 2.7). We can also prove the uniform turnpike theorem (Theorem 2.8) by the same arguments as in the case of myopic portfolios.

Proposition 3.15. *For $\gamma \in \{q, \alpha+1, 1\}$, there exist $C_\infty^\gamma \in \mathbb{S}_+^m$ and $\beta_\infty^\gamma \in \mathbb{R}^m$ such that*

$$\lim_{T \nearrow \infty} C^\gamma(t;T) = C_\infty^\gamma,$$
$$\lim_{T \nearrow \infty} \beta^\gamma(t;T) = \beta_\infty^\gamma$$

for any t. In addition, $B - \gamma\Lambda A - \Lambda\Lambda^\top C_\infty^\gamma$ is stable.

Proof. Because the statements for $\gamma = 0$ are obvious and those for $\gamma = 1$ are along the lines for the case of $\gamma \in (0,1)$, we consider only the case of $\gamma \in (0,1)$. As Remark 3.3 shows, $C^\gamma(\cdot;T)$ and $\beta^\gamma(\cdot;T)$

satisfy

$$\dot{C}^\gamma(t) - C^\gamma(t)^\top \Lambda\Lambda^\top C^\gamma(t) + K_1^\top C^\gamma(t) + C^\gamma(t)K_1 + \gamma(1-\gamma)A^\top A + \gamma R_2 = 0,$$
$$C^\gamma(T) = 0, \tag{36}$$

$$\dot{\beta}^\gamma(t) + \left\{K_1 - \Lambda\Lambda^\top C^\gamma(t)\right\}^\top \beta^\gamma(t) + C^\gamma(t)\{b - \gamma\Lambda a\} + \gamma(1-\gamma)A^\top a + \gamma r_1 = 0,$$
$$\beta^\gamma(T) = 0, \tag{37}$$

where $K_1 := B - \gamma\Lambda A$. By Theorems D.2 and D.3, we have to check that (K_1, Λ) is stabilizable and (C, K_1) is detectable, where $C \in \mathbb{R}^{n \times m}$ satisfies $C^\top C = \gamma(1-\gamma)A^\top A + \gamma R_2$. The stabilizability of (K_1, Λ) can be seen from the fact that

$$K_1 + \Lambda L = B$$

is stable when setting $L := \gamma A$. To see the detectability of (C, K_1), we consider two cases of (iii) in Assumption 2.5. If $R_2 = 0$, then $C = \sqrt{\gamma(1-\gamma)}A$ and setting $F = \sqrt{\dfrac{\gamma}{1-\gamma}}\Lambda$ implies that

$$FC + K_1 = B,$$

which means that (C, K_1) is detectable. If $\left(\gamma(1-\gamma)A^\top A + \gamma R_2\right)$ is positive definite, then $C := \left(\gamma(1-\gamma)A^\top A + \gamma R_2\right)^{\frac{1}{2}}$ is positive definite and thus (C, K_1) is detectable. As a result, by Theorems D.2 and D.3, $C_\infty^\gamma := \lim_{T \nearrow \infty} C^\gamma(t;T)$ and $\beta_\infty^\gamma := \lim_{T \nearrow \infty} \beta^\gamma(t;T)$ exist for any $t \geq 0$ and satisfy

$$-C_\infty^\gamma \Lambda\Lambda^\top C_\infty^\gamma + K_1^\top C_\infty^\gamma + C_\infty^\gamma K_1 + \gamma(1-\gamma)A^\top A + \gamma R_2 = 0,$$
$$\left\{K_1 - \Lambda\Lambda^\top C_\infty^\gamma\right\}^\top \beta_\infty^\gamma + C_\infty^\gamma\{b - \gamma\Lambda a\} + \gamma(1-\gamma)A^\top a + \gamma r_1 = 0,$$

and

$$K_1 - \Lambda\Lambda^\top C_\infty^\gamma = B - \gamma\Lambda A - \Lambda\Lambda^\top C_\infty^\gamma$$

is stable. $\qquad\square$

Proposition 3.16. *For* $\gamma \in \{q, \alpha+1, 1\}$,

$$\sup_{t \in [0,T]} \mathbb{E}^{\mathbb{Q}_T^\gamma}[|Y_t|^2] = O(1), \quad (T \nearrow \infty).$$

Proof. We define $\tilde{\beta}^\gamma, \tilde{C}^\gamma$ by

$$\tilde{\beta}^\gamma(t;T) := b - \gamma\Lambda a - \Lambda\Lambda^\top \beta^\gamma(t;T),$$
$$\tilde{C}^\gamma(t;T) := B - \gamma\Lambda A - \Lambda\Lambda^\top C^\gamma(t;T),$$

and the SDE (33) becomes

$$dY_t = \left(\tilde{\beta}^\gamma(t;T) + \tilde{C}^\gamma(t;T)Y_t\right)dt + \Lambda dW_t^{\mathbb{Q}_T^\gamma}, \quad Y_0 = y.$$

Because $\tilde{C}_\infty^\gamma := \lim_{T \nearrow \infty} \tilde{C}^\gamma(t;T) = B - \gamma\Lambda A - \Lambda\Lambda^\top C_\infty^\gamma$ is stable, there exists $U \in \mathbb{S}_{++}^m$ such that

$$(\tilde{C}_\infty^\gamma)^\top U + U\tilde{C}_\infty^\gamma < 0.$$

Therefore, there exist $T_1, \epsilon > 0$ such that

$$T - t \geq T_1 \Rightarrow \tilde{C}^\gamma(t;T)^\top U + U\tilde{C}^\gamma(t;T) < -\epsilon I_m,$$

where I_m is an $m \times m$ identity matrix. For $0 \leq s \leq t \leq T$,

$$\mathbb{E}^{\mathbb{Q}_T^\gamma}[\langle UY_t, Y_t\rangle] = \mathbb{E}^{\mathbb{Q}_T^\gamma}[\langle UY_s, Y_s\rangle] + \int_s^t \mathbb{E}^{\mathbb{Q}_T^\gamma}\left[\left\langle\left(\tilde{C}^\gamma(u;T)^\top U + U\tilde{C}^\gamma(u;T)\right)Y_u, Y_u\right\rangle + 2\left\langle U\tilde{\beta}^\gamma(u;T), Y_u\right\rangle\right]du$$
$$+ \text{Tr}(U^\top \Lambda\Lambda^\top)(t-s).$$

Because U is positive definite, the maximum and minimum of the eigenvalues, $\lambda_{\min}, \lambda_{\max} > 0$, satisfy

$$\lambda_{\min}|y|^2 \leq \langle Uy, y \rangle \leq \lambda_{\max}|y|^2 \tag{38}$$

for all $y \in \mathbb{R}^m$. Moreover, because the function $\tilde{\beta}$ is a bounded function of (t, T), there exist positive constants C such that

$$\frac{2}{\epsilon}|U\tilde{\beta}(t; T)|^2 + \text{Tr}(U^\top \Lambda\Lambda^\top) \leq C \tag{39}$$

for any t, T with $0 \leq t \leq T$. From the above inequalities (38) and (39), it follows that for $t \in [0, T - T_1]$,

$$\frac{d}{dt}\mathbb{E}^{\mathbb{Q}_T^\gamma}[\langle UY_t, Y_t \rangle] = \mathbb{E}^{\mathbb{Q}_T^\gamma}\left[\left\langle \left(\tilde{C}^\gamma(t; T)^\top U + U\tilde{C}^\gamma(t; T)\right)Y_t, Y_t\right\rangle + 2\left\langle U\tilde{\beta}^\gamma(t; T), Y_t\right\rangle\right] + \text{Tr}(U^\top \Lambda\Lambda^\top)$$

$$\leq -\epsilon\mathbb{E}^{\mathbb{Q}_T^\gamma}\left[|Y_t|^2\right] + 2\mathbb{E}^{\mathbb{Q}_T^\gamma}\left[\langle U\tilde{\beta}(t; T), Y_t\rangle\right] + \text{Tr}(U^\top \Lambda\Lambda^\top)$$

$$\leq -\frac{\epsilon}{2}\mathbb{E}^{\mathbb{Q}_T^\gamma}\left[|Y_t|^2\right] + \frac{2}{\epsilon}|U\tilde{\beta}(t; T)|^2 + \text{Tr}(U^\top \Lambda\Lambda^\top)$$

$$\leq -\frac{\epsilon}{2\lambda_{\max}}\mathbb{E}^{\mathbb{Q}_T^\gamma}[\langle UY_t, Y_t \rangle] + C,$$

where the second inequality follows from ϵ-Hölder's inequality. By Gronwall's inequality, there exists a constant $C = C(y)$ which depends only on y such that

$$\mathbb{E}^{\mathbb{Q}_T^\gamma}[\langle UY_t, Y_t \rangle] \leq C(y), \quad t \in [0, T - T_1].$$

The above inequalities and the right-hand side of (38) lead to

$$\mathbb{E}^{\mathbb{Q}_T^\gamma}\left[|Y_t|^2\right] \leq \frac{C(y)}{\lambda_{\min}}, \quad t \in [0, T - T_1].$$

Furthermore, because the length of $[T - T_1, T]$ is T_1, there exists a constant $C = C(y, T_1)$ which depends only on y and T_1 such that

$$\mathbb{E}^{\mathbb{Q}_T^\gamma}\left[|Y_t|^2\right] \leq C(y, T_1), \quad t \in [T - T_1, T].$$

As a result, the proposition follows. $\qquad\square$

Proposition 3.17. *For $\gamma \in \{q, \alpha + 1, 1\}$,*

$$\mathbb{E}^{\mathbb{Q}_T^\gamma}[|L_T|] = O(1), \quad (T \nearrow \infty).$$

Proof. Because $Dr(y) = r_1 + R_2 y$, $D\theta(y) = A$, $\nabla_y Y_t = e^{Bt}$, (17) becomes

$$L_T = \int_0^T ((r_1 + R_2 Y_t)^\top e^{Bt})^\top dt + \int_0^T (Ae^{Bt})^\top dW_t^{\mathbb{Q}_T^\gamma}$$

$$- \int_0^T (Ae^{Bt})^\top \left[\{(\gamma - 1)A + \Lambda^\top C^\gamma(t; T)\}Y_u + (\gamma - 1)a + \Lambda^\top \beta^\gamma(t; T)\right] dt.$$

Because B is stable, there exist $M, \omega > 0$ such that

$$|e^{Bt}| \leq Me^{-\omega t}, \ (t \geq 0).$$

Furthermore, C^γ, β^γ are bounded as functions of (t, T). Therefore, this proposition follows from Proposition 3.16. $\qquad\square$

Proof of Theorem 2.7. The result follows from Propositions 3.11 and 3.17. $\qquad\square$

Proof of Theorem 2.8. We can prove the theorem in the same way as in the proof of Theorem 2.5. $\quad\square$

3.4   Proofs for Sect. 2.4

Lemma 3.18. *Let U_1, U_2 be utility functions.*

(i) If there exist $f, \tilde{f} : (0, \infty) \to [0, \infty)$ such that

$$U_1'(x + f(x)) \le U_2'(x) \le U_1'(x + \tilde{f}(x)), \quad x > 0 \tag{40}$$

holds, then

$$|I_1(z) - I_2(z)| \le \max\left\{ f(I_2(z)), \tilde{f}(I_2(z)) \right\}, \quad z > 0$$

holds.

(ii) In addition to the assumption in (i), if there exists $g : (0, \infty) \to \mathbb{R}$ such that

$$|ART_1(x) - ART_2(x)| \le g(x), \quad x > 0 \tag{41}$$

holds and U_2 is the HARA utility, that is, $ART_2(x) = ax + b \ (x > 0)$ for some constants $a, b \in \mathbb{R}$, then

$$|zI_1'(z) - zI_2'(z)| \le g(I_1(z)) + |a| \max\left\{ f(I_2(z)), \tilde{f}(I_2(z)) \right\}, \quad z > 0$$

holds.

Proof.   (i) Substituting $x = I_2(z), z > 0$ in (40),

$$U_1'(I_2(z) + f(I_2(z))) \le z \le U_1'(I_2(z) + \tilde{f}(I_2(z)))$$

holds. Because I_1 is decreasing, we get

$$I_2(z) + \tilde{f}(I_2(z)) \le I_1(z) \le I_2(z) + f(I_2(z)),$$

which leads to

$$|I_1(z) - I_2(z)| \le \max\left\{ f(I_2(z)), \tilde{f}(I_2(z)) \right\}.$$

(ii) $zI_i'(z)$ can be represented by *the Arrow–Pratt measure of absolute risk tolerance* $ART_i(x) := -\dfrac{U_i'(x)}{U_i''(x)}$ as follows:

$$zI_i'(z) = \frac{z}{U_i''(I_i(z))} = \frac{U_i'(I_i(z))}{U_i''(I_i(z))} = -ART_i(I_i(z)).$$

Using (41) and the assumption that U_2 is the HARA utility, we obtain

$$\begin{aligned}
|zI_1'(z) - zI_2'(z)| &= |ART_1(I_1(z)) - ART_2(I_2(z))| \\
&\le |ART_1(I_1(z)) - ART_2(I_1(z))| + |ART_2(I_1(z)) - ART_2(I_2(z))| \\
&\le g(I_1(z)) + |a||I_1(z) - I_2(z)| \\
&\le g(I_1(z)) + |a| \max\left\{ f(I_2(z)), \tilde{f}(I_2(z)) \right\},
\end{aligned}$$

which completes the proof. $\square$

Let

$$U_1(x) := \sum_{i=1}^{n} \beta_i \frac{(\alpha_i x)^{p_i}}{p_i} = \sum_{i=1}^{n} w_i \frac{x^{p_i}}{p_i},$$

$$U_2(x) := w_n \frac{x^{p_n}}{p_n},$$

where $w_i := \beta_i \alpha_i^{p_i} > 0$.

Proof of Proposition 2.11. First, we prove the inequality (9). Because $U_1'(x) = \sum_{i=1}^{n} w_i x^{p_i - 1}$, $U_2'(x) = w_n x^{p_n - 1}$,

$$U_2'(x) \leq U_1'(x), \quad x > 0$$

holds. To look for $\beta \in \mathbb{R}$ that satisfies

$$U_1'(x + x^\beta) \leq U_2'(x)$$

for large enough $x > 0$, we define h by

$$h(x) := \frac{U_2'(x) - U_1'(x + x^\beta)}{w_n(x + x^\beta)^{p_n - 1}}$$

$$= \left(\frac{x}{x + x^\beta} \right)^{p_n - 1} - \sum_{i=1}^{n-1} \frac{w_i}{w_n}(x + x^\beta)^{p_i - p_n} - 1$$

for all $x > 0$. Then h', a derivative of h, is given by

$$h'(x) = \frac{(p_n - 1)(1 - \beta)x^{p_n - 2 + \beta}}{(x + x^\beta)^{p_n}} - \sum_{i=1}^{n-1} \frac{w_i}{w_n}(p_i - p_n)\frac{(x + x^\beta)^{p_i - 1}}{(x + x^\beta)^{p_n}}(1 + \beta x^{\beta - 1})$$

$$= -\frac{(p_n - p_{n-1})x^{p_n + \beta - 2}}{(x + x^\beta)^{p_n}} \left\{ \frac{1 - p_n}{p_n - p_{n-1}}(1 - \beta) - \sum_{i=1}^{n-1} \frac{w_i}{w_n} \cdot \frac{p_n - p_i}{p_n - p_{n-1}} \cdot (1 + x^{\beta - 1})^{p_i - 1} \left(\beta x^{p_i - p_n} + x^{1 + p_i - p_n - \beta} \right) \right\}.$$

If $\beta \in (1 + p_{n-1} - p_n, 1)$, then

$$1 + p_i - p_n - \beta < 1 + p_{n-1} - p_n - \beta < 0, \quad (i = 1, \ldots, n - 1),$$

which means that

$$\sum_{i=1}^{n-1} \frac{w_i}{w_n} \cdot \frac{p_n - p_i}{p_n - p_{n-1}} \cdot (1 + x^{\beta - 1})^{p_i - 1} \left(\beta x^{p_i - p_n} + x^{1 + p_i - p_n - \beta} \right) \to \sum_{i=1}^{n-1} \frac{w_i}{w_n} \cdot \frac{p_n - p_i}{p_n - p_{n-1}} \cdot (1 + 0)^{p_i - 1} (\beta \cdot 0 + 0) = 0$$

as $x \nearrow \infty$. Therefore, $h'(x) < 0$ holds for large enough $x > 0$. Combining this with the fact that

$$h(x) = \left(\frac{1}{1 + x^{\beta - 1}} \right)^{p_n - 1} - \sum_{i=1}^{n-1} \frac{w_i}{w_n} x^{p_i - p_n}(1 + x^{\beta - 1})^{p_i - p_n} - 1$$

$$\to 1 - \sum_{i=1}^{n-1} \frac{w_i}{w_n} \cdot 0 \cdot (1 + 0)^{p_i - p_n} - 1 \quad (x \nearrow \infty)$$

$$= 0$$

leads to $h(x) > 0$ for large enough $x > 0$. Using Lemma 3.18(i) with $f(x) = x^\beta, \tilde{f}(x) = 0$, and $\beta \in (1 + p_{n-1} - p_n, 1)$ leads to

$$|I_1(z) - I_2(z)| \leq \left(\frac{z}{w_n} \right)^{\beta(q_n - 1)}$$

for small enough $z > 0$, which means that the inequality (9) holds for any nonnegative $\beta \in (1 + p_{n-1} - p_n, 1)$ and a constant $K > 0$. Next, we prove the inequality (10). Because

$$|ART_1(x) - ART_2(x)| = \left| \frac{\sum_{i=1}^{n} w_i x^{p_i - 1}}{\sum_{i=1}^{n} w_i(1 - p_i)x^{p_i - 2}} - \frac{x}{1 - p_n} \right|$$

$$= \frac{\sum_{i=1}^{n-1}(p_n - p_i)w_i x^{p_i - 1}}{(1 - p_n)\left(\sum_{i=1}^{n} w_i(1 - p_i)x^{p_i - 2} \right)}$$

$$\leq \frac{\sum_{i=1}^{n-1}(p_n - p_i)w_i x^{p_i - 1}}{(1 - p_n)^2 w_n x^{p_n - 2}}$$

$$= \sum_{i=1}^{n-1} \frac{p_n - p_i}{(1 - p_n)^2} \cdot \frac{w_i}{w_n} \cdot x^{p_i - p_n + 1}$$

$$\leq C x^{p_{n-1} - p_n + 1}$$

holds for large enough $x > 0$ and a constant $C > 0$, Lemma 3.18(ii) implies that for small enough $z > 0$,

$$
\begin{aligned}
|zI_1'(z) - zI_2'(z)| &\leq CI_1(z)^{p_{n-1}-p_n+1} + \frac{1}{1-p_n}\left(\frac{z}{w_n}\right)^{\beta(q_n-1)} \\
&\leq C\left(\left(\frac{z}{w_n}\right)^{q_n-1} + \left(\frac{z}{w_n}\right)^{\beta(q_n-1)}\right)^{(p_{n-1}-p_n+1)^+} + \frac{1}{1-p_n}\left(\frac{z}{w_n}\right)^{\beta(q_n-1)} \\
&\leq \tilde{C}\left(z^{(q_n-1)(p_{n-1}-p_n+1)^+} + z^{\beta(q_n-1)(p_{n-1}-p_n+1)^+} + z^{\beta(q_n-1)}\right) \\
&\leq 3\tilde{C}z^{\beta(q_n-1)},
\end{aligned}
$$

where $\tilde{C}$ is a constant, and the last inequality follows from

$$
\beta(q_n-1) \leq (q_n-1)(p_{n-1}-p_n+1)^+ \leq \beta(q_n-1)(p_{n-1}-p_n+1)^+ \leq 0.
$$

We have completed the proof. $\qquad\square$

A   Appendix: Malliavin calculus

We recall some results on Malliavin calculus which allow us to derive an explicit stochastic flow representation for the optimal portfolio process $\hat{\pi}$. For details, see [34, 36, 37, 43].

Consider a complete probability space $(\Omega, \mathcal{F}, P)$ and a standard n-dimensional Brownian motion $W = (W^1, \ldots, W^n)^\top$ defined on $(\Omega, \mathcal{F}, P)$. We denote by $(\mathcal{F}_t)_{t\geq 0}$ the P-augmentation of the natural filtration generated by $W = (W_t)_{t\geq 0}$.

We introduce the Malliavin derivative operator as in [36]. Fix $T > 0$. Let $C_b^\infty(\mathbb{R}^m)$ be the space of infinitely differentiable functions on $\mathbb{R}^m$ which, together with all partial derivatives, are bounded. By $\mathscr{S}$ we denote the class of *smooth random variables*, namely, random variables of the form

$$
F = f(W_{t_1}, \ldots, W_{t_m}),
$$

where $(t_1, \ldots, t_m) \in [0, T]^m$ and the function $f = f(x^{11}, \ldots, x^{n1}, \ldots, x^{1m}, \ldots, x^{nm})$ belongs to $C_b^\infty(\mathbb{R}^{nm})$. For each $F \in \mathscr{S}$, the *Malliavin derivative* of F is the $L^2([0,T];\mathbb{R})^n$-valued random variable $DF = (D^1F, \ldots D^nF)$ with components

$$
D^iF(\cdot) := \sum_{j=1}^m \frac{\partial f}{\partial x^{ij}}(W_{t_1}, \ldots, W_{t_m})1_{[0,t_j]}(\cdot), \quad (i = 1, \ldots, n).
$$

Fix $p \in [1, \infty)$. Because we can view the operator D as an operator from $L^p(\Omega; \mathbb{R})$ to $L^p\left(\Omega; (L^2[0,T];\mathbb{R})^n\right)$ and D is closable, we denote the closure of D again by D and the domain of D in $L^p(\Omega;\mathbb{R})$ by $\mathbb{D}_{p,1}$. Thus, $\mathbb{D}_{p,1}$ is the closure of $\mathscr{S}$ with respect to the norm

$$
\begin{aligned}
||F||_{p,1} &:= ||F||_{L^p(\Omega;\mathbb{R})} + ||DF||_{L^p(\Omega;(L^2[0,T];\mathbb{R})^n)} \\
&= \mathbb{E}[|F|^p]^{\frac{1}{p}} + \mathbb{E}\left[\left(\int_0^T |DF(t)|^2 dt\right)^{\frac{p}{2}}\right]^{\frac{1}{p}}.
\end{aligned}
$$

Here, $|\cdot|$ denotes the Euclidean norm on $\mathbb{R}^n$. $\mathbb{D}_{p,1}$ is a Banach space with respect to the norm $||\cdot||_{p,1}$. Given $F \in \mathbb{D}_{p,1}$, DF is an $(L^2[0,T])^n$-valued random variable. To each DF, we can find a measurable process $[0,T] \times \Omega \ni (t,\omega) \mapsto D_tF(\omega) \in \mathbb{R}^n$ such that for almost all $\omega \in \Omega$, $D_tF(\omega) = DF(\omega)(t)$ holds for almost everywhere $t \in [0,T]$. Therefore, we identify $(L^2[0,T])^n$-valued random variable DF with $\mathbb{R}^n$-valued measurable process $(t,\omega) \mapsto D_tF(\omega)$ without further comment.

Remark A.1. Note that for real-valued random variable $F \in \mathbb{D}_{p,1}$, we define $DF = (D^1F, \ldots, D^nF)$ as a row vector; that is, DF is an $\mathbb{R}^{1\times n}$-valued stochastic process. For $\mathbb{R}^m$-valued random variable $F = (F_1, \ldots, F_m)^\top \in \mathbb{D}_{p,1}^m$ we define $DF = (D^jF_i)_{\substack{1\leq i\leq m \\ 1\leq j\leq n}}$, which is an $\mathbb{R}^{m\times n}$-valued stochastic process.

We collect well-known results on Malliavin calculus, that is, Clark's formula (Proposition A.1), chain rule (Proposition A.2), and Malliavin derivatives of Lebesgue integrals, stochastic integrals, and solutions to SDEs (Propositions A.3–A.5). Firstly, we quote Clark's formula for random variables in $\mathbb{D}_{1,1}$, which comes from [23].

Proposition A.1. *For every $F \in \mathbb{D}_{1,1}$ we have*

$$\mathbb{E}[F|\mathcal{F}_t] = \mathbb{E}[F] + \int_0^t \mathbb{E}\left[D_s F | \mathcal{F}_s\right] dW_s, \quad t \in [0, T].$$

The following proposition is a straightforward multidimensional version of Lemma A.1 in [36].

Proposition A.2. *Let $F = (F_1, \ldots F_m)^\top \in \mathbb{D}_{1,1}^m$. Let $\phi = (\phi^1, \ldots, \phi^k)^\top \in C^1(\mathbb{R}^m; \mathbb{R}^k)$. Assume that*

$$\mathbb{E}[|\phi^l(F)|] + \mathbb{E}\left[\left(\int_0^T \left|\sum_{i=1}^m \frac{\partial \phi^l}{\partial x_i}(F) D_t F_i\right|^2 dt\right)^{\frac{1}{2}}\right] < \infty \tag{42}$$

for all $l = 1, \ldots k$. Then $\phi(F) \in \mathbb{D}_{1,1}^k$ and

$$D(\phi(F)) = D\phi(F) DF,$$

where $D\phi = \left(\dfrac{\partial \phi^i}{\partial x_j}\right)_{\substack{1 \le i \le k \\ 1 \le j \le m}} : \mathbb{R}^m \to \mathbb{R}^{k \times m}$ is the Jacobi matrix of ϕ.

Remark A.2. Hölder's inequality implies that the condition (42) holds if $F \in \mathbb{D}_{p,1}^m$, $\phi^l(F) \in L^1$, $\dfrac{\partial \phi^l}{\partial x_i}(F) \in L^q$, $l = 1, \ldots k$, $i = 1, \ldots m$ for some $p \in [1, \infty)$, $q \in (1, \infty]$ such that $\dfrac{1}{p} + \dfrac{1}{q} = 1$. In particular, the condition (42) holds if $F \in \bigcap_{p \ge 1} \mathbb{D}_{p,1}^m$ and $\phi, D\phi$ are of polynomial growth.

Proposition A.3 is Lemma 5.1 in [24], Proposition A.4 is Proposition 2.3 in [36], and Proposition A.5 is Proposition 8.2 in [43].

Proposition A.3. *Let $u = (u_s)_{s \in [0,T]}$ be a real-valued, continuous, progressively measurable process such that*

(i) $u_s \in \mathbb{D}_{1,1}$ *for every $s \in [0, T]$,*

(ii) $\sup_{s \in [0,T]} \mathbb{E}\left[|u_s|^q\right] < \infty$ *for some $q > 1$, and* $\sup_{s \in [0,T]} \mathbb{E}\left[\int_0^T |D_t^j u_s|^4 dt\right] < \infty$ *for $j = 1, \ldots, n$,*

(iii) $s \mapsto D_t u(s, \omega)$ *is left (or right) continuous for almost every $(t, \omega) \in [0, T] \times \Omega$.*

Then $\int_0^T u_s ds \in \mathbb{D}_{1,1}$ and $D_t \int_0^T u_s ds = \int_t^T D_t u_s ds$.

Proposition A.4. *Let $u = (u^1, \ldots, u^n)^\top$ be an $\mathbb{R}^n$-valued progressively measurable process such that*

(i) $u_s \in \mathbb{D}_{1,1}^n$ *for every $s \in [0, T]$,*

(ii) $[0, T] \times \Omega \ni (x, \omega) \mapsto Du(s, \omega) \in (L^2[0, T])^{n^2}$ *admits a progressively measurable version,*

(iii)

$$|||u|||_{1,1} := \mathbb{E}\left[\left(\int_0^T |u_s|^2 ds\right)^{\frac{1}{2}}\right] + \mathbb{E}\left[\left\{\int_0^T \int_0^T |D_t u_s|^2 dt ds\right\}^{\frac{1}{2}}\right] < \infty,$$

where $|\cdot|$ denotes Euclidean norm on $\mathbb{R}^{n \times n}$.

Then $\int_0^T u_s^\top dW_s \in \mathbb{D}_{1,1}$ *and*

$$D_t \int_0^T u_s^\top dW_s = u_t^\top + \left(\int_0^T (D_t u_s)^\top dW_s \right)^\top.$$

Proposition A.5. *For $d \in \mathbb{N}$, we consider the d-dimensional SDE*

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW_t, \quad X_0 = x \in \mathbb{R}^d, \tag{43}$$

where $\mu = (\mu_1, \ldots \mu_d)^\top : \mathbb{R}^d \to \mathbb{R}^d$, $\sigma = (\sigma_{i,j})_{\substack{1 \le i \le d \\ 1 \le j \le n}} : \mathbb{R}^d \to \mathbb{R}^{d \times n}$ are continuously differentiable and satisfy

$$\sup_{x \in \mathbb{R}^d} \left(\left| \frac{\partial \mu_i}{\partial x_k}(x) \right| + \left| \frac{\partial \sigma_{i,j}}{\partial x_k}(x) \right| \right) < \infty$$

for $i, k = 1, \ldots, d$, $j = 1, \ldots, n$. Then (43) has a unique strong solution $X = (X^1, \ldots, X^d)^\top$ which satisfies the following:

(i) $X_s^k \in \bigcap_{p \ge 1} \mathbb{D}_{p,1}$, $k = 1, \ldots d$, $s \in [0, T]$;

(ii) $D_t X_s$ satisfies

$$D_t X_s = \sigma(X_t) + \int_t^s D\mu(X_u)D_t X_u du + \sum_{j=1}^n \int_t^s D\sigma_{\cdot j}(X_u)D_t X_u dW_u^j$$

for $t \in [0, s]$ and $D_t X_s = 0$ for $t \in (s, T]$, where $D\mu = \left(\frac{\partial \mu_i}{\partial x_j} \right)_{\substack{1 \le i \le d \\ 1 \le j \le d}}$, $D\sigma_{\cdot j} = \left(\frac{\sigma_{i,j}}{\partial x_l} \right)_{\substack{1 \le i \le d \\ 1 \le l \le d}};$

(iii) for $j = 1, \ldots, n$, $p \in [1, \infty)$,

$$\sup_{r \in [0,T]} \mathbb{E} \left[\sup_{s \in [0,T]} \left| D_r^j X_s^k \right|^p \right] < \infty;$$

(iv) $D_t X_s = \nabla_x X_s (\nabla_x X_t)^{-1} \sigma(X_t)$ for $t \in [0, s]$, where $\nabla_x X$ is an $\mathbb{R}^{d \times d}$-valued stochastic process satisfying

$$\nabla_x X_s = I + \int_0^s D\mu(X_u)\nabla_x X_u du + \sum_{j=1}^n \int_t^s D\sigma_{\cdot j}(X_u)\nabla_x X_u dW_u^j$$

for $s \in [0, T]$ and $I \in \mathbb{R}^{d \times d}$ is the identity matrix.

B   Appendix: Option pricing theory with stochastic factor models in complete markets

In this appendix, we recall well-known results on pricing and hedging problems in a complete market with a stochastic factor process (2). Under the equivalent martingale measure $\mathbb{Q}$, the dynamics of risky assets S and the stochastic factor process Y are denoted by

$$dS_t = \text{diag}(S_t) \left(\mathbf{1} r(Y_t)dt + \sigma(Y_t)dW_t^\mathbb{Q} \right), \qquad S_0 = s_0 \in \mathbb{R}_{++}^n,$$

$$dY_t = \tilde{b}(Y_t)dt + a(Y_t)dW_t^\mathbb{Q}, \qquad Y_0 = y \in \mathbb{R}^m,$$

where $\tilde{b}(y) := b(y) - a(y)\theta(y)$. Let $\mathcal{L}$ be a generator of (S, Y) under $\mathbb{Q}$; that is, for $f : [0, T] \times \mathbb{R}_{++}^n \times \mathbb{R}^m \to \mathbb{R}$, $\mathcal{L}f : [0, T] \times \mathbb{R}_{++}^n \times \mathbb{R}^m \to \mathbb{R}$ is defined by

$$\mathcal{L}f := D_s f^\top r(y)s + D_y f^\top \tilde{b}(y) + \frac{1}{2} \text{Tr} \left[\left\{ \Sigma\Sigma^\top \right\} (s, y) D^2 f \right], \quad (t, s, y) \in [0, T] \times \mathbb{R}_{++}^n \times \mathbb{R}^m,$$

where $\Sigma(s, y) := \begin{pmatrix} \text{diag}(s)\sigma(y) \\ a(y) \end{pmatrix} \in \mathbb{R}^{(n+m) \times n}$.

Theorem B.1. *Let $\Phi : \mathbb{R}^n_{++} \to \mathbb{R}$ and $u : [0, T] \times \mathbb{R}^n_{++} \times \mathbb{R}^m \to \mathbb{R}$ be a solution of the Cauchy problem*

$$\partial_t u + \mathcal{L}u - r(y)u = 0, \qquad\qquad on \quad [0, T) \times \mathbb{R}^n_{++} \times \mathbb{R}^m,$$
$$u(T, s, y) = \Phi(s), \qquad\qquad on \quad \mathbb{R}^n_{++} \times \mathbb{R}^m.$$

(i) Let $x := u(0, s_0, y)$ and $\pi = (\pi_t)_{t \in [0,T]}$ be a portfolio process satisfying

$$\pi_t^\top \sigma(Y_t) = D_s u^\top \operatorname{diag}(S_t)\sigma(Y_t) + D_y u^\top a(Y_t),$$

where $D_s u$ and $D_y u$ are evaluated at (t, S_t, Y_t). Then x is the replicating cost and π is the hedging portfolio. Indeed,

$$X_t^{x,\pi} = u(t, S_t, Y_t), \qquad t \in [0, T].$$

In particular,

$$X_T^{x,\pi} = \Phi(S_T).$$

(ii) Moreover, if $(H_t X_t^{x,\pi})_{t \in [0,T]}$ is a $\mathbb{P}$-martingale, then

$$u(t, s, y) = \mathbb{E}^{\mathbb{Q}}\left[\exp\left(-\int_t^T r(Y_v)dv\right)\Phi(S_T)\,\middle|\, S_t = s, Y_t = y\right]. \tag{44}$$

In particular, the replicating cost x is given by

$$x := u(0, s_0, y) = \mathbb{E}^{\mathbb{Q}}\left[\exp\left(-\int_0^T r(Y_v)dv\right)\Phi(S_T)\right]. \tag{45}$$

Proof. By the Ito formula,

$$\frac{1}{S_t^0}u(t, S_t, Y_t) = u(0, s_0, y) + \int_0^t \frac{1}{S_v^0}\left\{(\partial_t u + \mathcal{L}u - ru)(v, S_v, Y_v)dv + \left(D_s u^\top \operatorname{diag}(S_v)\sigma(Y_v) + D_y u^\top a(Y_v)\right)dW_v^{\mathbb{Q}}\right\}$$

$$= u(0, s_0, y) + \int_0^t \frac{1}{S_v^0}\left(D_s u^\top \operatorname{diag}(S_v)\sigma(Y_v) + D_y u^\top a(Y_v)\right)dW_v^{\mathbb{Q}}$$

$$= x + \int_0^t \frac{1}{S_v^0}\pi_v^\top \sigma(Y_v)dW_v^{\mathbb{Q}}$$

$$= \frac{X_t^{x,\pi}}{S_t^0},$$

which proves (i). If $(H_t X_t^\pi)_{t \in [0,T]}$ is a $\mathbb{P}$-martingale, then

$$u(t, S_t, Y_t) = X_t^{x,\pi}$$

$$= \frac{1}{H_t}\mathbb{E}_t\left[H_T X_T^{x,\pi}\right]$$

$$= \mathbb{E}_t^{\mathbb{Q}}\left[\exp\left(-\int_t^T r(Y_v)dv\right)\Phi(S_T)\right].$$

By the Markov property of (S, Y), we obtain (44) and (45). □

C Appendix: Relationship between stochastic control methods and martingale methods

In this appendix, we verify that the terminal wealth obtained via the dynamic programming approach matches the optimal terminal wealth derived from martingale duality methods. We prove the result in complete stochastic factor models given by (2).

Theorem C.1. *Let $V : [0,T] \times (0,\infty) \times \mathbb{R}^m$ be a classical solution to the HJB equation. Let $\hat{\pi} : [0,T] \times (0,\infty) \times \mathbb{R}^m$ be a candidate for optimal investment strategies. Furthermore, we assume that*

$$\lim_{t \nearrow T} V_x(t,x,y) = U'(x), \quad (x,y) \in (0,\infty) \times \mathbb{R}^m$$

and $(H_t X_t^{x,\hat{\pi}})_{t \in [0,T]}$ is a martingale for some $x > 0$. Then

$$X_T^{x,\hat{\pi}} = I(\hat{\lambda} H_T),$$

and $\hat{\pi}$ is the optimal feedback strategy.

Proof. We rewrite the HJB equation as

$$V_t + r(y)x V_x + b(y)^\top D_y V + \frac{1}{2}\mathrm{Tr}\left(aa^\top(y)D_{yy}^2 V\right) - \frac{V_{xx}}{2}|\sigma^\top(y)\hat{\pi}(t,x,y)|^2 = 0. \tag{46}$$

Differentiating (46) with respect to x, we have

$$\begin{aligned}
V_{tx} + r(y)V_x &+ r(y)x V_{xx} + b(y)^\top D_y V_x + \frac{1}{2}\mathrm{Tr}\left(aa^\top(y)D_{yy}^2 V_x\right) \\
&+ \hat{\pi}^\top \sigma(\theta V_{xx} + a^\top D_y V_{xx}) + \frac{V_{xxx}}{2}|\sigma^\top(y)\hat{\pi}(t,x,y)|^2 = 0.
\end{aligned} \tag{47}$$

That is,

$$V_{tx} + r(y)V_x + \mathcal{L}^{\hat{\pi}} V_x = 0,$$

where $\mathcal{L}^\pi$ is a controlled generator:

$$\mathcal{L}^\pi f := r(y)x f_x + b(y)^\top D_y f + \frac{1}{2}\mathrm{Tr}\left(aa^\top(y)D_{yy}^2 f\right) + \pi^\top \sigma(y)\left(\theta(y)f_x + a^\top(y)D_y f_x\right) + \frac{1}{2}|\sigma^\top(y)\pi|^2 f_{xx} = 0.$$

By the Ito formula,

$$\begin{aligned}
dV_x(t, X_t^{x,\hat{\pi}}, Y_t) &= (V_{xt} + \mathcal{L}^{\hat{\pi}_t} V_x)dt + \left\{V_{xx}\hat{\pi}_t^\top \sigma(Y_t) + D_y V_x^\top a(Y_t)\right\}dW_t \\
&= V_x\left(-r(Y_t)dt - \theta(Y_t)^\top dW_t\right).
\end{aligned}$$

Therefore,

$$V_x(t, X_t^{x,\hat{\pi}}, Y_t) = V_x(0,x,y)H_t.$$

Let $t \nearrow T$, then

$$U'(X_T^{x,\hat{\pi}}) = V_x(0,x,y)H_T,$$

which leads to

$$X_T^{x,\hat{\pi}} = I\left(V_x(0,x,y)H_T\right).$$

Because $H_t X_t^{x,\hat{\pi}}$ is a martingale,

$$x = \mathbb{E}\left[H_T X_T^{x,\hat{\pi}}\right] = \mathbb{E}\left[H_T I\left(V_x(0,x,y)H_T\right)\right]$$

holds and $\hat{\lambda} = V_x(0,x,y)$ by the uniqueness of $\hat{\lambda}$. We have completed the proof. $\qquad\square$

D Appendix: Matrix Riccati equation

We recall some facts stated in [21] about matrix Riccati differential equations. Let $T > 0$, $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times m}$, $C \in \mathbb{R}^{m \times n}$. We consider an $n \times n$ matrix solution $P = P(\cdot\,;T) : [0,T] \to \mathbb{R}^{n \times n}$ of the Riccati differential equation

$$\begin{aligned}
\dot{P}(t) - P(t)BB^\top P(t) + A^\top P(t) + P(t)A + C^\top C &= 0, \quad t \in [0,T], \\
P(T) &= 0.
\end{aligned} \tag{48}$$

First, we state the existence and uniqueness of (48).

Theorem D.1. *(i) There exists a nonnegative unique solution $P : [0, T] \to \mathbb{S}_+^n$ to (48).*

(ii) For any $0 \leq t_1 \leq t_2 \leq T$,

$$0 \leq P(t_2) \leq P(t_1)$$

holds.

Next, we state the asymptotic behaviors of the solution.

Definition D.1. (i) The pair (A, B) is said to be stabilizable if a matrix $L \in \mathbb{R}^{m \times n}$ exists such that $A + BL$ is stable (i.e., all its eigenvalues have negative real parts).

(ii) The pair (C, A) is said to be detectable if a matrix $F \in \mathbb{R}^{n \times m}$ exists such that $FC + A$ is stable.

Theorem D.2. *(i) If (A, B) is stabilizable, then there exists a finite nonnegative matrix $P_\infty \in \mathbb{S}_+^n$ such that*

$$\lim_{T \nearrow \infty} P(t; T) = P_\infty$$

for any t, and P_∞ satisfies the algebraic Riccati equation

$$-P_\infty BB^\top P_\infty + A^\top P_\infty + P_\infty A + C^\top C = 0.$$

(ii) If (A, B) is stabilizable and (C, A) is detectable, then $A - BB^\top P_\infty$ is stable.

The following theorem is Lemma 4.4 in [22].

Theorem D.3. *Let $K : [0, \infty) \to \mathbb{R}^{n \times n}$, $g : [0, \infty) \to \mathbb{R}^n$, and $f : [0, \infty) \to \mathbb{R}^n$ be a solution of the ordinary differential equation*

$$\dot{f}(t) = K(t) f(t) + g(t), \quad f(0) = 0.$$

Suppose that $K(t) \to \tilde{K} \in \mathbb{R}^{n \times n}$ and $g(t) \to \tilde{g} \in \mathbb{R}^n$ as $t \nearrow \infty$ and $\tilde{K}$ is a stable matrix. Then there exists $\tilde{f} = \lim_{t \nearrow \infty} f(t)$ that satisfies the equation

$$\tilde{K} \tilde{f} + \tilde{g} = 0.$$

Acknowledgement

References

[1] Ahn, D. H., Dittmar, R. F., & Gallant, A. R. (2002). Quadratic term structure models: Theory and evidence. The Review of financial studies, 15(1), 243-288.

[2] Branger, N., Chen, A., Mahayni, A., & Nguyen, T. (2023). Optimal collective investment: an analysis of individual welfare. Mathematics and Financial Economics, 17(1), 101-125.

[3] Bian, B., & Zheng, H. (2015). Turnpike property and convergence rate for an investment model with general utility functions. Journal of Economic Dynamics and Control, 51, 28-49.

[4] Bian, B., & Zheng, H. (2019). Turnpike property and convergence rate for an investment and consumption model. Mathematics and Financial Economics, 13(2), 227-251.

[5] Cox, J. C., & Huang, C. F. (1992). A continuous-time portfolio turnpike theorem. Journal of Economic Dynamics and Control, 16(3-4), 491-507.

[6] Detemple, J. B., Garcia, R., & Rindisbacher, M. (2003). A Monte Carlo method for optimal portfolios. The journal of Finance, 58(1), 401-446.

[7] Detemple, J., & Rindisbacher, M. (2010). Dynamic asset allocation: Portfolio decomposition formula and applications. The Review of Financial Studies, 23(1), 25-100.

[8] Dorfman, R., Samuelson, P. A., & Solow, R. M. (1958). Linear programming and economic analysis. New York, McGraw-Hill.

[9] Dybvig, P. H., Rogers, L. C. G., & Back, K. (1999). Portfolio turnpikes. The Review of Financial Studies, 12(1), 165-195.

[10] Fukaya, R. (2005). Application of Stochastic Flows to Optimal Portfolio Strategies. Journal of mathematical sciences, the University of Tokyo, 12(3), 349-397.

[11] Föllmer, H. & Schied, A. (2025). Stochastic Finance: An Introduction in Discrete Time. De Gruyter.

[12] Geng, T., & Zariphopoulou, T. (2025). Temporal and spatial turnpikes in Ito-diffusion markets under forward performance criteria. Numerical Algebra, Control and Optimization, 15(1), 243-272.

[13] Guasoni, P., Kardaras, C., Robertson, S., & Xing, H. (2014). Abstract, classic, and explicit turnpikes. Finance and stochastics, 18, 75-114.

[14] Hakansson, N. H. (1974). Convergence to isoelastic utility and policy in multiperiod portfolio choice. Journal of Financial Economics, 1(3), 201-224.

[15] Honda, T., & Kamimura, S. (2011). On the verification theorem of dynamic portfolio-consumption problems with stochastic market price of risk. Asia-Pacific Financial Markets, 18(2), 151-166.

[16] Huberman, G., & Ross, S. (1983). Portfolio Turnpike Theorems, Risk Aversion, and Regularly Varying Utility Functions. Econometrica, 51(5), 1345-1361.

[17] Huang, C. F., & Zariphopoulou, T. (1999). Turnpike behavior of long-term investments. Finance and Stochastics, 3, 15-34.

[18] Jensen, B. A., & Nielsen, J. A. (2016). How suboptimal are linear sharing rules?. Annals of Finance, 12(2), 221-243.

[19] Jin, X. (1998). Consumption and portfolio turnpike theorems in a continuous-time finance model. Journal of Economic Dynamics and Control, 22(7), 1001-1026.

[20] Karatzas, I., & Shreve, S. E., (1998). Methods of mathematical finance (Vol. 39, pp. xvi+-407). New York: Springer.

[21] Kucera, V. (1973). A review of the matrix Riccati equation. Kybernetika, 9(1), 42-61.

[22] Kuroda, K., & Nagai, H. (2002). Risk-sensitive portfolio optimization on infinite time horizon. Stochastics and Stochastic Reports, 73(3-4), 309-331.

[23] Karatzas, I., Ocone, D. L., & Li, J. (1991). An extension of Clark's formula. Stochastics: An International Journal of Probability and Stochastic Processes, 37(3), 127-131.

[24] Lakner, P. (1998). Optimal trading strategy for an investor: the case of partial information. Stochastic Processes and their Applications, 76(1), 77-97.

[25] Leippold, M., & Wu, L. (2002). Asset pricing under the quadratic class. Journal of Financial and Quantitative Analysis, 37(2), 271-295.

[26] Leland, H. (1972). On turnpike portfolios. In: Szego, G., Shell, K. (eds.) Mathematical Methods in Investment and Finance, p. 24. North-Holland, Amsterdam

[27] Liptser, R. S., & Shiryaev, A. N. (2013). Statistics of random processes: I. General theory (Vol. 5). Springer Science & Business Media.

[28] Liu, R., & Muhle-Karbe, J. (2013). Portfolio choice with stochastic investment opportunities: A User's guide. arXiv preprint arXiv:1311.1715.

[29] Mao, X. (2007). Stochastic differential equations and applications. Elsevier.

[30] Merton, R. C. (1969). Lifetime portfolio selection under uncertainty: The continuous-time case. The review of Economics and Statistics, 247-257.

[31] Merton, R. (1971). Optimum consumption and portfolio rules in a continuous-time model. Journal of Economic Theory, 3(4), 373-413.

[32] Mossin, J. (1968). Optimal Multiperiod Portfolio Policies. The Journal of Business, 41(2), 215-229.

[33] Nagai, H. (2003). Optimal strategies for risk-sensitive portfolio optimization problems for general factor models. SIAM journal on control and optimization, 41(6), 1779-1800.

[34] Nualart, D. (2006). The Malliavin calculus and related topics. Berlin, Heidelberg: Springer Berlin Heidelberg.

[35] Neumann, J. V. (1945). A model of general economic equilibrium. The Review of Economic Studies, 13(1), 1-9.

[36] Ocone, D. L., & Karatzas, I. (1991). A generalized Clark representation formula, with application to optimal portfolios. Stochastics: An International Journal of Probability and Stochastic Processes, 34(3-4), 187-220.

[37] Putschögl, W., & Sass, J. (2008). Optimal consumption and investment under partial information. Decisions in Economics and Finance, 31(2), 137-170.

[38] Qin, L., & Linetsky, V. (2017). Long-term risk: A martingale approach. Econometrica, 85(1), 299-312.

[39] Qin, L., & Linetsky, V. (2017). Long-term factorization of affine pricing kernels. Mathematics and Financial Economics, 11(4), 479-498.

[40] Qin, L., & Linetsky, V. (2018). Long-term factorization in Heath-Jarrow-Morton models. Finance and Stochastics, 22(3), 621-641.

[41] Ross, S. A. (1974). Portfolio turnpike theorems for constant policies. Journal of Financial Economics, 1(2), 171-198.

[42] Robertson, S., & Xing, H. (2017). Long-term optimal investment in matrix valued factor models. SIAM Journal on Financial Mathematics, 8(1), 400-434.

[43] Sass, J., & Haussmann, U. G. (2004). Optimizing the terminal wealth under partial information: The drift process as a continuous time Markov chain. Finance and Stochastics, 8(4), 553-577.

[44] Sun, J., & Yong, J. (2020). Stochastic linear-quadratic optimal control theory: Open-loop and closed-loop solutions. Springer Nature.

[45] Zariphopoulou, T. (2009). Optimal asset allocation in a stochastic factor model - an overview and open problems. In H. Albrecher, W. Runggaldier & W. Schachermayer (Ed.), Advanced Financial Modelling (pp. 427-456). Berlin, New York: De Gruyter.