Appl Math Optim (2014) 70:511–537
DOI 10.1007/s00245-014-9248-2

On Mean-Variance Hedging of Bond Options with Stochastic Risk Premium Factor

Shin Ichi Aihara · Arunabha Bagchi · Suresh K. Kumar

Published online: 13 April 2014

Abstract We consider the mean-variance hedging problem for pricing bond options using the yield curve as the observation. The model considered contains infinite-dimensional noise sources with the stochastically- varying risk premium. Hence our model is incomplete. We consider mean-variance hedging under the real world measure and obtain an explicit form of the optimal hedging strategy.

Keywords Mean-variance hedging · Indifference price · Kalman filter · Infinite-dimensional HMJ · Stochastic risk premium

Mathematics Subject Classification 91G20 · 91G80 · 60H15

1 Introduction

We study in this paper the optimal hedging problem for pricing options on bonds using the yield curve as the observation. There are various optimal hedging formulations

S. I. Aihara
Tokyo University of Science, Suwa, Toyohira 5000-1, Chino, Nagano, Japan
e-mail: aihara@rs.suwa.tus.ac.jp

A. Bagchi (✉)
FELab and Department of Applied Mathematics, University of Twente,
P.O. Box 217, 7500 Enschede, The Netherlands
e-mail: a.bagchi@utwente.nl

S. K. Kumar
Department of Mathematics, Indian Institute of Technology,
Mumbai 400 076, India
e-mail: suresh@math.iitb.ac.in

available in the literature. In this paper we only consider the mean-variance formulation. The mean-variance hedging problem has been studied under a risk-neutral measure. However since risk-neutral measure does not represent a statistical description of market events, the profit and loss of a portfolio may have a large variance while its "risk neutral" variance of the hedging error can be small. Therefore the mean-variance hedging under a real world measure has been introduced and studied extensively already, see for example [8,9].

In this article, we study the mean-variance hedging problem for the European option with bond as underlying asset. We use the general affine term structure model from [2]. Since general affine term structure model introduced in [2] has infinite noise sources, it describes an incomplete bond market, see [3,19] . This means one cannot perfectly hedge all contingent claims, for example options. The model [2] is described in the risk-neutral probability space. But as mentioned above mean-variance hedging has to be performed in the real world probability space. Hence one need to specify the bond market model in the real world. To describe the model given in [2] in the real world probability space, we use the risk-premium of the market.

Many statistical studies have been performed for the modeling of the risk-premium term, see [5,10,11]. Recently "predictability" in bond returns has been studied by Cochrane and Piazzesi [6]. Consistent with the findings of [6], Collin-Dufresne and Goldstein proposed a new model for the stochastic risk-premium factor, which is driven by feedback of noise sources of factor model in [7] . In this paper we adopt this stochastic premium model.

The stochastic risk-premium is not a tradable asset. Furthermore our market data consists of a finite number of bond data and the yield curve. This implies that the stochastic risk premium can not be reconstructed from the finite-dimensional observation data even though the randomness of the stochastic risk premium contains all components of the factor's randomness. Hence we have to study mean-variance hedging problem with partial information of the stochastic risk premium.

There are several articles dealing with partial information in finance, when the underlying assets are stocks, e.g. [12,13,15,17]. In our present paper where the underlying assets are bonds, our observation is the usual yield curve data for the zero-coupon bond and the time-series data of bonds used for constructing the self-financing portfolio. Using these data, we propose the Kalman filter technique to estimate the stochastic risk -premium. We apply the derived infinite-dimensional filter to the mean variance hedging using a method similar to the method proposed by Pham [17].

Our paper is organized as follows. In Sect. 2, we describe our bond market model in the real world. In Sect. 3, we review the filtering problem given in [2] and construct the Kalman filter algorithm for estimating the risk premium for the simple case that our observation information only contains the bond data. The mean-variance hedging problem is studied for European type options and Sect. 4 is devoted to calculate the optimal hedging strategy similar to the Black-Scholes delta. Augmenting the yield curve data and the portfolio as the new observation data, we reformulate the Kalman filter and the self-financing optimal mean variance portfolio is obtained in Sect. 5. In Sect. 6, we finally demonstrate the simulation studies to show the feasibility of the proposed hedging procedure.

2 Market Price of Risk

Consider the probablity space $(\Omega, \mathcal{F}, \mathcal{P})$ endowed with the filtration $\mathcal{F}_{t\geq 0}$. The above probability space represents the 'real world'. As stated in [12] for hedging the model is taken under the above *original* ("real-world" or "objective") measure, and hence we describe our model under the real-world measure.

The time variable t is defined on $[0, t_f]$ and the time-to-maturity variable x is defined on $G =]0, \hat{T}[$ and also the extended region $\tilde{G} =]0, \hat{T} + t_f[$[1]. We consider the Hilbert space $L^2(\tilde{G})$ (the space of square summable functions on $\tilde{G}$) with the inner product $(\cdot, \cdot)$ and $H^m(\tilde{G})$ (space of functions with their m-th derivatives in $L^2(\tilde{G})$). Our model for the instantaneous forward rate $f(t, x)$ for the time-to-maturity $x \in G =]0, \hat{T}[$ is

$$df(t,x) = \frac{\partial f(t,x)}{\partial x}dt + (\frac{1}{2}\frac{d}{dx}\tilde{q}(x) - \lambda(t)q_\lambda(x))dt + dw(t,x), \tag{1}$$

$$d\lambda(t) = (b + a\lambda(t))dt + (\sigma_\lambda(\cdot), dw(t,\cdot)),\ \lambda(0) = \lambda_o, \tag{2}$$

where q_λ, σ_λ are functions satisfying (A-4) given in Proposition 1, w is a two parameter Brownian motion process under $\mathcal{P}$, defined in $L^2(\tilde{G})$ with

$$E\{(\phi_1, w(t,\cdot))(\phi_2, w(t,\cdot))\} = t(\phi_1, Q\phi_2), \forall\phi_1, \phi_2 \in L^2(\tilde{G}), \tag{3}$$

with

$$Q\phi(\cdot) = \int_{\tilde{G}} q(x,y)\phi(y)dy, \quad \text{for } \phi \in L^2(\tilde{G}) \tag{4}$$

and[2] $\tilde{q}(x)$ is given by

$$\tilde{q}(x) = \int_0^x \int_0^x q(x_1, x_2)dx_1 dx_2. \tag{5}$$

Note that our model incorporates the market price of risk $\lambda(\cdot)$. We are using the model for the market price of risk proposed by Collin-Dufresne and Goldstein's working paper [7]. For simplicity, the risk-premium is given by the product of the stochastic part $\lambda(t)$ and the deterministic function $q_\lambda(x)$. From the fact that the bond risk premium contains all forward rates random property as suggested in [6], stochastic risk-premium is specified by (2).

Proposition 1 *We assume*

(A-1) $f_o \in L^2(\Omega, H^1(\tilde{G}))$

[1] From the property of the first order hyperbolic systems, the spatial region $\tilde{G} =]0, \hat{T} + t_f[$ is shrinking to $G =]0, \hat{T}[$ from $t = 0$ to t_f. See [1].

[2] The smoothness for q is given in Proposition 2.1.

(A-2) $\int_{\tilde{G}} \frac{\partial^2 q(x,y)}{\partial x \partial y}\Big|_{y=x} dx < \infty$
(A-3) $\lambda_o \in L^2(\Omega; R^1)$ *and*
(A-4) $q_\lambda \in H^1(\tilde{G}),\ \sigma_\lambda \in L^2(\tilde{G})$, *a and b are constants.*

Then

$$f \in L^2\left(\Omega; C([0, t_f]; H^1(]0, \hat{T}[))\right), \tag{6}$$

$$\lambda \in L^2\left(\Omega; C([0, t_f]; R^1)\right) \tag{7}$$

Proof The proof of this proposition is shown in Appendix 1. □

Set

$$d\tilde{w}(t, x) = dw(t, x) - \lambda(t) q_\lambda(x) dt. \tag{8}$$

Then we have the following Girsanov's theorem.

Proposition 2 *In addition to (A-4), we assume*

(A-5) $(q_\lambda, Q^{-1} q_\lambda) < \infty$,

where

$$Q^{-1}(\cdot) = \sum_{i=1}^{\infty} \frac{1}{\lambda_i} e_i (e_i, \cdot)$$

and $q(x, y) = \sum_{i=1}^{\infty} \lambda_i e_i(x) e_i(y)$ *for some orthonormal basis in* $L^2(\tilde{G})$. *Under the probability measure* $\tilde{\mathcal{P}}$ *defined by*

$$\frac{d\tilde{\mathcal{P}}}{d\mathcal{P}} = \exp\{\int_0^t \lambda(s)(q_\lambda, Q^{-1} dw(s, \cdot)) - \frac{1}{2} \int_0^t \lambda^2(s)(q_\lambda, Q^{-1} q_\lambda) ds\}, \tag{9}$$

$\tilde{w}$ *given by (8) is also a Brownian motion process .*

Proof This proof is presented in Appendix 2. □

In view of Proposition 2, it follows that under the probaility measure $\tilde{\mathcal{P}}$, the forward rate dynamics is given by the solution of

$$df(t, x) = \frac{\partial f(t, x)}{\partial x} dt + \frac{1}{2} \frac{d}{dx} \tilde{q}(x) dt + d\tilde{w}(t, x) \tag{10}$$

$$f(0, x) = f_o(x) \tag{11}$$

where $\tilde{w}(t, x)$ denotes the two parameter Brownian motion under $\tilde{\mathcal{P}}$ with

$$\tilde{E}\{\tilde{w}(t, x_1) \tilde{w}(t, x_2)\} = q(x_1, x_2) t.$$

The bond price $P(t, T)$ for a fixed maturity T is given by

$$P(t, T) = \exp\{-\int_0^{T-t} f(t, x)dx\}.$$

Noting that the exact short rate is given by $r(t) = f(t, 0)$, we find that

$$\tilde{P}(t, T) = \frac{P(t, T)}{\exp\{\int_0^t f(s, 0)ds\}} \tag{12}$$

becomes an $\tilde{\mathcal{F}}_t$-martingale, where $\tilde{\mathcal{F}}_t = \sigma\{f(s, \cdot); 0 \le s \le t\}$. Hence $\tilde{\mathcal{P}}$ is a risk-neutral measure.

3 Mean-Variance Hedging

We consider a simple form of a self financing portfolio which is constructed from two bonds to make the presentation of the article simple. Note that the case of portfolio with more than two bonds can be done along the same lines. Also we consider the Mean-Variance hedging for European call option and the case of general payouts,e.g. swaptions,interest rate caps, etc can be treated similarly. Our hedging problem is more challenging because, since the stochastic risk premium can not be directly observed from the market, the usual payout, e.g. call option can not be replicated in the usual manner. Hence we need to estimate the movement of the stochastic risk-premium from the observed market data and then we construct the portfolio to replicate the given pay off for minimizing the mean square error of the wealth of portfolio and the pay off, i.e., the mean variance hedging is introduced.

3.1 Self-financing Portfolio and Observation Data

We construct a self-financing portfolio for hedging the European type options. Consider a European call option with maturity T_m on a T_M-bond $P(t, T_M)$ such that $T_m < T_M$. For hedging we use a portfolio of the T_m-bond and T_M-bond. For clarity in the presentation, we first give an illustration using analogous discrete version of the portfolio.

- At $t = 0$ we set up a portfolio made up of x_0 units of $P(0, T_m)$. Hence the itinial wealth $V(0)$ becomes

$$V(0) = P(0, T_m)x_0.$$

- Now from the initial wealth, we purchase $\theta(\delta)$ units of $P(0, T_M)$. Hence at $t = \delta$, our wealth becomes

$$V(\delta) = \frac{P(0, T_m)x_0 - P(0, T_M)\theta(\delta)}{P(0, T_m)} P(\delta, T_m) + P(\delta, T_M)\theta(\delta).$$

The self-financing portfolio becomes for $t = \delta, 2\delta, \cdots$

$$\begin{cases} V(0) = P(0, T_m)x_0 \\ V(t) = \dfrac{V(t-\delta) - P(t-\delta, T_M)\theta(t)}{P(t-\delta, T_m)} P(t, T_m) + \theta(t)P(t, T_M) \end{cases} \tag{13}$$

where $\theta(t)$ is a $\mathcal{F}_{t-\delta}$ measurable process ($\mathcal{F}_t$-predictable process) and $\mathcal{F}_t$ denotes the data form the market at time t. From (13) it follows that the discounted wealth, i.e., $\frac{V(t)}{P(t,T_m)}$ satisfies

$$\frac{V(t)}{P(t, T_m)} = \frac{V(t-\delta)}{P(t-\delta, T_m)} + \theta(t)\Big\{\frac{P(t, T_M)}{P(t, T_m)} - \frac{P(t-\delta, T_M)}{P(t-\delta, T_m)}\Big\}.$$

Hence

$$V(n\delta) = P(n\delta, T_m)\left(x_0 + \sum_{i=1}^{n} \theta(i\delta) \left\{ \frac{P(i\delta, T_M)}{P(i\delta, T_m)} - \frac{P((i-1)\delta, T_M)}{P((i-1)\delta, T_m)} \right\} \right). \tag{14}$$

It is also possible to represent the wealth process:

$$\begin{aligned} V(t) - V(t-\delta) = {} & \frac{V(t-\delta)}{P(t-\delta, T_m)}\Big(P(t, T_m) - P(t-\delta, T_m)\Big) \\ & + \theta(t)\left(P(t, T_M) - P(t-\delta, T_M)\frac{P(t, T_m)}{P(t-\delta, T_m)}\right), \end{aligned}$$

i.e.,

$$\begin{aligned} V(n\delta) = {} & P(0, T_m)x_0 + \sum_{i=1}^{n} \left\{ \frac{V((i-1)\delta)}{P((i-1)\delta, T_m)} (P(i\delta, T_m) - P((i-1)\delta, T_m)) \right\} \\ & + \sum_{i=1}^{n} \theta(i\delta)\Big(P(i\delta, T_M) - P((i-1)\delta, T_M) \\ & - \frac{P((i-1)\delta, T_M)}{P((i-1)\delta, T_m)}(P(i\delta, T_m) - P((i-1)\delta, T_m))\Big) \end{aligned} \tag{15}$$

Hence noting that $\theta(t)$ is $\mathcal{F}_{t-\delta}$ measurable, from (14) the continuous version of $V(n\delta)$ becomes

$$V(t) = P(t, T_m)\left[x_0 + \int_0^t \theta(s) d\frac{P(s, T_M)}{P(s, T_m)} \right], \tag{16}$$

where $\theta(t) \in U_{ad}$ It also follows from (15) that

$$V(t) = P(0, T_m)x_0 + \int_0^t \left\{ (\frac{V(s)}{P(s, T_m)} - \theta(s)\frac{P(s, T_M)}{P(s, T_m)})dP(s, T_m) \right.$$
$$\left. + \theta(s)dP(s, T_M) \right\}. \tag{17}$$

– Now our portfolio is $(x_0, \theta(t))$. The mean-variaince hedging is to find the constant initial investment x_0 for $P(0, T_m)$ and the dynamic investment $\theta(s)$ for $P(s, T_M)$ under the self-financing situation.

Let Y_t denote the sigma algebra generated by the observation data from the market.

Proposition 3 *We assume that*

$$\theta \in U_{ad} = \left\{ \theta; \theta(t) \; is \; \mathcal{Y}_t - mesurable, with \; E\{\int_0^{T_M} \theta^2(t)P^2(t, T)dt\} < \infty \right\}.$$

Then the portfolio $(x_0, \ \theta(t))$ *is self-financing.*

Proof From (17), we find that

$$dV(t) = \left(\frac{V(t)}{P(t, T_m)} - \theta(t)\frac{P(t, T_M)}{P(t, T_m)} \right) dP(t, T_m) + \theta(t)dP(t, T_M).$$

This implies that $(\frac{V(\cdot)}{P(\cdot, T_m)} - \theta(\cdot)\frac{P(\cdot, T_M)}{P(\cdot, T_m)}, \theta(\cdot))$ of $(P(\cdot, T_m), P(\cdot, T_M))$ is self-financing, i.e. $(x_0, \ \theta(\cdot))$ is self-financing. □

Our information is $P(t, T_m)$ and $P(t, T_M)$ for $0 \leq t \leq T_m$. Since the stochastic risk premium $\lambda(\cdot)$ given by (2) is not a tradable asset, one need to estimate $\lambda(\cdot)$ from the two bond price processes $P(\cdot, T_m)$ and $P(\cdot, T_M)$. To find the relation between $\lambda(\cdot)$ and these bond data, we use the observation data

$$\tilde{Y}(t) = -\log \frac{P(t, T_M)}{P(t, T_m)}. \tag{18}$$

Noting that

$$\tilde{Y}(t) = \int_{T_m - t}^{T_M - t} f(t, x)dx,$$

from (1) we have

$$
\begin{aligned}
d\tilde{Y}(t) &= -f(t, T_M - t)dt + f(t, T_m - t)dt + \int_{T_m-t}^{T_M-t} df(t,x)dx \\
&= \int_{T_m-t}^{T_M-t} \left(\int_0^x q(x,z)dz - \lambda(t) q_\lambda(x) \right) dxdt + \int_{T_m-t}^{T_M-t} dw(t,x)dx \\
&= \frac{1}{2} \left[\int_{T_m-t}^{T_M-t} \int_{T_m-t}^{T_M-t} q(x_1,x_2)dx_1dx_2 + 2 \int_0^{T_m-t} \int_{T_m-t}^{T_M-t} q(x_1,x_2)dx_1dx_2 \right] dt \\
&\quad + \int_{T_m-t}^{T_M-t} dw(t,x)dx - \lambda(t) \int_{T_m-t}^{T_M-t} q_\lambda(x)dxdt. \qquad (19)
\end{aligned}
$$

The observation process $\tilde{Y}(t)$ can be rewritten as

$$
d\tilde{Y}(t) = -\lambda(t)H(t)q_\lambda dt + \frac{1}{2}F(t)dt + H(t)dw(t,\cdot), \qquad (20)
$$

where

$$
\begin{aligned}
F(t) &= \int_{T_m-t}^{T_M-t} \int_{T_m-t}^{T_M-t} q(x_1,x_2)dx_1dx_2 + 2 \int_0^{T_m-t} \int_{T_m-t}^{T_M-t} q(x_1,x_2)dx_1dx_2, \\
H(t)\phi(t,\cdot) &= \int_{T_m-t}^{T_M-t} \phi(t,x)dx, \ \phi \in C([0,t_f, L^2(\hat{G})).
\end{aligned}
$$

Clearly the observation noise covariance $\bar{q}(\cdot)$ is given by

$$
\bar{q}(t) = \int_{T_m-t}^{T_M-t} \int_{T_m-t}^{T_M-t} q(x_1,x_2)dx_1dx_2. \qquad (21)
$$

Hence the Kalman filter equation for $\hat{\lambda}(t) = E\{\lambda(t)|\mathcal{Y}_t\}$, where $\mathcal{Y}_t = \sigma\{\tilde{Y}(s); 0 \leq s \leq t\}$ is given by

$$
\begin{aligned}
d\hat{\lambda}(t) &= a\hat{\lambda}(t)dt + bdt + K_\lambda(t)d\tilde{\ell}(t), \qquad (22) \\
K_\lambda(t) &= \left(-P_\lambda(t) \int_{T_m-t}^{T_M-t} q_\lambda(x)dx + \int_G \int_{T_m-t}^{T_M-t} \sigma_\lambda(x)q(x,y)dxdy \right) \bar{q}^{-1}(t), \qquad (23)
\end{aligned}
$$

where P_λ is error covariance described by the o.d.e. (see [4,14])

$$\frac{dP_\lambda(t)}{dt} = 2aP_\lambda(t) + \int_G \int_G \sigma_\lambda(x)q(x,y)\sigma_\lambda(y)dxdy - \Big(- P_\lambda(t) \int_{T_m-t}^{T_M-t} q_\lambda(x)dx + \int_G \int_{T_m-t}^{T_M-t} \sigma_\lambda(x)q(x,y)dxdy\Big)^2 \bar{q}^{-1}(t), \quad (24)$$

and the innovation process $\tilde{\ell}(t)$ is defined by

$$\tilde{\ell}(t) = \tilde{Y}(t) - \tilde{Y}(0) - \int_0^t \left\{ \frac{1}{2}\bar{q}(s) + \bar{q}_2(s) - \hat{\lambda}(s)\bar{q}_\lambda(s) \right\} ds, \quad (25)$$

$$\bar{q}_2(s) = \int_0^{T_m-s} \int_{T_m-s}^{T_M-s} q(x_1,x_2)dx_1dx_2, \quad (26)$$

$$\bar{q}_\lambda(s) = \int_{T_m-s}^{T_M-s} q_\lambda(x)dx. \quad (27)$$

It is well-known that, see [14]

$$\tilde{\ell}(\cdot) \text{ is a } \mathcal{Y}_t - \text{Brownian motion with incremental covariance } \bar{q}(\cdot) \quad (28)$$

From (16) and (18), we have

$$\begin{aligned} \frac{V_t}{P(t,T_m)} &= x_0 + \int_0^t \theta(s)\, d\Big(\frac{P(s,T_M)}{P(s,T_m)}\Big) \\ &= x_0 + \int_0^t \theta(s)\, de^{-\tilde{Y}(s)} \\ &= x_0 - \int_0^t \theta(s)e^{-\tilde{Y}(s)}d\tilde{Y}(s) + \frac{1}{2}\int_0^t \theta(s)e^{-\tilde{Y}(s)}\bar{q}(s)ds. \end{aligned}$$

It follows from (25) that

$$\tilde{V}_t = \frac{V_t}{P(t,T_m)} = x_0 - \int_0^t \theta(s)e^{-\tilde{Y}(s)}d\tilde{\ell}(s) + \int_0^t \theta(s)e^{-\tilde{Y}(s)}\{-\bar{q}_2(s) + \hat{\lambda}(s)\bar{q}_\lambda(s)\}ds. \quad (29)$$

The payoff at maturity T_m of the European call option $(P(T_m, T_M) - K)^+$ is represented as follows:

$$(P(T_m, T_M) - K)^+ = \Big(\frac{P(T_m, T_M)}{P(T_m, T_m)} - K\Big)^+ = \Big(\exp(-\tilde{Y}(T_m)) - K\Big)^+$$
$$= \left(\exp(-\tilde{\ell}(T_m) - \int_0^{T_m} \{\frac{1}{2}\bar{q}(s) + \bar{q}_2(s) - \hat{\lambda}(s)\bar{q}_\lambda(s)\}ds - \tilde{Y}(0)) - K\right)^+, \quad (30)$$

Hence, noting that $(P(T_m, T_M) - K)^+$ is a functional of $\tilde{\ell}$ and

$$\lim_{t\to T_m} \tilde{V}_t = V_{T_m},$$

we set the mean-variance cost as

$$J(t, x_0, \theta) = E\{|\tilde{V}_t - E\{(P(T_m, T_M) - K)^+|\mathcal{Y}_t\}|^2\}, \quad (31)$$

for $x_0 \in R^1_+, \theta \in U_{ad}$ where

$$U_{ad} = \left\{\theta; \theta(t) \text{ is } \mathcal{Y}_t - \text{measurable, with } E\{\int_0^{T_M} \theta^2(t)P^2(t, T_M)dt\} < \infty\right\}.$$

Before proceeding to solve the above optimization problem, we calculate the so called indifference price $E\{(P(T_m, T_M) - K)^+|\mathcal{Y}_t\}$.

3.2 Explicit Form for $E\{(P(T_m, T_M) - K)^+|\mathcal{Y}_t\}$

In order to obtain the explicit form of $E\{(P(T_m, T_M) - K)^+|\mathcal{Y}_t\}$, we use the Gaussian nature of $\tilde{\ell}(\cdot)$. So we need the following auxiliary proposition.

Proposition 4 *Let*

$$A(t) = e^{-at} \int_t^{T_m} e^{as}\bar{q}_\lambda(s)ds, \ t \geq 0.$$

Then

$$\int_t^{T_m} \hat{\lambda}(s)\bar{q}_\lambda(s)ds \ = \ A(t)\hat{\lambda}(t) + b\int_t^{T_m} A(s)ds + \int_t^{T_m} A(s)K_\lambda(s)d\tilde{\ell}(s). \quad (32)$$

Proof See Appendix 3. □

Denote the indifference price as follows.

$$\hat{H}^t_{T_m} = E\{(P(T_m, T_M) - K)^+ | \mathcal{Y}_t\}.$$

Also set

$$m_1(t, \tilde{\ell}(t)) = -\tilde{\ell}(t) - \frac{1}{2}\int_0^{T_m} \bar{q}(s)ds \tag{33}$$

$$m_2(t, \hat{\lambda}(s); 0 \le s \le t)) = \int_0^t \hat{\lambda}(s)\bar{q}_\lambda(s)ds - \tilde{Y}(0) - \int_0^{T_m} \bar{q}_2(s)ds + A(t)\hat{\lambda}(t) + b\int_t^{T_m} A(s)ds \tag{34}$$

and

$$\tilde{K} = e^{-m_1(t,\tilde{\ell}(t)) - m_2(t,\hat{\lambda}(s);0\le s\le t)} K.$$

Theorem 1 *The indifference price is given by*

$$\hat{H}^t_{T_m} = \exp\left(m_1(t, \tilde{\ell}(t)) + m_2(t, \hat{\lambda}(s); 0 \le s \le t)) + \frac{R(t)}{2}\right)N(d_1(t)) - KN(d_2(t)), \tag{35}$$

where $N(\cdot)$ denotes the commutative distribution function of the standard Normal distribution,

$$R(t) = \int_t^{T_m} \bar{q}(s)ds - 2\int_t^{T_m} \bar{q}(s)K_\lambda(s)A(s)ds + \int_t^{T_m} A^2(s)K_\lambda(s)\bar{q}(s)K_\lambda(s)ds, \tag{36}$$

and

$$d_1(t) = R^{-1/2}(t)(m_1(t, \tilde{\ell}(t)) + m_2(t, \hat{\lambda}(s); 0 \le s \le t)) - \log K + R(t)),$$
$$d_2(t) = R^{-1/2}(t)(m_1(t, \tilde{\ell}(t)) + m_2(t, \hat{\lambda}(s); 0 \le s \le t)) - \log K).$$

Proof From (30) and Proposition 4 we have

$$\hat{H}^t_{T_m} = E\left\{\left(\exp\{-\tilde{\ell}(t) + \int_0^t \hat{\lambda}(s)\bar{q}_\lambda(s)ds - \int_0^{T_m}\{\frac{1}{2}\bar{q}(s) + \bar{q}_2(s)\}ds - \tilde{Y}(0)\}\right.\right.$$
$$\left.\left. \times \exp\{-(\tilde{\ell}(T_m) - \tilde{\ell}(t)) + \int_t^{T_m} \hat{\lambda}(s)\bar{q}_\lambda(s)ds\} - K\right)^+ | \mathcal{Y}_t\right\}$$

$$= E\left\{\left(\exp\{-\tilde{\ell}(t) + \int_0^t \hat{\lambda}(s)\bar{q}_\lambda(s)ds\right.\right.$$
$$- \int_0^{T_m} (\frac{1}{2}\bar{q}(s) + \bar{q}_2(s))ds - \tilde{Y}(0) + A(t)\hat{\lambda}(t) + b\int_t^{T_m} A(s)ds\}$$
$$\left.\left.\times \exp\{-(\tilde{\ell}(T_m) - \tilde{\ell}(t)) + \int_t^{T_m} A(s)K_\lambda(s)d\tilde{\ell}(s)\} - K\right)^+ |\mathcal{Y}_t\right\}$$
$$= \exp(m_1(t, \tilde{\ell}(t)) + m_2(t, \hat{\lambda}(s); 0 \le s \le t))$$
$$\times E\left\{\left(\exp\{-(\tilde{\ell}(T_m) - \tilde{\ell}(t)) + \int_t^{T_m} A(s)K_\lambda(s)d\tilde{\ell}(s)\} - \tilde{K}\right)^+ |\mathcal{Y}_t\right\}.$$

Noting that $-(\tilde{\ell}(T_m) - \tilde{\ell}(t)) + \int_t^{T_m} A(s)K_\lambda(s)d\tilde{\ell}(s)$ is Gaussian with zero mean and covariance $R(t)$ given by (36), this result is easily derived. □

4 Explicit forms of Optimal Hedging Strategies

Noting that $\hat{H}^t_{T_M}$ is a $\mathcal{Y}_t$ martingale, and since $\int_0^{\cdot} \hat{\lambda}(s)\bar{q}`_\lambda(s)ds$ is of bounded variation it follows from Theorem 1 that

$$d\hat{H}^t_{T_m} = \frac{\partial \hat{H}^t_{T_m}}{\partial m_1}\frac{\partial m_1}{\partial \tilde{\ell}}d\tilde{\ell}(t) + \frac{\partial \hat{H}^t_{T_m}}{\partial m_2}\frac{\partial m_2}{\partial \hat{\lambda}}K_\lambda(t)d\tilde{\ell}(t) \tag{37}$$

Using the well-known property, for $i = 1, 2$,

$$\exp(m_1 + m_2 + R/2)\frac{\partial N(R^{-1/2}(m_1 + m_2 - \log K + R))}{\partial m_i}$$
$$-K\frac{\partial N(R^{-1/2}(m_1 + m_2 - \log K))}{\partial m_i} = 0,$$

we have

$$\frac{\partial \hat{H}^t_{T_m}}{\partial m_i} = \hat{H}^t_{T_m}, \ i = 1, 2.$$

Hence

$$d\hat{H}^t_{T_m} = \hat{H}^t_{T_m}(-1 + A(t)K_\lambda(t))d\tilde{\ell}(t). \tag{38}$$

Defining the hedging error process $\epsilon(t) = \tilde{V}_t - \hat{H}^t_{T_m}$, we obtain

$$d\epsilon(t) = \{\hat{H}^t_{T_m}(1 - A(t)K_\lambda(t)) - \theta(t)e^{-\tilde{Y}(t)}\}d\tilde{\ell}(t) \\ + \{-\bar{q}_2(t) + \hat{\lambda}(t)\bar{q}_\lambda(t)\}\theta(t)e^{-\tilde{Y}(t)}dt. \tag{39}$$

Applying Ito's lemma to $|\epsilon(t)|^2$, the hedging cost $J(t, x_0, \theta)$ becomes

$$J(t, x_0, \theta) = E\{|x_0 - \hat{H}^0_{T_m}|^2\} + 2E\{\int_0^t \{-\bar{q}_2(s) + \hat{\lambda}(s)\bar{q}_\lambda(s)\} \\ \times \theta(s)e^{-\tilde{Y}(s)}(\tilde{V}_s - \hat{H}^s_{T_m})ds\} \\ + E\{\int_0^t |\hat{H}^s_{T_m}(1 - A(s)K_\lambda(s)) - \theta(s)e^{-\tilde{Y}(s)}|^2\bar{q}(s)\}ds. \tag{40}$$

Now from the fact that the RHS of (40) is a quadratic function with respect to x_0 and θ we can derive the explicit form of the optimal hedging strategy:

Theorem 2 *The optimal hedging strategy* (x^o_0, θ^o) *is given by*

$$x^o_0 = \frac{P(0, T_M)}{P(0, T_m)} \exp\Bigg(A(0)\hat{\lambda}(0) + b\int_0^{T_m} A(s)ds - \frac{1}{2}\int_0^{T_m} \bar{q}(s)ds \\ - \int_0^{T_m} \bar{q}_2(s)ds + \frac{R(0)}{2}\Bigg)N(d_1(0)) - KN(d_2(0)) \tag{41}$$

and

$$\theta^o(t) = \Bigg\{\exp\Bigg(A(t)\hat{\lambda}(t) + b\int_t^{T_m} A(s)ds - \frac{1}{2}\int_t^{T_m} \bar{q}(s)ds \\ - \int_t^{T_m} \bar{q}_2(s)ds + \frac{R(t)}{2}\Bigg)\ N(d_1(t)) - \frac{P(t, T_m)}{P(t, T_M)}KN(d_2(t))\Bigg\} \\ \times \Bigg(1 - A(t)K_\lambda(t) + \frac{1}{\bar{q}(t)}\{-\bar{q}_2(t) + \hat{\lambda}(t)\bar{q}_\lambda(t)\}\Bigg) \\ - \frac{1}{\bar{q}(t)}\Big\{-\bar{q}_2(t) + \hat{\lambda}(t)\bar{q}_\lambda(t)\Big\}\frac{V_t}{P(t, T_M)}, \tag{42}$$

where $d_1(t)$ and $d_2(t)$ are represented by

$$d_1(t) = R^{-1/2}(t)(-\log \frac{P(t,T_m)}{P(t,T_M)} + A(t)\hat{\lambda}(t) + b\int_t^{T_m} A(s)ds - \frac{1}{2}\int_t^{T_m} \bar{q}(s)ds - \int_t^{T_m} \bar{q}_2(s)ds - \log K + R(t)) \tag{43}$$

$$d_2(t) = d_1 - R^{1/2}(t). \tag{44}$$

Proof From (40), we have

$$x_0^o = \hat{H}_{T_m}^o \tag{45}$$

$$\theta^o(t) = e^{\tilde{Y}(t)}\left\{\hat{H}_{T_m}^t(1 - A(t)K_\lambda(t)) - \frac{1}{\bar{q}(s)}\{-\bar{q}_2(t) + \hat{\lambda}(t)\bar{q}_\lambda(t)\}\{\tilde{V}_t - \hat{H}_{T_m}^t\}\right\}. \tag{46}$$

In order to implement the optimal $\theta^o(t)$, we represent $\tilde{\ell}$ as functions of $P(t, T_M)$ and $P(t, T_m)$ and filter output $\hat{\lambda}(t)$. From (25) and (18) we have

$$\tilde{\ell}(t) = \log\left[\frac{P(0,T_M)}{P(0,T_m)}\frac{P(t,T_m)}{P(t,T_M)}\right] - \int_0^t \{\frac{1}{2}\bar{q}(s) + \bar{q}_2(s) - \hat{\lambda}(s)\bar{q}_\lambda(s)\}ds. \tag{47}$$

Hence we get

$$m(t,\tilde{\ell}(t)) + m_2(t,\hat{\lambda}(s); 0 \le s \le t)) = -\log \frac{P(t,T_m)}{P(t,T_M)} + A(t)\hat{\lambda}(t) + b\int_t^{T_m} A(s)ds - \frac{1}{2}\int_t^{T_m} \bar{q}(s)ds - \int_t^{T_m} \bar{q}_2(s)ds) \tag{48}$$

$$e^{\tilde{Y}(t)}\hat{H}_{T_m}^t = \exp\{A(t)\hat{\lambda}(t) + b\int_t^{T_m} A(s)ds - \frac{1}{2}\int_t^{T_m} \bar{q}(s) - \int_t^{T_m} \bar{q}_2(s)ds + \frac{R(t)}{2}\}N(d_1(t)) - \frac{P(t,T_m)}{P(t,T_M)}KN(d_2(t)). \tag{49}$$

Substituting (48) and (49) into (46), we obtain this theorem. □

5 Yield Curve Data

In the previous section, we only used the purchased bond data for estimating $\lambda(\cdot)$. Usually from the market one also get the yield curve data. It is possible to include these data for estimating the market price of risk term $\lambda(t)$. According to the results obtained previously, the optimal θ^o is a function of the estimate $\hat{\lambda}(t)$ and the form of the optimal θ^o is easily adjusted by using the estimate $\hat{\lambda}(t)$ from $\tilde{Y}(t)$ and yield curve data. To perform the Kalman filter by using these data, we need to estimate the factor process $f(t, x)$ whose estimate is not explicitly included in the optimal gain θ^o. However we can expect that the estimation results for $\hat{\lambda}$ are well improved including the yield curve data in our observation.

5.1 Observation Data from Yield Curve and Optimal θ^o

We include continuously compounded yields on zero-coupon bonds with fixed time-to-maturity

$$y_i(t) = \frac{1}{\tau_i}\int_0^{\tau_i} f(t, x)dx, \text{ for } \tau_1 < \tau_2 < \cdots < \tau_m \tag{50}$$

as additional observations. It follows from (1) that

$$\begin{aligned} dy_i(t) = &\frac{1}{\tau_i}(f(t, \tau_i) - f(t, 0))dt + \frac{1}{2\tau_i}\tilde{q}(\tau_i)dt \\ &- \lambda(t)\frac{1}{\tau_i}\int_0^{\tau_i} q_\lambda(x)dxdt + \frac{1}{\tau_i}\int_0^{\tau_i} dw(t, x)dx. \end{aligned}$$

Let $\mathbf{Y}(t) = [y_1(t), y_2(t), \cdots, y_m(t), \tilde{Y}(t)]', t \geq 0$ denote the augmented observation process. Set $\mathcal{Y}_t = \sigma\{\mathbf{Y}(s); 0 \leq s \leq t\}$. Clearly $\mathbf{Y}(\cdot)$ is given by

$$d\mathbf{Y}(t) = \mathbf{H}_\delta f(t, \cdot)dt - \lambda(t)\mathbf{H}(t)q_\lambda dt + \frac{1}{2}\mathbf{F}(t)dt + \mathbf{H}(t)dw(t, \cdot), \tag{51}$$

where

$$\mathbf{H}_\delta[\cdot] = \begin{bmatrix} H_\delta[\cdot] \\ 0 \end{bmatrix} = \begin{bmatrix} [\frac{1}{\tau_i}\int_G(\delta(x-\tau_i) - \delta(x))(\cdot)dx]_{m\times 1} \\ 0 \end{bmatrix}_{(m+1)\times 1},$$

$$\mathbf{F}(t) = \begin{bmatrix} [\frac{1}{\tau_i}\tilde{q}(\tau_i)]_{m\times 1} \\ \int_{T_m-t}^{T_M-t}\int_{T_m-t}^{T_M-t} q(x_1, x_2)dx_1dx_2 + 2\int_0^{T_m-t}\int_{T_m-t}^{T_M-t} q(x_1, x_2)dx_1dx_2 \end{bmatrix}_{(m+1)\times 1},$$

$$\mathbf{H}(t)q_\lambda = \begin{bmatrix} [\frac{1}{\tau_i}\int_0^{\tau_i} q_\lambda dx]_{m\times 1} \\ \int_{T_m-t}^{T_M-t} q_\lambda dx \end{bmatrix}_{(m+1)\times 1}$$

and $\delta(x)$ is a delta function,i.e., $\int_G \delta(x-\tau_i) f(t,x)dx = f(t,\tau_i)$. The observation noise covariance $\mathbf{\Phi}(\cdot)$ is given by

$$\mathbf{\Phi}(t) = \begin{bmatrix} \frac{1}{\tau_i\tau_j}\int_0^{\tau_i}\int_0^{\tau_j} q(x_1,x_2)dx_1dx_2 & \frac{1}{\tau_i}\int_0^{\tau_i}\int_{T_m-t}^{T_M-t} q(x_1,x_2)dx_1dx_2 \\ \frac{1}{\tau_j}\int_{T_m-t}^{T_M-t}\int_0^{\tau_j} q(x_1,x_2)dx_1dx_2 & \int_{T_m-t}^{T_M-t}\int_{Tm-t}^{T_M-t} q(x_1,x_2)dx_1dx_2 \end{bmatrix}_{(m+1)\times(m+1)},$$

and by using the method proposed in [2],we can show that $\Phi(t)$ is invertible. Hence without adding the artificial observation noise, we can derive the Kalman filter equation for the augmented observation $\mathbf{Y}$ (see [2] for detail derivations). The Kalman filter is given by

$$d\begin{pmatrix} \hat{f}(t,x) \\ \hat{\lambda}(t) \end{pmatrix} = \begin{pmatrix} \frac{\partial \hat{f}(t,x)}{\partial x} - q_\lambda(x)\hat{\lambda}(t) \\ a\hat{\lambda}(t) \end{pmatrix} dt + \begin{pmatrix} \frac{1}{2}\frac{d\tilde{q}(x)}{dx} \\ b \end{pmatrix} dt$$
$$+ \left(\mathbf{P}(t)\begin{pmatrix} \mathbf{H}_\delta^* \\ -\mathbf{H}^*(q_\lambda) \end{pmatrix} + \begin{pmatrix} \mathbf{H}^*(q) \\ (\sigma_\lambda, \mathbf{H}^*(q)) \end{pmatrix}\right)\Phi^{-1} d\boldsymbol{\ell}(t), \quad (52)$$

where the innovation process $\boldsymbol{\ell}(t) = [\ell(t)\ \tilde{\ell}(t)]^*$ is defined by

$$\begin{bmatrix} \ell(t) \\ \tilde{\ell}(t) \end{bmatrix} = \begin{bmatrix} Y(t) - Y(0) - \int_0^t (H_\delta \hat{f} - \hat{\lambda}(s) H q_\lambda + \frac{1}{2} F) ds \\ \tilde{Y}(t) - \tilde{Y}(0) - \int_0^t \{\frac{1}{2}\bar{q}(s) + \bar{q}_2(s) - \hat{\lambda}(s)\bar{q}_\lambda(s)\} ds \end{bmatrix}, \quad (53)$$

and $\bar{q}(s)$, $\bar{q}_2(s)$ and $\bar{q}_\lambda(s)$ are defined by (21), (26) and (27).

The exact form of $\mathbf{P}(t)$ will be listed in Appendix 4. Note that the innovation process $\boldsymbol{\ell}(t) = [\ell(t)\ \tilde{\ell}(t)]^*$ is a $\mathcal{Y}_t$-Brownian motion with incremental covariance $\Phi(\cdot)$. Hence there exists an $\mathbf{R^{m+1}}$-valued $\mathcal{Y}_t$-standard Brownian motion $B(\cdot)$ such that

$$d\boldsymbol{\ell}(t) = \Sigma(t) dB(t),\ t \geq 0, \quad (54)$$

where $\Sigma(t)\Sigma(t)^* = \Phi(t)$, $t \geq 0$. Let $\mathbf{e}_{m+1} = (0, \ldots, 0, 1)$. Then

$$d\tilde{\ell}(t) = \mathbf{e}_{m+1}\Sigma(t) dB(t),\ t \geq 0. \quad (55)$$

5.2 Some Remarks of Arbitrage Opportunities

Before deriving an optimal portfolio, we discuss about possible arbitrage opportunity in our model. The model (1)–(2) is an arbitrage free model under the risk neutral measure $\tilde{\mathcal{P}}$ with the whole information $\tilde{\mathcal{F}}_t = \sigma\{f(s,x); 0 \leq s \leq t\}$, i.e., $\tilde{P}(t,T)$ given by (12) becomes an $\tilde{\mathcal{F}}_t$-martingale. It may be possible to find $\lambda(t)$ from $\mathcal{F}_t$, because all randomness of $\lambda(t)$ come from the random sources of $f(t,x)$. In practice we can not observe the whole process $f(t,x)$. Some finite number of bonds and yield curves are observed, i.e., our obtained data structure is only a partial observation,

$$\mathcal{Y}_t(\text{our observation data}) \subset \tilde{\mathcal{F}}_t.$$

Hence we construct a Kalman filter to estimate λ. Now by using this estimate $\hat{\lambda}(t) = E\{\lambda(t)|\mathcal{Y}_t\}$ and with the aid of Proposition-2, we get

$$df(t,x) = \frac{\partial f(t,x)}{\partial x}dt + \frac{1}{2}\frac{d}{dx}\tilde{q}(x)dt - (\lambda(t) - \hat{\lambda}(t))dt + d\tilde{w}(t,x).$$

This equation does not satisfies the Musiela drift condition [16]. Hence this implies that under the partial observation $\mathcal{Y}_t$ we may have some arbitrage opportunities which can not be predicted from the data $\mathcal{Y}_t$, if $(\lambda(t) - \hat{\lambda}(t))$ is not negligible. This is the main reason to introduce the mead-variance hedging, instead of the usual hedging obtained from the standard pricing formula. Furthermore under the world of the partial observation data $\mathcal{Y}_t$, the dynamics of the forward rates is the Kalman filter:

$$d\begin{pmatrix} \hat{f}(t,x) \\ \hat{\lambda}(t) \end{pmatrix} = \begin{pmatrix} \frac{\partial \hat{f}(t,x)}{\partial x} - q_\lambda(x)\hat{\lambda}(t) \\ a\hat{\lambda}(t) \end{pmatrix} dt + \begin{pmatrix} \frac{1}{2}\frac{d\tilde{q}(x)}{dx} \\ b \end{pmatrix} dt + \mathbf{K}(t,x)\Phi^{-1}d\boldsymbol{\ell}(t)$$

where the Kalman gain $\mathbf{K}(t,x) = [K_1(t,x)\ K_2(t,x)]'$. In this case, using Proposition-2[3] again , and setting $\int_0^{T-t} K_1(t,x)dx\Phi^{-1}d\tilde{\boldsymbol{\ell}}(t) = \int_0^{T-t} K_1(t,x)dx\Phi^{-1}d\boldsymbol{\ell}(t) + \int_0^{T-t}[\frac{1}{2}(\frac{d\tilde{\mathbf{K}}(t,x)}{dx} - \frac{d\tilde{q}(x)}{dx}) + q_\lambda(x)\hat{\lambda}(t)]dxdt$, we derive

$$d\hat{f}(t,x) = \frac{\partial \hat{f}(t,x)}{\partial x}dt + \frac{1}{2}\frac{d}{dx}\tilde{\mathbf{K}}(t,x)dt + K_1(t)\Phi^{-1}d\tilde{\boldsymbol{\ell}}(t), \tag{56}$$

where $\tilde{\mathbf{K}}(t,x) = \int_0^x K_1(t,z)dz\Phi^{-1}(\int_0^x K_1(t,z)dz)'$. This equation satisfies the Musiela drift condition, i.e., $\hat{P}(\tilde{t},T) = \frac{\exp\{-\int_t^{T-t}\hat{f}(t,x)dx\}}{\exp\{\int_0^t \hat{f}(s,0)ds\}}$ is a $\mathcal{Y}_t$-martingale. Hence under the partial observation $\mathcal{Y}_t$, the Kalman filter (56) is an arbitrage free model. For the mean-variance hedging problem, we only use the output $\hat{\lambda}(t)$ as mentioned in Sect. 5.

5.3 Optimal Portfolio

We use arguments analogous to that of Sect. 4, to characterize optimal portfolio. Since, our observation is the enlarged filtration $\boldsymbol{\mathcal{Y}}_t$, we reset the admissible set for θ as

$$U_{ad} = \left\{\theta;\theta(t) \text{ is } \boldsymbol{\mathcal{Y}}_t - \text{measurable with } E\{\int_0^{T_M}\theta(t)P^2(t,T_M)ds\} < \infty\right\}.$$

[3] We assume that $\frac{|\tilde{q}(x)|+|\int_0^x q_\lambda(y)dy|}{\tilde{\mathbf{K}}(t,x)} \leq C$(independent of t and $x > 0$) .

Noting that the yields $y_1(t), y_2(t), \cdots, y_m(t)$ are not tradable assets, we choose the same form of the self-financing portfolio as given in (29) for the augmented data $\mathcal{Y}_t$. Hence we need to calculate the following indifference price:

$$\hat{H}^t_{T_m} = E\{(P(T_m, T_M) - K)^+|\mathcal{Y}_t\}. \tag{57}$$

Now from (52), we get

$$d\hat{\lambda}(t) = (a\hat{\lambda}(t) + b)dt + K_\lambda d\boldsymbol{\ell}(t),$$

where $\boldsymbol{\ell}(t) = [\ell(t)\ \tilde{\ell}(t)]^*$ and the $1 \times (m+1)$-dimensional gain K_λ satisfies

$$\begin{bmatrix} K_f \\ K_\lambda \end{bmatrix} = \left(\mathbf{P}(t)\begin{pmatrix} \mathbf{H}^*_\delta \\ -\mathbf{H}(q_\lambda) \end{pmatrix} + \begin{pmatrix} q\mathbf{H}^* \\ (\sigma_\lambda, q\mathbf{H}^*) \end{pmatrix}\right)\Phi^{-1}.$$

As in Sect. 4, using Ito's formula, we get

$$\begin{aligned} d\hat{H}^t_{T_m} &= \frac{\partial \hat{H}^t_{T_m}}{\partial m_1}\frac{\partial m_1}{\partial \tilde{\ell}} d\tilde{\ell}(t) + \frac{\partial \hat{H}^t_{T_m}}{\partial m_2}\frac{\partial m_2}{\partial \hat{\lambda}} K_\lambda(t) d\boldsymbol{\ell}(t) \\ &= (-\mathbf{e}_{m+1} + A(t)K_\lambda(t))\hat{H}^t_{T_m} d\boldsymbol{\ell}(t). \end{aligned} \tag{58}$$

Hence the hedging error ϵ_t becomes

$$\begin{aligned} d\epsilon_t &= ((\mathbf{e}_{m+1} - A(t)K_\lambda(t))\hat{H}^t_{T_m} - \theta(t)e^{-\tilde{Y}(t)}\mathbf{e}_{m+1})d\boldsymbol{\ell}(t) \\ &\quad + \{\bar{q}_2(t) + \hat{\lambda}(t)\bar{q}_\lambda(t)\}\theta(t)e^{-\tilde{Y}(t)}dt. \end{aligned} \tag{59}$$

From Ito's lemma, we obtain

$$\begin{aligned} J(t, x_0, \theta) = E\{|x_0 - \hat{H}^0_{T_m}|^2\} \\ &+ 2E\{\int_0^t \{\bar{q}_2(s) + \hat{\lambda}(s)\bar{q}_\lambda(s)\}\theta(s)e^{-\tilde{Y}(s)}(\tilde{V}_s - \hat{H}^s_{T_m})ds\} \\ &+ E\{\int_0^t (\hat{H}^s_{T_m}(\mathbf{e}_{m+1} - A(s)K_\lambda(s)) - \theta(s)e^{-\tilde{Y}(s)}\mathbf{e}_{m+1})\Phi \\ &\times (\hat{H}^s_{T_m}(\mathbf{e}_{m+1} - A(s)K_\lambda(s)) - \theta(s)e^{-\tilde{Y}(s)}\mathbf{e}_{m+1})^*\}ds. \end{aligned} \tag{60}$$

It is also possible to get

$$\hat{H}^t_{T_m} = \exp\left(m_1(t, \tilde{\ell}(t)) + m_2(t, \hat{\lambda}(s); 0 \le s \le t)) + \frac{R(t)}{2}\right) N(d_1(t)) - KN(d_2(t)), \tag{61}$$

where m_1, m_2, d_1 and d_2 are same forms in Sect. 4 and the covariance R is reset as

$$R(t) = \int_t^{T_m} \bar{q}(s)ds - 2\int_t^{T_m} A(s)K_\lambda(s)\Phi \mathbf{e}^*_{m+1}ds + \int_t^{T_m} A^2(s)K_\lambda(s)\mathbf{\Phi} K^*_\lambda(s)ds.$$

Therefore, $(x_0^o, \theta^o) = \text{argmin}_{x_0 \in R^1, \theta \in U_{ad}} J(x_0, \theta)$ satisfies

$$\begin{cases} \theta^o(t) = \frac{e^{\tilde{Y}(t)}}{\mathbf{e}_{m+1}\Phi \mathbf{e}^*_{m+1}} \Big\{(\mathbf{e}_{m+1} - A(t)K_\lambda(t))\Phi \mathbf{e}^*_{m+1} \hat{H}^t_{Tm} \\ \qquad\qquad -(\bar{q}_2(t) + \hat{\lambda}(t)\bar{q}_\lambda(t))(\tilde{V}_t - \hat{H}^t_{T_m})\Big\} \\ x_0^o \quad = \hat{H}^0_{T_m} \end{cases}$$

Thus we have the following theorem.

Theorem 3 *The optimal mean-variance hedge* $(x_0^o, \theta^o(\cdot))$ *exists and is given by*

$$x_0^o = \frac{P(0, T_M)}{P(0, T_m)} \exp\left(A(0)\hat{\lambda}(0) + b\int_0^{T_m} A(s)ds - \frac{1}{2}\int_0^{T_m} \bar{q}(s)ds \right.$$
$$\left. - \int_0^{T_m} \bar{q}_2(s)ds + \frac{R(t)}{2}\} N(d_1(t)) + \frac{R(0)}{2}\right) N(d_1(0)) - KN(d_2(0)) \tag{62}$$

and

$$\theta^o(t) = \left\{\exp\left(A(t)\hat{\lambda}(t) + b\int_t^{T_m} A(s)ds - \frac{1}{2}\int_t^{T_m} \bar{q}(s)ds - \int_t^{T_m} \bar{q}_2(s)ds + \frac{R(t)}{2}\right)\right.$$
$$\left. \times N(d_1(t)) - \frac{P(t, T_m)}{P(t, T_M)} KN(d_2(t))\right\}$$
$$\times \left(1 - A(t)K_\lambda(t)\Phi \mathbf{e}^*_{m+1}\frac{1}{\bar{q}(t)} + \frac{1}{\bar{q}(s)}\{\bar{q}_2(t) + \hat{\lambda}(t)\bar{q}_\lambda(t)\}\right)$$
$$-\frac{1}{\bar{q}(t)}\left\{\bar{q}_2(t) + \hat{\lambda}(t)\bar{q}_\lambda(t)\right\} \frac{V_t}{P(t, T_M)}. \tag{63}$$

Table 1 System parameters

c	a_r	$\sigma_{q\lambda}$	σ_r	σ_ℓ	a	b
0.1627	3.3114	−36	0.2949	0.15	−2	0.1

Table 2 Yield and bond parameters

τ_1	τ_2	τ_3	τ_4	τ_5	τ_6	τ_7	T_m	T_M
1	2	3	5	7	10	20	0.5	0.75

6 Simulation Studies

In this digital simulation study, from [2] we set [4]

$$q(x_1, x_2) = \sigma^2 \sum_{i=1}^{20} \frac{1}{i^2} \exp(-cx_1) \sin\left(\frac{\pi i x_1}{30}\right) \exp(-cx_2) \sin\left(\frac{\pi i x_2}{30}\right) + \sigma_r^2 \exp(-a_r(x_1 + x_2))$$

$$q_\lambda(x) = \sigma_{q\lambda} \exp(-0.1x^2) \int_0^x q(x, y)dy$$

and

$$\sigma_\lambda(\cdot) = \sigma_\ell \int_{\tilde{G}} (\cdot)dx.$$

The system parameters are given in Table 1 where c, a_r, σ_r and q are set from the experimental results for US-bond data in [2]. Other parameters are artificially set.

To simulate the yield curve and bond data, we used the parameters for the yield and bond data as shown in Table 2.

Now we generated 100 samples of f, the yield and bond data. One of the simulated f and the yield curve $[y_1, \cdots, y_7]$ and $P(:, T_M)/P(:, T_m)$ are shown in Figs. 1 and 2, respectively.

First we show the results for estimating the stochastically- varying risk premia by using $\tilde{Y}$ and $\mathbf{Y}$ in Fig. 3.

We also present the estimate of $f(t, x)$ in Fig. 4.

As we expected, the estimated value of $\lambda(t)$ by using $\mathbf{Y}$ is much better than the estimated value by using $\tilde{Y}$. Now we shall present the hedging result for $K = 0.72$ in Fig. 5.

Finally we demonstrate the histogram of the hedging results for 100 sample paths in Figs. 6 and 7 for the yield and bond case, respectively.

[4] In this digital simulation study, $e_i(x)$ is only summed up to 20 terms. Hence (A-5) is also satisfied.

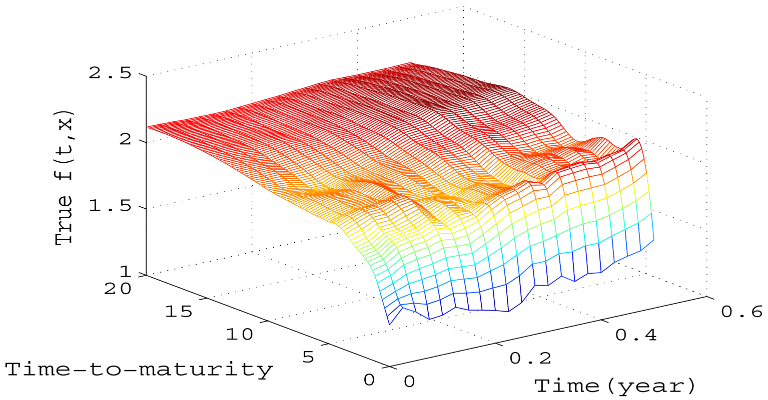

Fig. 1 True state $f(t, x)$

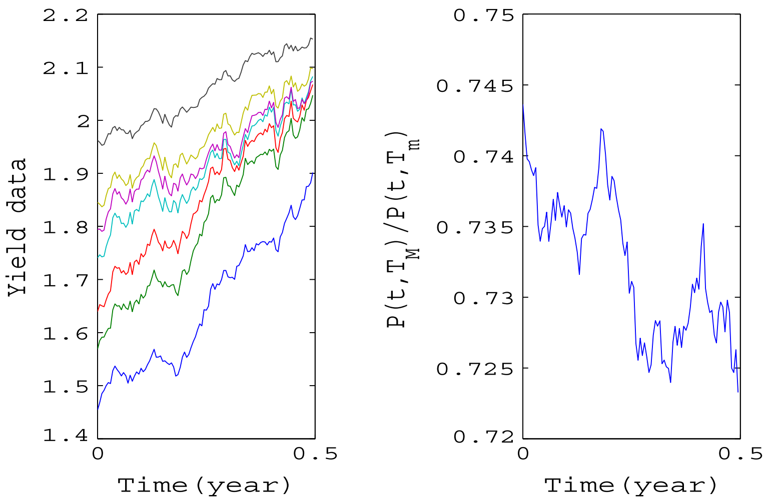

Fig. 2 Yield curve and $\frac{P(\cdot,T_M)}{P(\cdot,T_m)}$

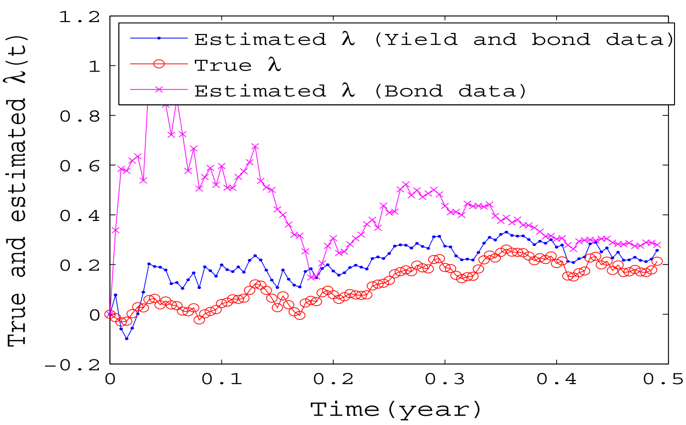

Fig. 3 Estimation of $\lambda(t)$

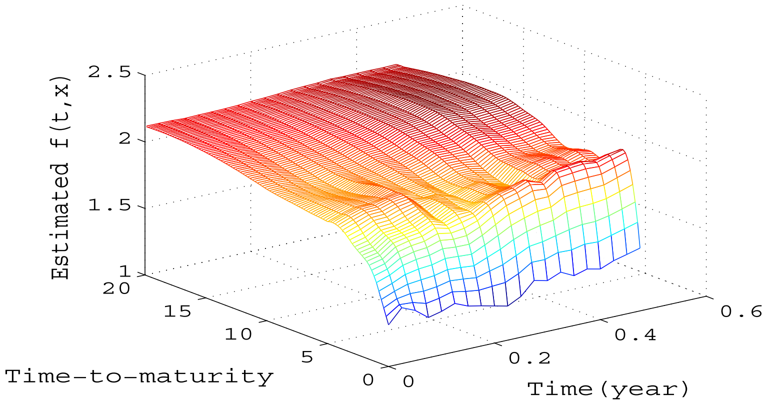

Fig. 4 Estimate of $f(t, x)$

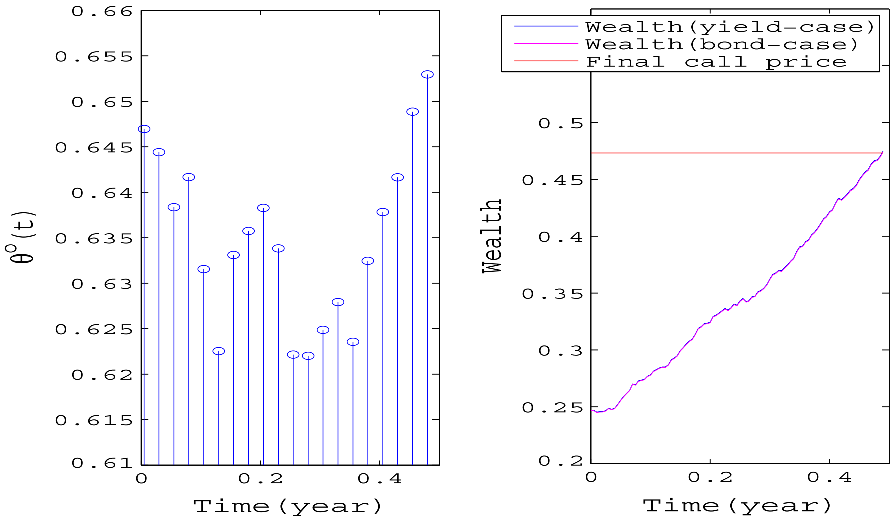

Fig. 5 Optimal $\theta^o(t)$ and its wealth process

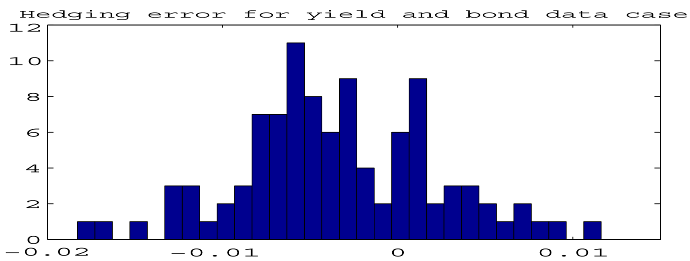

Fig. 6 Hedging error by using $[y_1, \cdots, y_7, \tilde{Y}]$

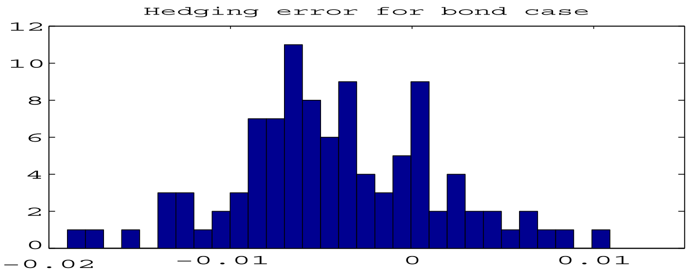

Fig. 7 Hedging error by using $\tilde{Y}$ only

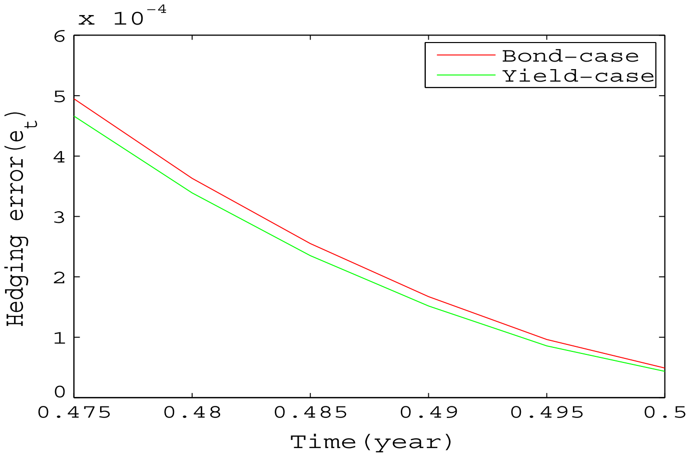

Fig. 8 Time evolution of the optimal cost

To show the feasibility of the proposed method in Sect. 5, we show the time evolution for the hedging error e_t

$$e_t = \frac{1}{100} \sum_{\text{sample}=1}^{100} (V_t(x_0^o, \theta^o : \text{sample}) - (P(t_f, T_M; \text{sample}) - K)^+)^2$$

for both yield and bond data cases in Fig. 8. The difference of the cost performance for these cases is not so big. Although in this simulation we only consider the scaler stochastic premium, the multi dimensional case will be really improved for the cost performance by using the yield curve data as the observations.

Appendix 1: Proof of Proposition 1

The Hilbert space valued stochastic integral has been well defined. See [18] for details. Here we need to check the integrability of stochastic integral $\int_0^t (\sigma_\lambda(\cdot), dw(s, \cdot)))$. From (2), we have

$$\lambda(t) = e^{at}\lambda_o + \int_0^t e^{a(t-s)} b ds + \int_0^t e^{a(t-s)} (\sigma_\lambda, dw(s)). \tag{64}$$

It follows from (A-2) and (A-4) that

$$(\sigma_\lambda, Q\sigma_\lambda) = \int_{\tilde{G}} \sigma_\lambda(x) \int_{\tilde{G}} q(x, y)\sigma_\lambda(y) dy dx$$
$$\leq |\sigma_\lambda|_{L^2(\tilde{G})} Tr(Q) < \infty.$$

Hence the 3rd term of (64) (stochastic integral) can be well defined, and (7) is derived.

The mild form of (1) becomes

$$f(t, x) = f_o(x + t) + \int_0^t (\frac{1}{2}\frac{d}{dx}\tilde{q}(x + t - s) - \lambda(s)q_\lambda(x + t - s))ds$$
$$+ \sum_{i=1}^{\infty} \int_0^t e_i(x + t - s)\sqrt{\lambda_i} d\beta_i(s), \tag{65}$$

where we use the following representation from [18]:

$$w(t, x) = \sum_{i=1}^{\infty} \sqrt{\lambda_i} e_i(x)\beta_i(t),$$

and where $\{\beta_i(t)\}$ is an R^1-valued standard Brownian motion process and $\{e_i(x)\}_{i=1}^{\infty}$ is an orthonormal basis in $L^2(\tilde{G})$ with values in H^1. It is easy to show that

$$\sum_{i=1}^{\infty} \int_0^{t_f} \lambda_i \int_0^{\hat{T}} e_i^2(x + t_f - s) dx ds = \sum_{i=1}^{\infty} \int_0^{t_f} \lambda_i \int_{t_f - s}^{\hat{T} + t_f - s} e_i^2(y) dy ds$$
$$\leq t_f \sum_{i=1}^{\infty} \lambda_i \int_0^{\hat{T} + t_f} e_i^2(y) dy = t_f \int_{\tilde{G}} q(x, x) dx = t_f Tr\{Q\} < \infty.$$

Hence the 3rd term of (65) is also well defined in $L^2(G)$. From (A-2) and (A-4), we find that $\tilde{q} \in H^1(\tilde{G})$ and $q_\lambda \in H^1(\tilde{G})$. Noting that $\lambda \in L^2(\Omega; C([0, t_f]; R^1))$, we get

$$E\{\sup_{0 \leq t \leq t_f} \int_0^{\hat{T}} f^2(t, x) dx\} \leq \text{Const.}$$

This implies that $f \in L^2(\Omega; C([0,t_f]; L^2(G)))$. To show $\frac{\partial f}{\partial x} \in L^2(\Omega; C([0,t_f]; L^2(G)))$, we repeat the above procedure. It follows from (A-2) that

$$E\left\{\sup_{0\le t\le t_f}\int_0^{\hat{T}}|\frac{\partial}{\partial x}\int_0^t\sum_{i=1}^{\infty}e_i(x+t-s)\sqrt{\lambda_i}d\beta_i(s)|^2dx\right\}$$
$$\le \int_0^{\hat{T}}(\int_0^{t_f}\sum_{i=1}^{\infty}(\frac{\partial e_i(x+t_f-s)}{\partial x}\sqrt{\lambda_i})^2ds)dx$$
$$\le t_f\int_{\tilde{G}}\frac{\partial^2 q(x,y)}{\partial x\partial y}_{|y=x}dx<\infty.$$

Consequently, from (A-1),(A-2) and (A-4) we find that $\frac{\partial f}{\partial x} \in L^2(\Omega; C([0,t_f]; L^2(G)))$, i.e., $f \in L^2(\Omega; C([0,t_f]; H^1(]0,\hat{T}[)))$. From Sobolev's imbedding theorem, it follows that $f \in L^2(\Omega; C([0,t_f]; C([0,\hat{T}])))$. From this $r(t) = f(t,0)$ can be defined.

Appendix 2: Proof of Proposition 2.2

From (A-5), we can prove this proposition by repeating the well known technique in Bensoussan [4]. Set

$$M(t) = \exp\left(\int_0^t \lambda(s)(q_\lambda, Q^{-1}dw(s,\cdot)\right) - \frac{1}{2}\int_0^t \lambda^2(s)(q_\lambda, Q^{-1}q_\lambda)ds).$$

Noting that from Proposition 2.1., $M(t)$ is a local martingale with respect to $\mathcal{P}$, we have

$$E\{M(t)\} \le 1.$$

The rest of this proof is to show that $E\{M(t)\} = 1$. For details see the proof of Lemma 4.1.4 in [4] p. 77.

Appendix 3: Proof of Proposition 4

Noting that A is a solution of

$$\frac{dA(s)}{ds} + aA(s) + \bar{q}_\lambda(s) = 0,\ A(T_m) = 0 \tag{66}$$

and $\hat{\lambda}$ is also a solution of

$$d\hat{\lambda}(t) = (a\hat{\lambda}(t) + b)dt + K_{\lambda}(t)d\tilde{\ell}(t), \tag{67}$$

we obtain

$$d(\hat{\lambda}(t)A(t)) = \frac{dA(t)}{dt}\hat{\lambda}(t)dt + A(t)(a\hat{\lambda}(t) + b)dt + A(t)K_{\lambda}(t)d\tilde{\ell}(t).$$

Hence

$$-\hat{\lambda}(t)A(t) = \int_t^{T_m} (\frac{dA(t)}{ds} + aA(s))\hat{\lambda}(s)ds + \int_t^{T_m} A(s)dsb + \int_t^{T_m} A(s)K_{\lambda}(s)d\tilde{\ell}(s). \tag{68}$$

Substituting (66) into (68), we obtain (32).

Appendix 4: The Kernels of P(t) Equation

The kernel forms are listed in the partial differential equation form;

$$\begin{aligned}
\frac{\partial p_{ff}(t,x,y)}{\partial t} &= \frac{\partial p_{ff}(t,x,y)}{\partial x} + \frac{\partial p_{ff}(t,x,y)}{\partial y} - q_{\lambda}(x)p_{\lambda f}(t,y) - p_{f\lambda}(t,x)q_{\lambda}(y) \\
&\quad - \left[\frac{p_{ff}(t,x,\tau_i) - p_{ff}(t,x,0)}{\tau_i} - p_{f\lambda}(t,x)\frac{1}{\tau_i}\int_0^{\tau_i} q_{\lambda}(z)dz + \frac{1}{\tau_i}\int_0^{\tau_i} q(x,z)dz\right]_{1\times m} \\
&\quad \times\Phi^{-1}\left[\frac{p_{ff}(t,\tau_j,y) - p_{ff}(t,0,y)}{\tau_j} - p_{f\lambda}(t,y)\frac{1}{\tau_j}\int_0^{\tau_j} q_{\lambda}(z)dz \right.\\
&\quad \left. + \frac{1}{\tau_j}\int_0^{\tau_j} q(x,z)dz\right]_{m\times 1} + q(x,y).
\end{aligned}$$

$$\begin{aligned}
\frac{\partial p_{f\lambda}(t,x)}{\partial t} &= \frac{\partial p_{f\lambda}(t,x)}{\partial x} - q_{\lambda}(x)p_{\lambda\lambda}(t) + ap_{f\lambda}(t,x) \\
&\quad - \left[\frac{p_{ff}(t,x,\tau_i) - p_{ff}(t,x,0)}{\tau_i} - p_{f\lambda}(t,x)\frac{1}{\tau_i}\int_0^{\tau_i} q_{\lambda}(z)dz + \frac{1}{\tau_i}\int_0^{\tau_i} q(x,z)dz\right]_{1\times m} \Phi^{-1} \\
&\quad \times\left[\frac{p_{f\lambda}(t,\tau_j) - p_{f\lambda}(t,0)}{\tau_j} - p_{\lambda\lambda}(t)\frac{1}{\tau_j}\int_0^{\tau_j} q_{\lambda}(z)dz \right.\\
&\quad \left. + \frac{1}{\tau_j}\int_0^{\tau_j}\int_G \sigma_{\lambda}(z)q(z,y)dzdy\right]_{m\times 1} + \int_G \sigma_{\lambda}(y)q(y,x)dy.
\end{aligned}$$

$$\frac{dp_{\lambda\lambda}(t)}{dt} = 2ap_{\lambda\lambda}(t)$$
$$- \left[\frac{p_{f\lambda}(t,\tau_i) - p_{f\lambda}(t,0)}{\tau_i} - p_{\lambda\lambda}(t)\frac{1}{\tau_i}\int_0^{\tau_j} q_\lambda(z)dz \right.$$
$$\left. + \frac{1}{\tau_i}\int_0^{\tau_i}\int_G \sigma_\lambda(z)q(z,y)dzdy \right]_{1\times m} \Phi^{-1} \times \left[\frac{p_{f\lambda}(t,\tau_j) - p_{f\lambda}(t,0)}{\tau_j} \right.$$
$$\left. - p_{\lambda\lambda}(t)\frac{1}{\tau_j}\int_0^{\tau_j} q_\lambda(z)dz + \frac{1}{\tau_j}\int_0^{\tau_j}\int_G \sigma_\lambda(z)q(z,y)dzdy \right]_{m\times 1}$$
$$+ \int_G\int_G \sigma_\lambda(x)q(x,y)\sigma_\lambda(y)dxdy.$$

References

1. Aihara, S., Bagchi, A.: Stochastic hyperbolic dynamics for infinite-dimensional forward rates and option pricing. Math. Financ. **15**(1), 27–47 (2005)
2. Aihara, S., Bagchi, A.: Identification of affine term structures from yield curve data. IJTAF **13**, 27–47 (2010)
3. Barski, M., Jakubowski, J., Zabczyk, J.: On incompleteness of bond markets with infinite number of random factors. Math. Financ. **21**, 541–556 (2010). doi:10.1111/j.1467-9965.2010.00,438.
4. Bensoussan, A.: Stochastic Control of Partially Observable Systems. Cambridge University Press, Cambridge (1992)
5. Campbell, J., Shiller, R.: Yield spreads and interest rate movements: A bird's eye view. Rev. Econ. Stud. **58**(3), 495–514 (1991)
6. Cochrane, J., Piazzesi, M.: Bond risk premia. Am. Econ. Rev. **95**(1), 138–160 (2005)
7. Collin-Dufresne, P., Goldstein, R.S.: Generalizing the affine framework to HJM and random field models. Working paper. http://sfi.epfl.ch/cms/site/sfi/lang/en/CollinDufresne/WorkingPapers (2002)
8. Cont, R., Tankov, P.: Financial Modelling with Jump Processes. Chapman & Hall, London (2004)
9. Carmona, R.: Indifference Pricing: Theory and Applications. Princeton University Press, Princeton (2009)
10. Duffee, G.: Term premia and interest rate forecasts in affine models. J. Financ. **57**(1), 405–443 (2002)
11. Fama, E., Bliss, R.: The information in long-maturity forward rates. Am. Econ. Rev. **77**(4), 680–692 (1987)
12. Heath, D., Platen, E.: A comparison of two quadratic approaches to hedging in incomplete markets. Math. Financ. **11**, 385–413 (2001)
13. Laurent, J.P., Pham, H.: Dynamic programming and mean-variance hedging. Financ. Stochast. **3**, 83–110 (1999)
14. Liptser, R., Shiryaev, A.: Statistics of Randon Processes I. Springer, Berlin (1997)
15. Mania, M., Tevzadze, R., Tronjadze, T.: Mean-variance hedging under partial information. SIAM J. Control Optim. **47**, 2381–2409 (2008)
16. Musiela, M., Rutkowski, M.: Martingale Methods in Financial Modeling. Springer, New York (1997)
17. Pham, H.: Mean-variance hedging for partially observed drift processes. IJTAF **4**, 263–284 (2001)
18. Prato, G.D., Zabczyk, J.: Stochastic Equations in Infinite Dimensions. Cambridge University Press, Cambridge (1992)
19. Taflin, E.: Bond market completeness and attainable contingent claims. Financ. Stoch. **9**, 429–452 (2005)