Data Informed Solution Estimation for Forward Backward Stochastic Differential Equations

Feng Bao, * Yanzhao Cao, †Jiongmin Yong ‡

Abstract

Forward backward stochastic differential equation (FBSDE) systems were introduced as a probabilistic description for parabolic type partial differential equations. Although the probabilistic behavior of the FBSDE system makes it a natural mathematical model in many applications, the stochastic integrals contained in the system generate uncertainties in the solutions which makes the solution estimation a challenging task. In this paper, we assume that we could receive partial noisy observations on the solutions and introduce an optimal filtering method to make data informed solution estimation for FBSDEs.

1 Introduction

In this work, we study a type of data informed solution estimation method for forward backward stochastic differential equation (FBSDE) systems. The FBSDE system is composed by two stochastic differential equations (SDEs) – a forward SDE and a backward SDE (BSDE). It was first introduced as a probabilistic interpretation for a class of semi-linear parabolic type partial differential equations (PDEs), which have been extensively studied by scientists and engineers with massive applications in industry [19]. Since FBSDEs describe the exactly same physics modeled by parabolic PDEs, they found applications in all the areas that their PDE counterparts are applicable. On the other hand, as a pure SDE system, the FBSDE system can be solved through stochastic algorithms which are typically scalable. This would allow numerical methods for solving FBSDEs to take the advantage of large scale parallel processing techniques which are the main components of modern supercomputing.

The theoretical foundation of FBSDEs is the Feynman-Kac formula, which is the bridge that links classic SDEs with parabolic PDEs through conditional expectations, and the solution of PDEs gives some statistical description for the random samples following the dynamics of SDEs. In an FBSDE system, the forward SDE coincides with a classic SDE and the BSDE plays the role of the parabolic PDE such that the BSDE provides the statistical behavior for the forward SDE random samples. Therefore, the forward SDE describes some stochastic dynamical system and the BSDE gives how the dynamics are incorporated into the physical model, and thus one FBSDE system contains both SDE side and PDE side of the Feynman-Kac formula. In this way, FBSDEs could explain stochastic dynamical systems in a more natural way, which would explore

*Department of Mathematics, Florida State University, Tallahassee, Florida, (bao@math.fsu.edu).
†Department of Mathematics and Statistics, Auburn University, Auburn, Alabama.
‡Department of Mathematics and Statistics, University of Central Florida, Orlando, Florida.

even broader applications [2, 15, 18]. There are many efforts on solving FBSDEs numerically such like [13, 23, 27, 28] and many more. All of these efforts provide effective methods to obtain the numerical solution of BSDEs as a function of the random variable that is governed by the forward SDE, which successfully capture the structural connection between the forward SDE and the BSDE in the FBSDE system. To the best of our knowledge, most existing numerical methods for solving FBSDEs target on approximating the function relation between the forward SDE and the BSDE without considering the randomness contained in the forward SDE. However, as a diffusion process, the forward SDE adopts uncertainties as it propagates in time. Therefore, conventional numerical methods that just approximate solutions as functions are not sufficient in many practical applications, and its necessary to design solution estimation methods that take uncertainties in the forward SDE into consideration. With rapid development in data collection capability, we are able to gather observational data that contains information for solutions of FBSDEs. In this connection, we introduce a computational framework that applies data assimilation methods to estimate solutions of FBSDEs based on observational data.

Data assimilation is a mathematical discipline that aims to find estimations for the state of some stochastic dynamical system based on noisy partial observations of the system. The major technique that accomplishes the mission of data assimilation is the optimal filter which provides the best estimation for the target state as a conditional expectation conditioned on the observational information. In an optimal filtering problem, the stochastic dynamical system that we want to estimate is usually named the state process and the observations are described by an observation process. Except to the linear case of the problem, which can be solved analytically by the Kalman filter method, most optimal filtering methods for nonlinear problems try to approximate the conditional probability density function (pdf) for the state process, which is also called the filtering density. An important method for the nonlinear filtering problem is the extended Kalman filter which solves a linearized problem and then apply the Kalman filter method to the corresponding linear filtering problem [10]. However, due to the linearity assumption for the classic Kalman filter method, the extended Kalman filter does not provide reliable results for highly nonlinear problems, which is the case in this work. One of the most successful methods that solves the nonlinear filtering problem is the particle filter method, which is also known as the sequential Monte Carlo method [1, 5, 9, 12, 16, 17]. The central idea of the particle filter method is to approximate the desired conditional pdf by random samples, which are called "particles", and then resample particles to adjust their distribution based on observational data. The particle filter method is an effective method and could deal with the nonlinearity well. The main drawbacks of the particle filter method are low accuracy and low stability due to the Monte Carlo sampling/resampling. In this way, although the particle filter method could solve nonlinear filtering problems, it's not suitable to the task of solution estimation for FBSDEs. In this work, on the other hand, we apply a novel optimal filtering method, named the backward SDE filter, to solve the optimal filtering problem. The main theme of the backward SDE filter is to solve a forward backward doubly stochastic differential equation (FBDSDE) system whose solution is equivalent to the filtering density up to a normalization [4, 8]. The backward SDE filter enjoys the advantages of the particle filter as a stochastic computing algorithm, and it is also more accurate and more stable than the particle filter since the solution of the FBDSDE system is the analytical filtering density as desired for the optimal filtering problem.

The computational framework for the solution estimation for FBSDE systems is composed by two components: First of all, we solve the target FBSDE system numerically and obtain approximate solutions without using observational data. Secondly, we derive an optimal filtering problem that aims to estimate FBSDE solutions and then solve it with the backward SDE filter to get data informed solution estimations for the FBSDE system. The rest of this paper is organized as follows. In Section 2, we give a brief statement for the solutions estimation problem that we try to solve. In Section 3, we introduce the backward SDE filter approach to achieve the goal of this paper. Section 4 discusses numerical implementation of the computational framework and Section 5 shows numerical experiments to demonstrate the performance of our algorithm.

2 Problem statement

In this section, we introduce the general formulation of the problem that we aim to solve. We start our discussion by introducing the forward backward stochastic differential equation (FBSDE) system in Section 2.1. Our primary goal in this work is to estimate solutions of FBSDEs and filter out the uncertainties contained in the solutions caused by stochastic integrals. Then, in Section 2.2 we propose to apply the optimal filtering method to make the best estimation for the solutions based on observational data.

2.1 Coupled forward backward stochastic differential equations

We consider the following FBSDE system defined on the probability space $(\Omega, \mathcal{F}, P)$

$$\begin{aligned} dX_t &= b(X_t, y_t, z_t)dt + \sigma(X_t, y_t, z_t)dW_t, \qquad & X_0 &= \xi \\ dy_t &= -f(X_t, y_t, z_t)dt + z_t dW_t, \qquad & y_T &= \Psi(X_T), \end{aligned} \tag{2.1}$$

where

$$\begin{gathered} b : [0,T] \times \mathbb{R}^d \times \mathbb{R}^m \times \mathbb{R}^{m\times l} \to \mathbb{R}^d, \\ \sigma : [0,T] \times \mathbb{R}^d \times \mathbb{R}^m \times \mathbb{R}^{m\times l} \to \mathbb{R}^{d\times l}, \\ f : [0,T] \times \mathbb{R}^d \times \mathbb{R}^m \times \mathbb{R}^{m\times l} \to \mathbb{R}^m \\ \Psi : \mathbb{R}^d \to \mathbb{R}^m \end{gathered}$$

are given functions and ξ is a random variable following some given probability distribution. The first equation in the system (2.1) is a stochastic differential equation (SDE) propagating from time 0 to t, where W_t is a standard Brownian motion in $\mathbb{R}^d$ and $\int_0^t \cdot dW_t$ is an Itô type stochastic integral; the second equation is a so-called backward stochastic differential equation (BSDE), which propagates from a future time T (with a given side condition $y_T = \Psi$) to the current time t. To distinguish the first equation in (2.1) from the BSDE, we sometimes call it the *forward SDE*. Denote $\mathcal{F}_t^W := \sigma\{W_s; 0 \leq s \leq t\}$ to be the σ-algebra generated by the Brownian motion W_t that contains all the information generated by X_t. It can be shown [22] that under proper assumptions for functions b, σ, f and Ψ, there exists a unique solution $\mathcal{V}_t$ adapted to $\mathcal{F}_t^W$ for the above FBSDE system, where $\mathcal{V}_t := (X_t, y_t, z_t)$ is a triple and z_t is the martingale representation for y_t. In addition, y_t and z_t are functions of X_t, i.e. $y_t = y_t(X_t)$, $z_t = z_t(X_t)$, and one can prove that

$$y_t = \nabla\sigma(X_t, y_t, z_t) z_t, \tag{2.2}$$

which connects the solutions y_t and z_t.

The FBSDE system has applications in many areas. One significant application of FBSDEs is in finance theory. For example, the forward SDE in (2.1) could model the dynamics of stock prices in the financial market. In this application, the drift term b is usually interpreted as the stock appreciation rates which constantly determine the change of stock values and σ is the volatility matrix which reflects uncertainties of stocks in the market. Therefore, the solution X_t describes the stock values at time t. On the other hand, a specially designed BSDE could model the option price for the stocks given a specific exercise price at a future time T and z_t gives the investment portfolio which is the combination of stocks to satisfy the option [11, 18]. In the situation that some agent holds large portion of option which could influence stock prices, the functions b and σ would contain the option price y_t, and therefore the FBSDE system becomes coupled. Another remarkable application of the FBSDE system is in stochastic optimal control theory. If the forward stochastic process X_t in (2.1) is controlled by some control terms, it is usually named the "controlled process". Then, we can find a BSDE corresponding to the adjoint stochastic process of the controlled process. In this connection, the FBSDE system would result the so-called "stochastic maximum principle" for the stochastic optimal control problem [14, 24].

In most areas that FBSDEs are applicable, we want to find the solution $\mathcal{V}_t$ of (2.1). Since y_t and z_t are functions of X_t, for any given point $x\in\mathbb{R}^d$ such that $X_t=x$, the solution for (2.1) is expressed as $X_t=x$, $y_t=y_t(x)$ and $z_t=z_t(x)$. In this way, traditional numerical approaches for solving FBSDEs focus on finding approximations for functions $y_t(\cdot)$ and $z_t(\cdot)$, and leave X_t as a random variable. However, as a random variable, X_t contains uncertainties which is generated by the Brownian motion through the Itô integrals. In practice, we usually don't have the exact value of X_t as $X_t=x$, and the solution pair (y_t,z_t) of the FBSDE system (2.1) as functions of X_t are also random variables due to the randomness caused by the uncertainties in X_t. In order to estimate the solution triple (X_t,y_t,z_t), in this work we assume that we are able to collect observational data on X_t and we propose to use optimal filtering methods, which will be discussed in the following subsection, to filter out the uncertainties in the random variable X_t and make data informed estimations for solutions of FBSDEs.

2.2 Optimal filtering problem

The optimal filtering problem is usually formulated by the following state-space model on the probability space $(\Omega,\mathcal{F},P)$:

$$\begin{aligned} dS_t &= h(S_t)+\gamma dW_t, && \text{(State)} \\ dM_t &= g(S_t)+dB_t, && \text{(Observation)} \end{aligned} \tag{2.3}$$

where S_t and M_t are two stochastic processes which take values in $\mathbb{R}^d$ and $\mathbb{R}^k$ respectively; $h:\mathbb{R}^d\to\mathbb{R}^d$ and $g:\mathbb{R}^d\to\mathbb{R}^k$ are two functions; W_t and B_t are two independent Brownian motions that bring uncertainties to S_t and M_t; and we assume that S_0 follows a initial distribution p_0. The stochastic process S_t is called the "state process" that describes the state of some stochastic dynamical model, and γ in the state process is the covariance matrix for the state noise W_t. M_t is usually named by the "observation process", which provides noise perturbed measurements for S_t through the observation function $g(\cdot)$. The goal of the optimal filtering problem is to find the best estimation for $\Phi(S_t)$ based on observations, where $\Phi(\cdot)$ is a given test function. Mathematically, we want to find the conditional expectation for $\Phi(S_t)$ given the observational information $\mathcal{M}_t:=\sigma(M_s,0\le s\le t)$, i.e. $\tilde{\Phi}(S_t):=E[\Phi(S_t)|\mathcal{M}_t]$. The solution $\tilde{\Phi}(S_t)$ of the optimal filtering problem is often called the "optimal filter" of Φ and it can be expressed by the integral of $\Phi(\cdot)$ multiplied by a density function p_t, i.e.

$$\tilde{\Phi}(S_t) = E[\Phi(S_t)|\mathcal{M}_t] = <\Phi(x), p_t(x)>, \tag{2.4}$$

where the inner product $<\cdot,\cdot>$ indicates the spatial integral with respect to x over the $\mathbb{R}^d$ space and $p_t = p(X_t|\mathcal{M}_t)$ is the conditional probability density function (pdf) of the state S_t given $\mathcal{M}_t$, which is usually called the filtering density. In most practical approaches for solving the optimal filtering problem, we aim to find the filtering density p_t and use the inner product (2.4) to calculate the solution $\tilde{\Phi}(S_t)$.

With the aforementioned optimal filtering framework, in this paper we propose to gather observations M_t for the solution X_t in the FBSDE system. Actually, in most applications the forward SDE in (2.1) describes the state of some dynamical model, which is usually observable. Specifically, we assume that we have the following observation process

$$dM_t = g(X_t)dt + dB_t, \tag{2.5}$$

where $g(\cdot)$ is the observation function that collects information about X_t. Our goal in this work is to find the estimate solution $\hat{\mathcal{V}}_t := (\hat{X}_t, \hat{y}_t, \hat{z}_t)$ as

$$\hat{X}_t = E[X_t|\mathcal{M}_t], \quad \hat{y}_t = E[y_t|\mathcal{M}_t], \quad \hat{z}_t = E[z_t|\mathcal{M}_t], \tag{2.6}$$

where the conditional expectation has incorporated the information contained in observational data. To calculate $\hat{X}_t$, $\hat{y}_t$ and $\hat{z}_t$, we need two components: (i) the solution (X_t, y_t, z_t) of the FBSDE system (2.1); and (ii) the conditional pdf of X_t given observational information, which is the filtering density p_t that we need to find in the optimal filtering problem. When we get the conditional distribution of the forward SDE X_t, we could calculate $\hat{X}_t$, $\hat{y}_t$ and $\hat{z}_t$ by $\hat{X}_t = <x, p_t(x)>$, $\hat{y}_t = <y_t(x), p_t(x)>$ and $\hat{z}_t = <z_t(x), p_t(x)>$. The procedure to obtain the second component is the standard optimal filtering problem and the first component is typically achieved through numerical methods for solving FBSDEs. While both components are well studied, to the best of our knowledge, this is the first attempt to introduce a computational framework that combine these two efforts together to get effective data informed estimations for the solution of FBSDEs.

Apparently, accurate approximations for p_t in the optimal filtering problem is the key component in calculating $\hat{X}_t$, $\hat{y}_t$ and $\hat{z}_t$. Well known optimal filtering methods such like the ensemble Kalman filter and the particle filter are Monte Carlo based methods. Although they are effective optimal filtering techniques in solving many application problems, due to the fact that both Kalman filters and particle filters solve approximate problems for the original optimal filtering problem and the low accuracy of these methods typically occurs in Monte Carlo type approaches, these methods are not suitable for the task in this work as we need an accurate approximation for the density function to calculate the inner product $<\cdot,\cdot>$. In order to provide a more accurate approximation for the conditional pdf, in what follows we introduce a backward SDE filter approach to solve the optimal filtering problem and thus derive an accurate numerical method to estimate the data informed solution $\hat{\mathcal{V}}_t$ for the coupled FBSDE system.

3 Backward SDE filter for FBSDEs

The mathematical foundation of the backward SDE filter is the forward backward doubly stochastic differential equation (FBDSDE) system. In this section, we first introduce the backward SDE filter that solves the general optimal filtering problem (2.3) through the FBDSDE system. Then, we present formulation of the backward SDE filter for estimating solutions of coupled FBSDEs. To simplify our presentation, in the rest of this paper we assume that σ is an invariant constant.

The goal of the backward SDE filter is to find the conditional pdf, i.e. the filtering density p_t, for the state S_t in the optimal filtering problem. Then the optimal filter would be obtained by taking the inner product of the test function and the conditional pdf as indicated in (2.4). However, instead of getting the filtering density in the original probability space $(\Omega, \mathcal{F}, P)$, we are going to derive the conditional pdf for the target state in an induced probability space $(\Omega, \mathcal{F}, \tilde{P})$, where $\tilde{P}$ is defined by the Radon-Nikodym derivative, i.e. $\frac{d\tilde{P}}{dP} = Q_0^t$. The function Q_s^t is the Girsanov transformation function defined by

$$Q_s^t := \exp\Big(\int_s^t g^*(S_r) dM_r - \frac{1}{2}\int_s^t |g(S_r)|^2 dr\Big) \tag{3.7}$$

where g^* is the transpose of g and we denote $Q^t := Q_0^t$. From the Girsanov transformation theorem, we know that when the observation function g in (2.3) satisfies the Novikov's condition, the observation process M_t can be considered as a standard Brownian motion under the induced probability measure $\tilde{P}$ [7, 25]. It's important to observe that

$$E[\Phi(S_t)\big|\mathcal{M}_t] \propto \tilde{E}[\Phi(S_t)Q^t\big|\mathcal{M}_t]. \tag{3.8}$$

Therefore, the conditional pdf that we obtain in $(\Omega, \mathcal{F}, \tilde{P})$ is equivalent to the filtering density in the original probability space up to a normalization factor.

In connection with the optimal filtering problem (2.3), for any given time instant $t > 0$ we introduce the following FBDSDE system in the probability space $(\Omega, \mathcal{F}, \tilde{P})$,

$$\begin{aligned} dS_r &= h(S_r)dr + \gamma dW_r, & S_0 &= x, \\ dU_r &= V_r dW_r - g(S_r)U_r d\overleftarrow{M}_r, & U_t &= \Phi(S_t), \end{aligned} \tag{3.9}$$

where h, g and γ are the same coefficients in the optimal filtering problem, the side condition Φ is the test function in (2.4), W_t is the Brownian motion in the state process and M_t is the observation process which is a Brownian motion under the induced probability $\tilde{P}$. We can see that the first equation in (3.9) is a standard SDE, which is equivalent to the state process in the optimal filtering problem. The second equation is called a backward doubly stochastic differential equation (BDSDE). It contains two stochastic integrals $\int \cdot dW_r$ and $\int \cdot d\overleftarrow{M}_r$. The second integral, i.e. $\int \cdot d\overleftarrow{M}_r$, is an Itô integral integrated backward which is also named "backward Itô integral". Specifically, for a quasi-uniform temporal partition $s = t_0 < t_1 < \cdots < t_{N-1} < t_N = t$, the backward Itô integral for a stochastic process ϕ_t is defined by

$$\int_s^t \phi_r d\overleftarrow{M}_r := \lim_{\Delta \to 0} \sum_{n=0}^{N} \phi_{t_{n+1}}(M_{t_{n+1}} - M_{t_n}),$$

where $\Delta = \max\limits_{0 \le i \le N-1}(t_{n+1} - t_n)$. We can see that the BDSDE has similar structure to a BSDE except that there's an extra backward Itô integral involved in the equation. The solution of the above FBDSDE system is also a triple (S_t, U_t, V_t), where V_t is the martingale representation of U_t. It has been shown that under certain assumptions for h, g and Φ, the FBDSDE system (3.9) has a unique solution [21].

One important property of the FBDSDE system (3.9) is that the solution U satisfies the following optimal filtering type Feynman-Kac formula

$$U_0 = \tilde{E}\big[\Phi(S_t)Q^t \big| \mathcal{M}_t, S_0 = x\big], \tag{3.10}$$

where $\tilde{E}$ is the expectation under the induced probability measure $\tilde{P}$. The proof of (3.10) involves applications of Girsanov transformation, martingale representation and a type of Itô formula for two-sided Itô integrals [20], and we refer to [4] for details of the proof. Although the equivalent relation (3.10) shows that the solution U_t of the FBDSDE system (3.9) provides information about the conditional expectation of $\Phi(S_t)Q^t$ given $\mathcal{M}_t$, the extra condition $S_0 = x$ indicates that (3.10) is a conditional expectation for a given deterministic point at time $t = 0$, and therefore U_0 does not provide the desired solution of the optimal filtering problem.

In order to obtain the filtering density that leads to the solution, one can prove that the adjoint process for U_t, denoted by $\tilde{Y}_t$, is the conditional pdf for the state process. Specifically, on time interval $[0,t]$ we introduce the following stochastic differential system in the probability space $(\Omega, \mathcal{F}, \tilde{P})$

$$\begin{aligned} d\overleftarrow{X}_r &= h(\overleftarrow{X}_r)dr - \gamma d\overleftarrow{W}_r, & \overleftarrow{X}_t &= x, \\ d\tilde{Y}_r &= -\sum_{j=1}^{d} \frac{\partial h_j}{\partial x_j}(\overleftarrow{X}_r)\tilde{Y}_r dr - \tilde{Z}_r d\overleftarrow{W}_r + g^*(\overleftarrow{X}_r)\tilde{Y}_r dM_r, & \tilde{Y}_0 &= p_0(X_0), \end{aligned} \tag{3.11}$$

where $\frac{\partial h_j}{\partial x_j}$ is the partial differential operator with respect to the j-th component of $\overleftarrow{X}_t$, g^* is the transpose of the observation function g, and p_0 is the distribution of S_0. We notice that in (3.11) the dW integrals are backward Itô integrals, i.e. $\int \cdot d\overleftarrow{W}_r$, and the dM integral is now a standard forward Itô integral, i.e. $\int \cdot dM_r$. Since the initial condition for the first equation is given at time t, i.e. $\overleftarrow{X}_t = x$, it propagates backward in time from t to 0, which is consistent with the backward Itô integral. Therefore, the first equation in (3.11) is a standard SDE with inverse time index. On the other hand, the second equation contains both the standard forward Itô integral and the backward Itô integral, which follows the structure of BDSDEs except that the propagation direction is from 0 to t. In this way, the stochastic differential system (3.11) is an FBDSDE system with inverse time index, and the existence and uniqueness for the solution triple $(\overleftarrow{X}_t, \tilde{Y}_t, \tilde{Z}_t)$ in (3.11) are guaranteed.

Under some regularity assumptions, we can prove that U_t and $\tilde{Y}_t$ are adjoint stochastic processes, which means the inner product of U_t and $\tilde{Y}_t$ for any time instant t is a constant almost everywhere (see [3, 4]), i.e. $< U_0, \tilde{Y}_0 > = < U_t, \tilde{Y}_t >$, $\forall t > 0$. Then, from the side conditions $U_t = \Phi$ and $Y_0 = p_0$ introduced in (3.9) and (3.11) respectively, we know that

$$\tilde{E}\big[\Phi(S_t)Q^t \big| \mathcal{M}_t\big] = < U_0, p_0 > = < U_0, \tilde{Y}_0 > = < U_t, \tilde{Y}_t > = < \Phi, \tilde{Y}_t >,$$

which leads to

$$< \Phi, \tilde{Y}_t > = \tilde{E}\big[\Phi(S_t)Q^t \big| \mathcal{M}_t\big] \propto E\big[\Phi(S_t) \big| \mathcal{M}_t\big]. \tag{3.12}$$

Therefore, the inner product of a test function Φ and the solution $\tilde{Y}_t$ of the time inverse FBDSDE system (3.11) is proportional to the conditional expectation of Φ, which means that the solution $\tilde{Y}_t$ is the unnormalized filtering density as desired in the optimal filtering problem.

Now, we shall discuss the backward SDE filter approach to estimate the solution $\mathcal{V}_t$ of coupled FBSDEs. Recall that we have observational data in the form of some observation process M_t to collect information about X_t. The contribution of the backward SDE filter in this framework is to derive the conditional pdf of X_t given the observational information. Since the solutions y_t and z_t are both functions of X_t, with the conditional pdf of X_t and choosing y_t and z_t to be the test function Φ, the optimal filters $\tilde{y}_t$ and $\tilde{z}_t$ (as defined in (2.4)) would be our estimations for y_t and z_t.

To proceed, we first describe the optimal filtering problem based on the FBSDE system (2.1). Since the stochastic process X_t is the target state, we let the forward SDE to be the state process and use it to replace S_t in the optimal filtering problem to get the following state-space model

$$\begin{aligned} dX_t &= b(X_t, y_t, z_t) + \sigma dW_t, && \text{(State)} \\ dM_t &= g(X_t) + dB_t, && \text{(Observation)} \end{aligned} \tag{3.13}$$

where M_t is the observation process that provides noisy observations for the state process X_t. It's important to mention that the state process in the optimal filtering problem (3.13) contains solutions y_t and z_t of the coupled FBSDE system. Therefore, we need to solve for y_t and z_t in (2.1) with the side condition $y_T = \Psi$ in order to obtain the state process X_t. In addition, since y_t and z_t are functions of X_t, the dynamics b in (3.13) is a combined function i.e. $b(\cdot, y_t(\cdot), z_t(\cdot))$, and the dynamics $h(\cdot)$ in the optimal filtering problem (2.3) is now in the form of $h(x) = b(x, y_t(x), z_t(x))$. We also want to recall that although y_t and z_t are stochastic processes, the uncertainty of y_t and z_t is from X_t as functions $y_t(X_t)$ and $z_t(X_t)$. As we discussed above, the solution $\tilde{Y}_t$ of the time inverse FBDSDE system (3.11) is equivalent to the unnormalized filtering density that leads to the solution of the optimal filtering problem. It follows from the derivation of the backward SDE filter through the relation between the FBDSDE system (3.11) and the optimal filtering problem (2.3) that the FBDSDE system for the optimal filtering problem (3.13) is

$$\begin{aligned} d\overleftarrow{X}_r &= b(\overleftarrow{X}_r, \overleftarrow{y}_r, \overleftarrow{z}_r)dr - \sigma d\overleftarrow{W}_r, && \overleftarrow{X}_t = x, \\ d\tilde{Y}_r &= -\sum_{j=1}^{d} \frac{\partial b_j}{\partial x_j}(\overleftarrow{X}_r, \overleftarrow{y}_r, \overleftarrow{z}_r)\tilde{Y}_r dr - \tilde{Z}_r d\overleftarrow{W}_r + g^*(\overleftarrow{X}_r)\tilde{Y}_r dM_r, && \tilde{Y}_0 = p_0(S_0), \end{aligned} \tag{3.14}$$

where p_0 is the distribution of ξ, which is the initial condition for X_0, the back arrows on y_t and z_t indicate that $\overleftarrow{y_t}$ and $\overleftarrow{z}_t$ are both functions of $\overleftarrow{X}_t$ in (3.14). When we get the solution $\tilde{Y}_t$ in the induced probability space $(\Omega, \mathcal{F}, \tilde{P})$, in order to obtain the estimations in the original probability space $(\Omega, \mathcal{F}, P)$, we apply the Girsanov transform to $\tilde{Y}_t$ and get $Y_t := Q^t \tilde{Y}_t / C$, where C is a normalization factor. Therefore, the estimation for the solution of FBSDEs, i.e. $\hat{\mathcal{V}}_t = (\hat{X}_t, \hat{y}_t, \hat{z}_t)$ defined in (2.6), is given by

$$\hat{X}_t = < X_t, Y_t >, \quad \hat{y}_t = < y_t, Y_t >, \quad \hat{z}_t = < z_t, Y_t >. \tag{3.15}$$

As a result, the aforementioned framework includes two steps: (i) the construction step for the appropriate optimal filtering problem (3.13) based on the solutions of FBSDEs; and (ii) the backward SDE filter step to solve the corresponding filtering problem. Since in most practical problems, it's very difficult to find analytical solutions for coupled FBSDEs and optimal filtering problems are not explicitly solvable when the state-space model is a nonlinear system, in the following section we shall introduce a numerical approach to implement steps (i)-(ii).

4 Numerical approach

The numerical implementation of the backward SDE filter approach for estimating solutions of coupled FBSDEs is composed by two computational tasks: (i) numerical solution for FBSDEs and (ii) numerical solution for FBDSDEs. In this section, we shall discuss numerical schemes for solving FBSDEs and FBDSDEs separately. Then, we construct the general computational framework that combines the numerical solutions of FBSDEs and FBDSDEs to achieve the goal of solution estimation for FBSDEs. For the convenience of presentation, in this work we introduce our algorithms in the one dimensional case. The multi-dimensional cases can be obtained in a similar way.

4.1 Numerical solution for FBSDEs

We first derive numerical methods for solving the FBSDE system (2.1). To proceed, we introduce a temporal partition $\Pi_N(0,T) := \{0 = t_0 < t_1 < t_2 < \cdots < t_N = T\}$ on the time interval $[0,T]$, where N is the number of partition steps, $T > 0$ is a given positive time instant, and we denote $\Delta t_n := t_{n+1} - t_n$ to be the stepsize at the temporal partition step n and denote $\Delta W_{t_n} := W_{t_{n+1}} - W_{t_n}$. Then, for $n = 0,1,2,\ldots,N-1$, we consider the FBSDE system on the sub-interval $[t_n, t_{n+1}]$ in the integral form as following

$$\begin{aligned} X_{t_{n+1}} &= X_{t_n} + \int_{t_n}^{t_{n+1}} b(X_t,y_t,z_t)dt + \int_{t_n}^{t_{n+1}} \sigma dW_t, && (SDE) \\ y_{t_n} &= y_{t_{n+1}} + \int_{t_n}^{t_{n+1}} f(X_t,y_t,z_t)dt - \int_{t_n}^{t_{n+1}} z_t dW_t. && (BSDE) \end{aligned} \tag{4.16}$$

To derive a numerical solution for (4.16), we first discuss approximation schemes for the integrals in the system. Since we aim to describe the general methodology of applying optimal filtering methods to estimate solutions of FBSDEs, in this work we use classic numerical integration methods to approximate integrals. Specifically, for the SDE in (4.16), we apply the left point formula to approximate the deterministic integral and get

$$X_{t_{n+1}} = X_{t_n} + b(X_{t_n}, y_{t_n}, z_{t_n})\Delta t_n + \sigma \Delta W_{t_n} + R_X^n, \tag{4.17}$$

where $R_X^n := \int_{t_n}^{t_{n+1}} b(X_t,y_t,z_t)dt - b(X_{t_n},y_{t_n},z_{t_n})\Delta t_n$ is the approximation error. It's worthy to mention that since we assume that σ is an invariant constant, it's straightforward that $\sigma \Delta W_{t_n} = \int_{t_n}^{t_{n+1}} \sigma dW_t$. For the BSDE in the system, we apply the Euler-Maruyama scheme to approximate the stochastic integral, and apply the right point formula to approximate the deterministic integral. In this way, we get

$$y_{t_n} = y_{t_{n+1}} + f(X_{t_{n+1}}, y_{t_{n+1}}, z_{t_{n+1}})\Delta t_n - z_{t_n}\Delta W_{t_n} + R_Y^n, \tag{4.18}$$

where $R_Y^n := \int_{t_n}^{t_{n+1}} f(X_t,y_t,z_t)dt - f(X_{t_{n+1}},y_{t_{n+1}},z_{t_{n+1}})\Delta t_n + z_{t_n}\Delta W_{t_n} - \int_{t_n}^{t_{n+1}} z_t dW_t$ is the approximation error.

Based on the discretization formulas (4.17) and (4.18), we derive numerical schemes for the FBSDE system on the time interval $[t_n, t_{n+1}]$. For the solution X_t, it's easy to observe that by dropping the approximation error term R_X^n in (4.17), we obtain an approximation scheme for X_t directly.

For solutions y_t and z_t, we derive their approximation schemes from (4.18). Define $\mathbb{E}_X^n[\cdot] := E[\cdot | X_{t_n}]$ as a conditional expectation given the random variable X_{t_n}. Then, we take the conditional expectation $\mathbb{E}_X^n$ on both sides of (4.18) to get

$$y_{t_n} = \mathbb{E}_X^n[y_{t_{n+1}}] + \mathbb{E}_X^n[f(X_{t_{n+1}}, y_{t_{n+1}}, z_{t_{n+1}})]\Delta t_n + \mathbb{E}_X^n[R_Y^n], \tag{4.19}$$

where we have used the facts that $\mathbb{E}_X^n[z_{t_n}\Delta W_{t_n}]=0$ and y_{t_n} is $\mathcal{F}_{t_n}^W$ adapted which leads to $\mathbb{E}_X^n[y_{t_n}]=y_{t_n}$. From (4.19), we can see that by dropping the conditional expectation of the error term R_Y^n, i.e. $\mathbb{E}_X^n[R_Y^n]$, we obtain an approximation for y_t. In order to derive a numerical scheme for z_t, we multiply ΔW_{t_n} on both sides of the discretized equation (4.18). Then, we take the conditional expectation $\mathbb{E}_X^n$ to get

$$\begin{aligned}\mathbb{E}_X^n[y_{t_n}\Delta W_{t_n}] = &\ \mathbb{E}_X^n[y_{t_{n+1}}\Delta W_{t_n}]+\mathbb{E}_X^n[f(X_{t_{n+1}},y_{t_{n+1}},z_{t_{n+1}})\Delta W_{t_n}]\Delta t_n\\ &-\mathbb{E}_X^n[z_{t_n}(\Delta W_{t_n})^2]+\mathbb{E}_X^n[R_Y^n\Delta W_{t_n}].\end{aligned}$$

Since y_{t_n} and z_{t_n} are independent from ΔW_{t_n}, we have $\mathbb{E}_X^n[y_{t_n}\Delta W_{t_n}]=0$ and $\mathbb{E}_X^n[z_{t_n}(\Delta W_{t_n})^2]=z_{t_n}\Delta t_n$. Therefore, the above equation can be rewritten as

$$\begin{aligned}z_{t_n}\Delta t_n = &\ \mathbb{E}_X^n[y_{t_{n+1}}\Delta W_{t_n}]+\mathbb{E}_X^n[f(X_{t_{n+1}},y_{t_{n+1}},z_{t_{n+1}})\Delta W_{t_n}]\Delta t_n\\ &+\mathbb{E}_X^n[R_Y^n\Delta W_{t_n}].\end{aligned} \tag{4.20}$$

Now we combine (4.17), (4.19) and (4.20), and drop the error terms R_X^n, $\mathbb{E}_X^n[R_Y^n]$ and $\mathbb{E}_X^n[R_Y^n\Delta W_{t_n}]$ to get

$$\begin{aligned}X_{t_{n+1}} &\approx X_{t_n}+b(X_{t_n},y_{t_n},z_{t_n})\Delta t_n+\sigma\Delta W_{t_n},\\ y_{t_n} &\approx \mathbb{E}_X^n[y_{t_{n+1}}]+\mathbb{E}_X^n[f(X_{t_{n+1}},y_{t_{n+1}},z_{t_{n+1}})]\Delta t_n,\\ z_{t_n}\Delta t_n &\approx \mathbb{E}_X^n[y_{t_{n+1}}\Delta W_{t_n}]+\mathbb{E}_X^n[f(X_{t_{n+1}},y_{t_{n+1}},z_{t_{n+1}})\Delta W_{t_n}]\Delta t_n.\end{aligned} \tag{4.21}$$

Denote X_{n+1} to be our numerical solution for X_t at time instant t_{n+1}, and y_n and z_n to be numerical solutions for y_t and z_t at time instant $t=t_n$, we could obtain numerical schemes for X_{n+1}, y_n and z_n from (4.21). Specifically, for the time level $n=N-1,N-2,\cdots,2,1,0$, let $y_N=y_T$, $z_N=z_T$ and derive z_T from Ψ (from the identity (2.2)), we introduce the numerical schemes to solve the FBSDE system (2.1) recursively on the time interval $[t_n,t_{n+1}]$ as following

$$\begin{aligned}X_{n+1} &= X_n+b(X_n,y_n,z_n)\Delta t_n+\sigma\Delta W_{t_n},\\ y_n &= \mathbb{E}_X^n[y_{n+1}]+\mathbb{E}_X^n[f(X_{n+1},y_{n+1},z_{n+1})]\Delta t_n,\\ z_n\Delta t_n &= \mathbb{E}_X^n[y_{n+1}\Delta W_{t_n}]+\mathbb{E}_X^n[f(X_{n+1},y_{n+1},z_{n+1})\Delta W_{t_n}]\Delta t_n.\end{aligned} \tag{4.22}$$

In the scheme (4.22) we observe that the approximation for X_{n+1} depends on y_n and z_n. On the other hand, X_{n+1} is also required to get y_n and z_n. In this way, iteration is needed when calculating numerical solutions for X_t, y_t and z_t. To obtain a solvable scheme for coupled FBSDEs, we introduce the following algorithm: With the initialization $y_N=y_T$ and $z_N=z_T$, and for the recursive step $n=N-1,N-2,\cdots,2,1,0$, we solve for X_{n+1}, y_n and z_n iteratively in a loop. Specifically, we let $y_n^0=y_{n+1}$, $z_n^0=z_{n+1}$ at the time step n. For $l=0,1,2,\cdots$, the iteration version of schemes (4.22) is given by

$$\begin{aligned}X_{n+1}^{l+1} &= X_n+b(X_n,y_n^l,z_n^l)\Delta t_n+\sigma\Delta W_{t_n},\\ y_n^{l+1} &= \mathbb{E}_X^n[y_{n+1}]+\mathbb{E}_X^n[f(X_{n+1}^{l+1},y_{n+1},z_{n+1})]\Delta t_n,\\ z_n^{l+1}\Delta t_n &= \mathbb{E}_X^n[y_{n+1}\Delta W_{t_n}]+\mathbb{E}_X^n[f(X_{n+1}^{l+1},y_{n+1},z_{n+1})\Delta W_{t_n}]\Delta t_n.\end{aligned} \tag{4.23}$$

When certain criteria, which is typically set to $\max(\|y_n^{l+1}-y_n^l\|,\|z_n^{l+1}-z_n^l\|)<\epsilon$ (or $l+1\le L$) for some user defined parameter $\epsilon>0$ (or $L\in\mathbb{N}$), is satisfied, we let $X_{n+1}=X_{n+1}^{l+1}$, $y_n=y_n^{l+1}$, and $z_n=z_n^{l+1}$.

Remark: At each recursive step $t_{n+1}\to t_n$, every term in the scheme (4.23) is $\mathcal{F}_{t_n}^W$ adapted. Therefore, X_n should be considered as known and some effective spatial representation for X_{t_n} can be used to construct X_n [8, 26].

4.2 Numerical solution for FBDSDEs

Next, we derive numerical methods for solving the time inverse FBDSDE system (3.14) corresponding to the optimal filtering problem (3.13) as the foot stone in our numerical approach. For the same temporal partition $\Pi_N(0,T)$, we consider the following FBDSDE system on the time interval $[t_n, t_{n+1}]$:

$$\begin{aligned}
\overleftarrow{X}_{t_n} =& \overleftarrow{X}_{t_{n+1}} - \int_{t_n}^{t_{n+1}} b(\overleftarrow{X}_s, \overleftarrow{y}_s, \overleftarrow{z}_s) ds + \int_{t_n}^{t_{n+1}} \sigma d\overleftarrow{W}_r, \\
\tilde{Y}_{t_{n+1}} =& Y_{t_n} - \int_{t_n}^{t_{n+1}} b'(\overleftarrow{X}_s, \overleftarrow{y}_s, \overleftarrow{z}_s) \tilde{Y}_s ds - \int_{t_n}^{t_{n+1}} \tilde{Z}_r d\overleftarrow{W}_r + \int_{t_n}^{t_{n+1}} g^*(\overleftarrow{X}_r) \tilde{Y}_r dM_r,
\end{aligned} \tag{4.24}$$

where b is the drift term in the state process and b' denotes the first order spatial derivative with respect to the $\overleftarrow{X}$ variable. Similar to the discussion on FBSDEs, we derive the numerical solution by discretizing the integrals first. For the SDE in the system (the first equation in (4.24)), we notice that the stochastic integral is a backward Itô integral that integrates backward in time from t_{n+1} to t_n with a given initial variable $\overleftarrow{X}_{t_{n+1}}$, which is equivalent to a standard SDE except that the propagation direction is backward. In this way, we use the right point formula to discretize the deterministic integral to be consistent with the propagation direction, and we have $\int_{t_n}^{t_{n+1}} \gamma d\overleftarrow{W}_r = \gamma \Delta W_{t_n}$ since we assume that γ is an invariant constant in this work. As a result, we obtain the following discretization equation

$$\overleftarrow{X}_{t_n} = \overleftarrow{X}_{t_{n+1}} - b(\overleftarrow{X}_{t_{n+1}}, \overleftarrow{y}_{t_{n+1}}, \overleftarrow{z}_{t_{n+1}}) \Delta t_n + \sigma \Delta W_{t_n} + \overleftarrow{R}_X^n, \tag{4.25}$$

where $\overleftarrow{R}_X^n := b(\overleftarrow{X}_{t_{n+1}}, \overleftarrow{y}_{t_{n+1}}, \overleftarrow{z}_{t_{n+1}}) \Delta t_n - \int_{t_n}^{t_{n+1}} b(\overleftarrow{X}_s, \overleftarrow{y}_s, \overleftarrow{z}_s) ds$ is the approximation error. For the BDSDE (the second equation in (4.24)), we can see that there are two stochastic integrals in the equation with different propagation directions, i.e. $\int_{t_n}^{t_{n+1}} \tilde{Z}_r d\overleftarrow{W}_r$ is a backward Itô integral and $\int_{t_n}^{t_{n+1}} g^*(\overleftarrow{X}_r) \tilde{Y}_r dM_r$ is a forward Itô integral where M_t is a standard Brownian motion under the induced probability measure $\tilde{P}$. Based on the propagation direction, we use the Euler-Maruyama scheme to approximate the stochastic integrals as

$$\int_{t_n}^{t_{n+1}} \tilde{Z}_r d\overleftarrow{W}_r = \tilde{Z}_{t_{n+1}} \Delta W_{t_n} + \overleftarrow{R}_W^n$$

and

$$\int_{t_n}^{t_{n+1}} g^*(\overleftarrow{X}_r) \tilde{Y}_r dM_r = g^*(\overleftarrow{X}_{t_n}) \tilde{Y}_{t_n} \Delta M_{t_n} + \overleftarrow{R}_M^n,$$

where we denoted $\Delta M_{t_n} := M_{t_{n+1}} - M_{t_n}$, and $\overleftarrow{R}_W^n = \int_{t_n}^{t_{n+1}} \tilde{Z}_r d\overleftarrow{W}_r - \tilde{Z}_{t_{n+1}} \Delta W_{t_n}$, $\overleftarrow{R}_M^n = \int_{t_n}^{t_{n+1}} g^*(\overleftarrow{X}_r) \tilde{Y}_r dM_r - g^*(\overleftarrow{X}_{t_n}) \tilde{Y}_{t_n} \Delta M_{t_n}$ are approximation errors. Substituting the above approximations into the second equation in (4.24) and applying the left point formula to approximate the deterministic integral, we obtain

$$\tilde{Y}_{t_{n+1}} = \tilde{Y}_{t_n} - b'(\overleftarrow{X}_{t_n}, \overleftarrow{y}_{t_n}, \overleftarrow{z}_{t_n}) \tilde{Y}_{t_n} \Delta t_n - \tilde{Z}_{t_{n+1}} \Delta W_{t_n} + g^*(\overleftarrow{X}_{t_n}) \tilde{Y}_{t_n} \Delta M_{t_n} + \overleftarrow{R}_Y^n, \tag{4.26}$$

where $\overleftarrow{R}_Y^n$ is the approximation error for integrals in (4.26) with

$$\overleftarrow{R}_Y^n := \overleftarrow{R}_M^n - \overleftarrow{R}_W^n + b'(\overleftarrow{X}_{t_n}, \overleftarrow{y}_{t_n}, \overleftarrow{z}_{t_n}) \tilde{Y}_{t_n} \Delta t_n - \int_{t_n}^{t_{n+1}} b'(\overleftarrow{X}_s, \overleftarrow{y}_s, \overleftarrow{z}_s) \tilde{Y}_s ds.$$

In the approximation equation (4.26), we have used the classic Euler-Maruyama scheme to describe the stochastic integrals and the left point formula to approximate the deterministic integral. Higher order methods require more sophisticated mathematical derivation due to the fact that the stochastic integrals have different propagation directions and more complicated Itô expansion is needed. We refer to [6] for discussions of higher order schemes. On the other hand, we want to point out that in this work we solve the FBDSDE system in the induced probability space $(\Omega,\mathcal{F},\tilde{P})$ and then transfer the solution back to the original probability space to solve the optimal filtering problem. In this way, the effectiveness of Girsanov transformation has stronger influence to the accuracy of the backward SDE filter. To derive a numerical solution for $\tilde{Y}_t$, we define $\tilde{\mathbb{E}}_X^{n+1}[\cdot]:=\tilde{E}[\cdot|\overleftarrow{X}_{t_{n+1}},\mathcal{M}_{t_{n+1}}]$ and take the conditional expectation $\tilde{\mathbb{E}}_X^{n+1}$ on both sides of (4.26) to get

$$\begin{aligned}\tilde{Y}_{t_{n+1}} =&\tilde{\mathbb{E}}_X^{n+1}[\tilde{Y}_{t_n}]-\tilde{\mathbb{E}}_X^{n+1}[b'(\overleftarrow{X}_{t_n},,\overleftarrow{y}_{t_n},\overleftarrow{z}_{t_n})\tilde{Y}_{t_n}]\Delta t_n\\ &+\tilde{\mathbb{E}}_X^{n+1}[g^*(\overleftarrow{X}_{t_n})\tilde{Y}_{t_n}]\Delta M_{t_n}+\tilde{\mathbb{E}}_X^{n+1}[\overleftarrow{R}_{\tilde{Y}}^n],\end{aligned}\tag{4.27}$$

where we have used that fact that $\tilde{Y}_{t_{n+1}}=\tilde{\mathbb{E}}_X^{n+1}[\tilde{Y}_{t_{n+1}}]$ and $\tilde{\mathbb{E}}_X^{n+1}[Z_{t_{n+1}}\Delta W_{t_n}]=Z_{t_{n+1}}\tilde{\mathbb{E}}_X^{n+1}[\Delta W_{t_n}]=0$ since $\overleftarrow{X}_t$ propagates backward.

Combining (4.25) and (4.27), and dropping the error terms $\overleftarrow{R}_X^n$ and $\tilde{\mathbb{E}}_X^{n+1}[\overleftarrow{R}_{\tilde{Y}}^n]$, we obtain the following approximations for $\overleftarrow{X}_{t_n}$ and $\tilde{Y}_{t_{n+1}}$

$$\begin{aligned}\overleftarrow{X}_{t_n}&\approx\overleftarrow{X}_{t_{n+1}}-b(\overleftarrow{X}_{t_{n+1}},\overleftarrow{y}_{t_{n+1}},\overleftarrow{z}_{t_{n+1}})\Delta t_n+\sigma\Delta W_{t_n},\\ \tilde{Y}_{t_{n+1}}&\approx\tilde{\mathbb{E}}_X^{n+1}[\tilde{Y}_{t_n}]-\tilde{\mathbb{E}}_X^{n+1}[b'(\overleftarrow{X}_{t_n},\overleftarrow{y}_{t_n},\overleftarrow{z}_{t_n})\tilde{Y}_{t_n}]\Delta t_n+\tilde{\mathbb{E}}_X^{n+1}[g^*(\overleftarrow{X}_{t_n})\tilde{Y}_{t_n}]\Delta M_{t_n}.\end{aligned}\tag{4.28}$$

From the formulation of the backward SDE filter, we observe that only the solution $\tilde{Y}_t$ appears in the inner product (3.12) that provides the conditional pdf for the state, and $\tilde{Z}_t$ is not needed to estimate the conditional expectation for a test function Φ. At the same time, we also notice that an approximation for $\tilde{Z}_t$ is not necessary to approximate $\tilde{Y}_t$ from (4.28). Therefore, obtaining numerical approximations for $\overleftarrow{X}_{t_n}$ and $\tilde{Y}_{t_{n+1}}$ is sufficient.

For the time level $n=0,1,2,\cdots,N-1$, we introduce the recursive numerical schemes for the FBDSDE system (3.11) as follows.

$$\begin{aligned}\overleftarrow{X}_n&=\overleftarrow{X}_{n+1}-b\Big(\overleftarrow{X}_{n+1},y_{n+1}(\overleftarrow{X}_{n+1}),z_{n+1}(\overleftarrow{X}_{n+1})\Big)\Delta t_n+\sigma\Delta W_{t_n},\\ \tilde{Y}_{n+1}&=\tilde{\mathbb{E}}_X^{n+1}[\tilde{Y}_n]-\tilde{\mathbb{E}}_X^{n+1}\Big[b'\Big(\overleftarrow{X}_n,y_n(\overleftarrow{X}_n),z_n(\overleftarrow{X}_n)\Big)\tilde{Y}_n\Big]\Delta t_n+\tilde{\mathbb{E}}_X^{n+1}[g^*(\overleftarrow{X}_n)\tilde{Y}_n]\Delta M_{t_n}.\end{aligned}\tag{4.29}$$

where $\overleftarrow{X}_n$ is our numerical solution for $\overleftarrow{X}_t$ at time t_n and $\tilde{Y}_{n+1}$ is the numerical solution for $\tilde{Y}_t$ at t_{n+1} as desired in the backward SDE filter. Also, in the above scheme we have used approximations for solutions y_{t_n}, z_{t_n} of FBSDEs, i.e. y_n, z_n that we obtained by solving (2.1) through the scheme (4.23).

4.3 Computational framework

With the numerical schemes (4.23) and (4.29), we now introduce our computational framework that estimates solutions of FBSDEs by using the backward SDE filter. Recall that we aim to estimate $\hat{\mathcal{V}}_t=(\hat{X}_t,\hat{y}_t,\hat{z}_t)$ based on the observational information $\mathcal{M}_t$, i.e.

$\hat{X}_t = E[X_t|\mathcal{M}_t]$, $\hat{y}_t = E[y_t|\mathcal{M}_t]$, $\hat{z}_t = E[z_t|\mathcal{M}_t]$. The optimal filtering problem that finds $\hat{\mathcal{V}}_t$ is described by a state process

$$dX_t = b(X_t, y_t, z_t) + \sigma dW_t,$$

which coincides with the forward SDE in FBSDE (2.1); and an observation process

$$dM_t = g(X_t) + dB_t,$$

which provides observational information $\mathcal{M}_t$ about X_t. From the discussions in Section 3, we know that the FBDSDE system corresponding to the above optimal filtering problem is

$$\begin{aligned} d\overleftarrow{X}_r &= b(\overleftarrow{X}_r, \overleftarrow{y}_r, \overleftarrow{z}_r)dr - \sigma d\overleftarrow{W}_r, && \overleftarrow{X}_t = x, \\ d\tilde{Y}_r &= -b'(\overleftarrow{X}_r, \overleftarrow{y}_r, \overleftarrow{z}_r)\tilde{Y}_r dr - \tilde{Z}_r d\overleftarrow{W}_r + g^*(\overleftarrow{X}_r)\tilde{Y}_r dM_r, && \tilde{Y}_0 = p_0(X_0), \end{aligned} \tag{4.30}$$

where the solution $\tilde{Y}_t$ of the FBDSDE system (4.30) is the un-normalized filtering density for the state X_t.

The computational framework in this work follows the methodology of the backward SDE filter. We first solve the FBSDE system (2.1) through scheme (4.23) to obtain approximate solutions for y_t and z_t so that we have the complete formulation of the optimal filtering problem. Then we solve the FBDSDE system (4.30) to obtain the approximate solution for $\tilde{Y}_t$ by using the scheme (4.29). Since we solve the FBDSDE system under the induced probability $\tilde{P}$, finally we transfer the solution $\tilde{Y}_t$ to the original probability space by using the Girsanov transformation. Specifically, with the approximate solution $\tilde{Y}_{n+1}$ at time instant t_{n+1}, we let $Y_{n+1} := Q_{t_n}^{t_{n+1}} \tilde{Y}_{n+1}/C$ be the approximate filtering density as desired, where C is a normalization factor.

In what follows, we describe our computational framework and summarize the entire algorithm.

Initialization: Generate a temporal partition $\Pi_N(0,T)$. Initialize Y_0 by setting $Y_0 = p_0$, and initialize y_N and z_N by using the side condition given in (2.1). Define a positive constant $\epsilon > 0$ (or $L \in \mathbb{N}$) as the stopping criteria for iteration.

Numerical solution for the FBSDE system: For $n = N-1, N-2, \cdots, 1, 0$, solve the FBSDE system by using the iteration scheme (4.23) to get $\{y_n\}_{n=0}^{N-1}$ and $\{z_n\}_{n=0}^{N-1}$.

Implement the backward SDE filter: For $n = 0, 1, 2, \cdots, N-2, N-1$,

- *Numerical solution for the FBDSDE system:* Solve the FBDSDE system corresponding to the optimal filtering problem (3.13) by using the scheme (4.29) to get Y_{n+1}.

- *Girsanov transformation:* Apply the Girsanov transform to obtain the filtering density $Y_{n+1} (= Q_{t_n}^{t_{n+1}} \tilde{Y}_{n+1}/C)$ under the original probability space $(\Omega, \mathcal{F}, P)$.

- *Solution estimation:* Calculate $\hat{X}_{n+1} = < x, Y_{n+1} >$, $\hat{y}_{n+1} = < y_{n+1}, Y_{n+1} >$, $\tilde{z}_{n+1} = < z_{n+1}, Y_{n+1} >$.

In order to implement the above algorithm, we need a set of spatial points to describe the random variables X_t and $\overleftarrow{X}_t$, which can be considered as the spatial approximation for our algorithm. There are many effective methods to achieve this goal, such like the uniform tensor product mesh, sparse grid/adaptive sparse grid mesh [26], quasi-Monte Carlo based meshfree points [6], and adaptive meshfree points based on sequential Monte Carlo methods [8]. In addition, approximation methods to simulate conditional expectations are needed in both schemes (4.23) and (4.29). Well known approaches include the Monte Carlo method, quasi-Monte Carlo methods and numerical integration methods [29].

In the following section, we demonstrate the performance of our optimal filtering method for estimating solutions of FBSDEs, and we use the classic uniform mesh to represent X_t and $\overleftarrow{X}_t$, and use Gaussian-Hermite quadrature to approximate conditional expectations.

5 Numerical experiments

In this section, we verify the performance of our algorithm. The main theme of our approach is to use the inner product of the *conditional pdf* for the state of the forward SDE in the target FBSDE system and the *solutions of FBSDEs* to obtain data informed solution estimations. Therefore, the primary computational tasks of this algorithm are numerical solutions for FBSDEs and numerical solutions for FBDSDEs, which would provide conditional pdf for the forward SDE X_t. In the first example we present a demonstrate experiment to confirm the convergence of our algorithm under theoretical framework by calculating the inner product of solutions of FBSDEs and solutions of FBDSDEs. In the second example we examine the effectiveness of our algorithm in estimating solutions of coupled FBSDEs with observational data. In Example 3, we solve a linear stochastic feedback control problem to demonstrate the practical application potential of the computational framework we introduced in this work.

5.1 Example 1.

Consider the following coupled FBSDE system

$$\begin{aligned} dX_t &= y_t dt + \sigma dW_t \\ dy_t &= \Big(\frac{\sin(X_t+\alpha)}{\sigma} z_t - \frac{1}{2}\sigma^2 y_t\Big) dt + z_t dW_t. \end{aligned} \tag{5.31}$$

with analytical solution $y_t = \sin(X_t+\alpha)$ and $z_t = \cos(X_t+\alpha)\sigma$, where α is a given constant and σ is the diffusion coefficient in the forward SDE. Then, we introduce the following FBDSDE system that mimics the time inverse FBDSDE system corresponding to an optimal filtering problem

$$\begin{aligned} d\overleftarrow{X}_t &= y_t dt + \sigma d\overleftarrow{W}_t \\ dY_t &= Y_t\Big(2(\overleftarrow{X}_t-\beta)y_t - \big(-1+2(\overleftarrow{X}_t-\beta)^2\big)\sigma^2 + \frac{1}{8}\Big) dt \\ &\quad + Z_t d\overleftarrow{W}_t + \frac{1}{2}Y_t dB_t. \end{aligned} \tag{5.32}$$

Although the deterministic integral in the BDSDE part of the above FBDSDE system is not exactly $\int b'(X_s, y_s, z_s) ds$ as desired in the aforementioned framework, it provides a close analogue which can be used to examine the convergence performance of our methodology. We can derive that the analytic solution Y_t for (5.32) is $Y_t = \exp\big(-(\overleftarrow{X}_t-\beta)^2 + \frac{B_t}{2}\big)$, which looks like a typical unnormalized Gaussian distribution.

In this example, we choose the parameters in (5.31) and (5.32) as $\alpha = 1$, $\beta = -1$, $\sigma = 0.25$, and show the convergence trend in approximating y_t, z_t in Table 1. The "Error

Table 1: Example 1: Approximation errors for y_t, z_t

Stepsize	$\Delta t = 2^{-3}$	$\Delta t = 2^{-4}$	$\Delta t = 2^{-5}$	$\Delta t = 2^{-6}$	$\Delta t = 2^{-7}$	CR
Error y_t	$2.508e-02$	$1.052e-02$	$5.049e-03$	$2.431e-03$	$1.189e-03$	1.091
Error z_t	$7.915e-03$	$4.723e-03$	$2.346e-03$	$1.148e-03$	$5.652e-04$	0.965

y_t" and "Error z_t" that we present in the table are errors in L^2 norm corresponding to the given partition stepsize, i.e. $\Delta t = 2^{-3}, 2^{-4}, 2^{-5}, 2^{-6}, 2^{-7}$, and "CR" stands for *convergence rate*. We can see from the table that the convergence rate for the numerical solutions of FBSDEs is first order. In Table 2, we show the root mean square errors (RMSE) in L^2 norm for the Y_t approximation. In this experiment, we also use partition stepsize $\Delta t = 2^{-3}, 2^{-4}, 2^{-5}, 2^{-6}, 2^{-7}$ to be consistent with the (y_t, z_t) approximations, and repeat the numerical experiment 300 time to calculate RMSEs. As a result, we can see from the table that numerical scheme (4.29) for solving the FBDSDE system has half order convergence rate, i.e. $CR = 0.687$ in this experiment.

Table 2: Example 1: Approximation errors for Y_t

Stepsize	$\Delta t = 2^{-3}$	$\Delta t = 2^{-4}$	$\Delta t = 2^{-5}$	$\Delta t = 2^{-6}$	$\Delta t = 2^{-7}$	CR
RMSE Y_t	$3.005e-02$	$1.902e-02$	$1.172e-02$	$7.8493e-03$	$4.316e-03$	0.687

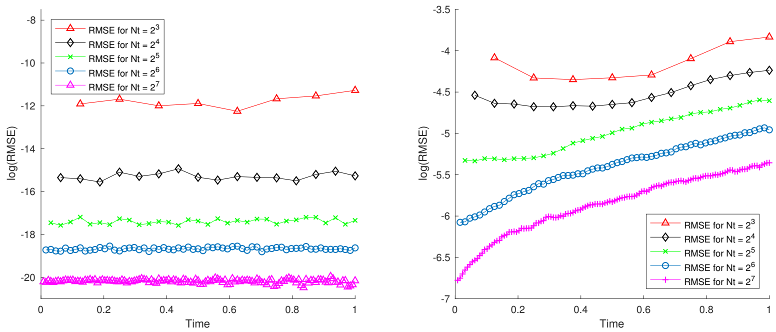
Since $\hat{X}_t$ is the mean value of X_t variable whose accuracy is based on the Y_t approximation that is presented in Table 2, in this example we examine numerical approximations for $< y_t, Y_t >$ and $< z_t, Y_t >$, which are proportional to $\hat{y}_t$ and $\hat{z}_t$ under the optimal filtering framework. In Figure 1, we plot the logarithmic RMSEs of $< y_t, Y_t >$ approximations and $< z_t, Y_t >$ approximations in subplot (a) and subplot (b), respectively. In each subplot, we use red, black, green, blue and magenta curves to represent approximation logarithmic RMSEs with stepsize $\Delta t = 2^{-3}, 2^{-4}, 2^{-5}, 2^{-6}, 2^{-7}$, respectively. From the figure, we can see the convergence trend of our algorithm with decreasing temporal stepsizes under the theoretical framework,

5.2 Example 2.

In this example, we present the effectiveness of our algorithm in estimating the solution $\mathcal{V}_t$ of the FBSDE system with observational data. To proceed, we consider the following FBSDE system

$$\begin{aligned} dX_t &= \Big(y_t - \arctan(X_t) - \frac{t}{2}\Big)dt + \sigma dW_t, \\ dy_t &= \Big(\frac{1}{2} - \frac{X_t}{(1+X_t^2)}\sigma z_t\Big)dt + z_t dW_t, \end{aligned} \tag{5.33}$$

where the SDE and BSDE are coupled. We can derive that the solutions y_t and z_t of the above FBSDE system are $y_t = \arctan(X_t) + \frac{t}{2}$ and $z_t = \frac{1}{1+X_t^2}\sigma$. In order to estimate the



(a) Convergence for $<y_t, Y_t>$ (b) Convergence for $<z_t, Y_t>$

Figure 1: Convergence of the estimations: subplot (a) shows estimation errors for $<y_t, Y_t>$; subplot (b) shows estimation errors for $<z_t, Y_t>$.

solution, in this example we assume that we receive observational data on the solution X_t of the FBSDE system (5.33). Specifically, we introduce the following observation process to give us X_t measurements

$$dM_t = X_t dt + R dB_t,$$

where R is the variance for the noise B_t and the observation function is linear, i.e. $g(X_t) = X_t$. As a result, we aim to solve the following optimal filtering problem

$$\begin{aligned} dX_t =& \big(y_t - \arctan(X_t) - \frac{t}{2}\big)dt + \sigma dW_t, \\ dM_t =& X_t dt + R dB_t, \end{aligned} \tag{5.34}$$

We observe that with the analytical solution $y_t = \arctan(X_t) + \frac{t}{2}$, the forward SDE in (5.33) is equivalent to a standard Brownian motion. Therefore, given the solution y_t of the FBSDE system, the optimal filtering problem (5.34) is equivalent to the following linear version

$$\begin{aligned} dX_t =& \sigma dW_t, \\ dM_t =& X_t dt + R dB_t, \end{aligned} \tag{5.35}$$

which can be solved analytically by the classic Kalman filter. Since the filtering problem (5.35) does not depend on the FBSDE system and the Kalman filter provides the *exact solution*, we use the Kalman filter in this example as the benchmark reference performance for the estimation.

In Figure 2, we present the estimation for $\hat{X}_t$ over the time interval $[0,2]$, which can be considered as the tracking performance for the state of X_t given the observation M_t. The observation gap we choose in this experiment is $\Delta t = 0.04$, therefore we track X_t for 50 steps, i.e. $N = 50$. The black curve marked by “plus signs” shows the real state of X_t; the red curve marked by “crosses” represent the estimation obtained by the backward SDE filter approach for the optimal filtering problem (5.34); and the blue curve marked by “circles” is the Kalman filter estimation for the equivalent linear problem (5.35). We can see from this figure that the backward SDE filter which implements our computational framework provides very similar results to the Kalman filter which only works

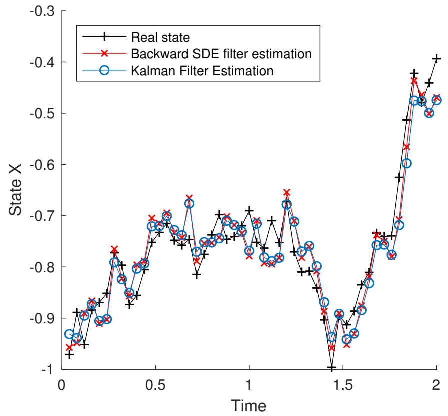


Figure 2: Estimation for the state of X_t. (Example 2)

for simplified problems. We want to point out that although the Kalman filter method does not always match the real state in this experiment, it gives the best estimate for the state based on observations due to the noises contained in the observational data.

Next, we verify the capability of our method in solving nonlinear problems, especially in dealing with nonlinear observations. In the following experiment, we use a nonlinear observation function to collect observational information, i.e. let

$$dM_t = (X_t)^2 dt + R dB_t, \tag{5.36}$$

where the data is collected through the nonlinear function $(X_t)^2$. In Figure 3, we present

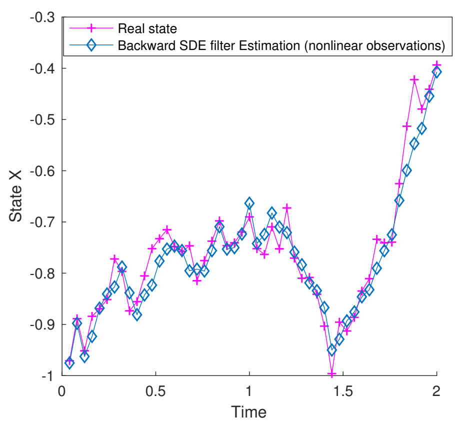
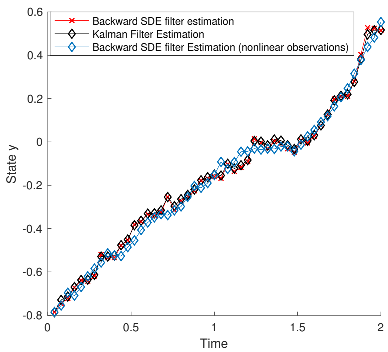


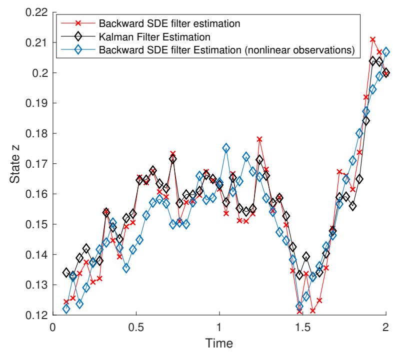
Figure 3: Estimation for the state of X_t with indirect observations. (Example 2)

the backward SDE filter estimation for X_t by using the nonlinear observation process (5.36). The magenta curve marked by "plus signs" is the real state and the blue curve marked by "diamonds" is the backward SDE filter estimation. We can see that although we use nonlinear observations for the state, our method still provides close estimation for the state.

To examine the solution estimation performance for the FBSDE system, in Figure 4 we plot the estimate solutions y_t and z_t, i.e. $\hat{y}_t$ and $\hat{z}_t$, in subplots (a) and (b) respectively. In each subplot, we use the red curve marked by "crosses" to represent



(a) Estimation for y_t



(b) Estimation for z_t

Figure 4: Estimation for the state of y_t and z_t. (Example 2)

the backward SDE filter estimation with the linear observation introduced in (5.34); the blue curve marked by "diamonds" shows the backward SDE filter estimation with the nonlinear observation process (5.36); the black curve marked by "diamonds" is the Kalman filter estimation derived from the equivalent linear filtering problem (5.35) with analytical solutions y_t and z_t, which performs as the reference solution. From this experiment, we can see that our estimations are very close to the reference estimation obtained by the Kalman filter, which indicates the effectiveness of our computational framework in estimating solutions of FBSDEs.

5.3 Example 3.

In this example, we apply our method to solve a feedback stochastic optimal control problem. The optimal control part of the problem is the classic linear-quadratic (LQ) model, which is described by the following controlled state equation

$$\begin{cases} dX_s = [AX_s + Bu_s]ds + CdW_s, & s \in [t,T], \\ X_t = x, \end{cases} \tag{5.37}$$

where A, B and C are given constants and u_t is the control process that governs the equation of X_t. In the stochastic optimal control problem, we want to determine a control process u_t that minimizes the convex cost functional which is defined as following

$$J(t,x;u) = \mathbb{E}\Big[\int_t^T \big(q(X(s)) + |u(s)|^2\big)ds + h(X(T))\Big].$$

Here, $q(\cdot)$ and $h(\cdot)$ are convex coercive functions. It can be shown that the optimal control u_t uniquely exists to minimize the cost function, i.e. $\inf_{u_t \in \mathcal{U}} J(t,x;u)$, and satisfies the expression $u_t = -By_t$, where y_t is the solution of the following coupled FBSDE

$$\begin{cases} dX_s = \left[AX_s - B^2 y_s\right] ds + C dW_s \\ dy_s = -\left[Ay_s - q_x(X_s)\right] ds + z_s dW_s \\ X_t = x, \quad y_T = h_x(X_T). \end{cases} \tag{5.38}$$

As a result, solving the aforementioned optimal control problem is equivalent to solving the FBSDE system (5.38), and we can obtain the optimal control u_t by using the numerical solution for y_t which could be very accurate. However, the control strategy we obtain in this way only gives us a function of X_t and in practice the control should depend on the state of X_t which is actually unknown. In order to make real time estimation for X_t and make adjustment to better design the control strategy u_t, we assume that we could receive observations on the state equation X_t and denote the observation process as

$$M_t = X_t + \dot{B}_t.$$

Therefore, the optimal filtering estimate $\hat{X}_t$ would give us the timely feedback for the current state of controlled process and the estimate solution $\hat{y}_t$ for the FBSDE system (5.38) would lead to the control strategy that we should use at the time instant t based on the observational data.

To implement the numerical experiment, we choose $q(x) = 0.15\ x^2$ and $h(x) = 0.1x^2$ in the cost function. For the controlled process, we let $A = 0.1$, $B = 0.5$ and $C = 0.25$, and assume that our initial guess for X_0 to be $\hat{X}_0 = 0$ while the actual state is $X_0 = 0.1$. We assume that we receive observations in every 0.02 second, i.e. $\Delta t = 0.02$ and estimate the controlled process and the control process for 1 second, i.e. $N = 50$. Since this is a classic benchmark problem which can be well solved by the Kalman filter, we compare our estimate with the estimate obtained by using the Kalman filter method. In Figure 5,

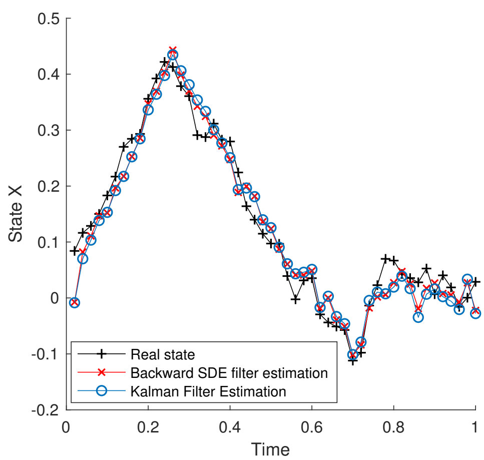


Figure 5: Estimation for the state X_t. (Example 3)

we plot the estimate for X_t over the time interval $[0,1]$ to demonstrate the performance of filtering methods in state estimation. The black curve marked by "plus signs" is the real controlled state that we observe; the red curve marked by "crosses" is the backward SDE filter estimate for the state of X_t; the blue curve marked by "circles" shows the estimate obtained by using the Kalman filter method. We can see that both methods provide similar results, which are very close to the real state of the controlled dynamics.

Finally, we present the data informed estimation for the solution y_t, i.e. $\hat{y}_t$, given the observational information contained in M_t. With the estimated $\hat{y}_t$, we can derive the real time feedback optimal control as $\hat{u}_t = -B\hat{y}_t$. In Figure 6, we show the estimation

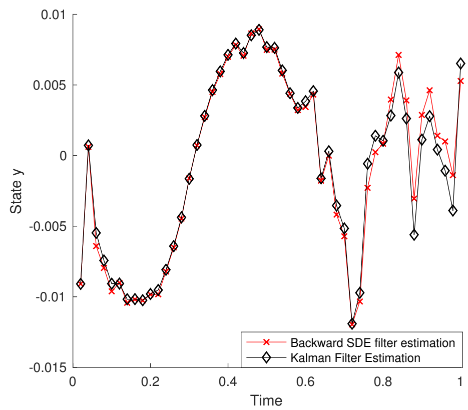


Figure 6: Example 3: Estimation for the state of y_t.

performance by using the backward SDE filter and the Kalman filter. The red curve marked by "crosses" is the estimate obtained by using the backward SDE filter and the black curve marked by "diamonds" is the estimate obtained by using the Kalman filter which we use in this experiment as the benchmark reference performance. We can see that the backward SDE filter estimate is very close to the Kalman filter estimate, which is also consistent with the estimation for the controlled process (demonstrated in Figure 5). We also want to mention that our method is not restricted to the linear case of the feedback stochastic optimal control problem, which is the only case that the Kalman filter is applicable. In this way, our method would allow us to solve broader application problems.

6 Acknowledgement

This work is partially supported by the Scientific Discovery through Advanced Computing (SciDAC) program funded by U.S. Department of Energy, Office of Science, Advanced Scientific Computing Research through CompFUSE project and FASTMath Institute. The first author also acknowledges support by U.S. National Science Foundation under Contract DMS-1720222.

References

[1] C. Andrieu, A. Doucet, and R. Holenstein. Particle markov chain monte carlo methods. *J. R. Statist. Soc. B*, 72(3):269–342, 2010.

[2] F. Bao, R. Archibald, and P. Maksymovych. Backward sde filter for jump diffusion processes and its applications in material sciences. *Communications in Computational Physics*, to appear.

[3] F. Bao, Y. Cao, and X. Han. Forward backward doubly stochastic differential equations and optimal filtering of diffusion processes, arxiv: 1509.06352. 2016.

[4] Feng Bao, Yanzhao Cao, and Hongmei Chi. Adjoint forward backward stochastic differential equations driven by jump diffusion processes and its application to nonlinear filtering problems. *Int. J. Uncertain. Quantif.*, 9(2):143–159, 2019.

[5] Feng Bao, Yanzhao Cao, Xiaoying Han, and Jinglai Li. Efficient particle filtering for stochastic Korteweg–de Vries equations. *Stoch. Dyn.*, 17(2):1750008, 20, 2017.

[6] Feng Bao, Yanzhao Cao, Amnon Meir, and Weidong Zhao. A first order scheme for backward doubly stochastic differential equations. *SIAM/ASA J. Uncertain. Quantif.*, 4(1):413–445, 2016.

[7] Feng Bao, Yanzhao Cao, and Weidong Zhao. A backward doubly stochastic differential equation approach for nonlinear filtering problems. *Commun. Comput. Phys.*, 23(5):1573–1601, 2018.

[8] Feng Bao and Vasileios Maroulas. Adaptive meshfree backward SDE filter. *SIAM J. Sci. Comput.*, 39(6):A2664–A2683, 2017.

[9] A. J. Chorin and X. Tu. Implicit sampling for particle filters. *Proc. Nat. Acad. Sc. USA*, 106:17249–17254, 2009.

[10] Pierre Del Moral, Aline Kurtzmann, and Julian Tugaut. On the stability and the concentration of extended Kalman-Bucy filters. *Electron. J. Probab.*, 23:Paper No. 91, 30, 2018.

[11] N. El Karoui, S. Peng, and M. C. Quenez. Backward stochastic differential equations in finance. *Math. Finance*, 7(1):1–71, 1997.

[12] N.J Gordon, D.J Salmond, and A.F.M. Smith. Novel approach to nonlinear/non-gaussian bayesian state estimation. *IEE PROCEEDING-F*, 140(2):107–113, 1993.

[13] Jiequn Han, Arnulf Jentzen, and Weinan E. Solving high-dimensional partial differential equations using deep learning. *Proc. Natl. Acad. Sci. USA*, 115(34):8505–8510, 2018.

[14] Yuecai Han, Shige Peng, and Zhen Wu. Maximum principle for backward doubly stochastic control systems with applications. *SIAM J. Control Optim.*, 48(7):4224–4241, 2010.

[15] E. Hirvijoki, C. Liu, G. Zhang, D. del Castillo-Negrete, and D. Brennan. A fluid-kinetic framework for self-consistent runaway-electron simulations. *Physics of Plasmas*, (25):062507, 2018.

[16] Kai Kang, Vasileios Maroulas, Ioannis Schizas, and Feng Bao. Improved distributed particle filters for tracking in a wireless sensor network. *Comput. Statist. Data Anal.*, 117:90–108, 2018.

[17] H. R. Kunsch. Particle filters. *Bernoulli*, 19(4):1391–1403, 2013.

[18] Jin Ma and Jiongmin Yong. *Forward-backward stochastic differential equations and their applications*, volume 1702 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1999.

[19] É. Pardoux and S. Peng. Backward stochastic differential equations and quasilinear parabolic partial differential equations. In *Stochastic partial differential equations and their applications (Charlotte, NC, 1991)*, volume 176 of *Lecture Notes in Control and Inform. Sci.*, pages 200–217. Springer, Berlin, 1992.

[20] É. Pardoux and P. Protter. A two-sided stochastic integral and its calculus. *Probab. Theory Related Fields*, 76(1):15–49, 1987.

[21] Étienne Pardoux and Shi Ge Peng. Backward doubly stochastic differential equations and systems of quasilinear SPDEs. *Probab. Theory Related Fields*, 98(2):209–227, 1994.

[22] Shige Peng and Zhen Wu. Fully coupled forward-backward stochastic differential equations and applications to optimal control. *SIAM J. Control Optim.*, 37(3):825–843, 1999.

[23] Jie Yang, Guannan Zhang, and Weidong Zhao. A first-order numerical scheme for forward-backward stochastic differential equations in bounded domains. *J. Comput. Math.*, 36(2):237–258, 2018.

[24] Jiongmin Yong and Xun Yu Zhou. *Stochastic controls*, volume 43 of *Applications of Mathematics (New York)*. Springer-Verlag, New York, 1999. Hamiltonian systems and HJB equations.

[25] Moshe Zakai. On the optimal filtering of diffusion processes. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 11:230–243, 1969.

[26] Guannan Zhang, Max Gunzburger, and Weidong Zhao. A sparse-grid method for multi-dimensional backward stochastic differential equations. *J. Comput. Math.*, 31(3):221–248, 2013.

[27] Jianfeng Zhang. A numerical scheme for BSDEs. *Ann. Appl. Probab.*, 14(1):459–488, 2004.

[28] Weidong Zhao, Lifeng Chen, and Shige Peng. A new kind of accurate numerical method for backward stochastic differential equations. *SIAM J. Sci. Comput.*, 28(4):1563–1581, 2006.

[29] Weidong Zhao, Yu Fu, and Tao Zhou. New kinds of high-order multistep schemes for coupled forward backward stochastic differential equations. *SIAM J. Sci. Comput.*, 36(4):A1731–A1751, 2014.