A FORWARD-BACKWARD ALGORITHM FOR STOCHASTIC CONTROL PROBLEMS

Using the stochastic maximum principle as an alternative to dynamic programming

Stephan E. Ludwig[1], Justin A. Sirignano[2], Ruojun Huang[3], George Papanicolaou[4]

[1]*Department of Mathematics, Heidelberg University, INF 288, Heidelberg, Germany*

[2]*Department of Management Science and Engineering, Stanford University, Stanford, USA*

[3]*Department of Statistics, Stanford University, Stanford, USA*

[4]*Department of Mathematics, Stanford University, 380 Sloan Hall, Stanford, USA*

StephanLudwig@gmx.de, jasirign@stanford.edu, hruojun@stanford.edu, papanico@math.stanford.edu

Keywords: OPTIMAL STOCHASTIC CONTROL : STOCHASTIC MAXIMUM PRINCIPLE : FORWARD-BACKWARD STOCHASTIC DIFFERENTIAL EQUATIONS

Abstract: An algorithm for solving continuous-time stochastic optimal control problems is presented. The numerical scheme is based on the stochastic maximum principle (SMP) as an alternative to the widely studied dynamic programming principle (DDP). By using the SMP, (Peng, 1990) obtained a system of coupled forward-backward stochastic differential equations (FBSDE) with an external optimality condition. We extend the numerical scheme of (Delarue and Menozzi, 2006) by a Newton-Raphson method to solve the FBSDE system and the optimality condition simultaneously. As far as the authors are aware, this is the first fully explicit numerical scheme for the solution of optimal control problems through the solution of the corresponding extended FBSDE system. We discuss possible numerical advantages to the DDP approach and consider an optimal investment-consumption problem as an example.

1 INTRODUCTION

We consider continuous-time stochastic control problems where the state variable is a controlled stochastic process of Markovian type and the objective function depends on the state and on the control. These type of problems typically appear in mathematical finance and economics. The most common method to solve these problems is the dynamic programming principle (DPP), which leads to the well-known Hamilton-Jacobi-Bellman (HJB) equation. Various numerical schemes take advantage of the DDP's discrete version by performing a backward algorithms or directly solving the HJB partial differential equation using a finite difference scheme.

In this paper, we consider an alternative approach to the problem based on the stochastic maximum principle (SMP), which leads to a system of coupled forward-backward stochastic differential equations (FBSDE) plus an external optimality condition. This was first studied by (Peng, 1990). It is well known that a quasilinear PDE has a FBSDE representation, which is an extension of the well-known Feynman-Kac formula. However, the FBSDE representation cannot be directly applied to the HJB equation unless the optimal control is known as an explicit function. Instead, by using the SMP, we obtain a coupled FBSDE system for the adjoint equations. The coupling arises through the additional optimality condition only.

In addition to reviewing briefly the connection between stochastic control problems and FBSDE systems, the main objective of this paper is to present a complete numerical algorithm by obtaining approximate solutions to a certain class of optimal control problems. We need to use advanced numerical methods for the FBSDE because of the coupling which arises from dependence of the state process on the controls and is therefore connected to the controlled objective function. Therefore, we take advantage of an existing numerical scheme for coupled FBSDEs, initially proposed by (Delarue and Menozzi, 2006) and extend it to satisfy the optimality condition.

The paper's outline is as follows. After the problem definition in section 2, we briefly derive the corresponding FBSDE representation for the adjoint equations and state a verification theorem in section 3. In section 4 we discretize the time-continuous problem and provide the numerical scheme using a Markov chain approximation and a Newton-Raphson method for the optimization. A brief comparison of the computational costs between our method and the standard dynamic programming approach together with first results from an application on an investment-consumption problem are presented in section 5. The paper concludes with some outlook to further development.

2 PROBLEM STATEMENT

Throughout the paper we assume a given probability space $(\Omega, F, \mathbb{P})$, endowed with a d-dimensional Brownian motion $(W_t)_{t\geq 0}$, whose natural filtration is denoted by $\{F_t\}_{t\geq 0}$.

Consider the following problem. The dynamics of a controlled diffusion process X_t, which represents the state of our system, are given by:

$$dX_t = b(X_t, \pi_t)dt + \sigma(X_t, \pi_t)dW_t, \quad X_0 = x \in \mathbb{R}^n, \quad (1)$$

and the goal is to maximize a given objective function with finite time horizon $[0,T]$ over admissible controls $\tilde{\pi} = \{\pi_t\}_{t\in[0,T]} \in \mathbb{A}$:

$$J(t,x,\tilde{\pi}) := E_t^{\tilde{\pi}}\left[\int_t^T f(s,X_s,\pi_s)ds + g(X_T)\middle| X_t = x\right]. \quad (2)$$

Here, $\mathbb{A}$ is the set of all progressively F_t-measurable controls which takes its values π_t in a compact set $A \subset \mathbb{R}^r$. If it exists, we will denote the optimal control by:

$$\tilde{\pi}^* := \arg\max_{\tilde{\pi}\in\mathbb{A}} J(t,x,\tilde{\pi}), \quad \forall (t,x) \in [0,T]\times\mathbb{R}^n, \quad (3)$$

and the value function by:

$$v(t,x) := \sup_{\tilde{\pi}\in\mathbb{A}} J(t,x,\tilde{\pi}), \quad \forall (t,x) \in [0,T]\times\mathbb{R}^n. \quad (4)$$

2.1 General conditions

For each fixed $\pi \in A$, we ensure the existence of a unique solution to the controlled forward SDE (1) by the following assumptions:

- $b(\cdot,\pi), \sigma(\cdot,\pi)$ are uniformly Lipschitz continuous with respect to x:

$$\forall \pi \in A,\ d|b(x_1,\pi) - b(x_2,\pi)| + |\sigma(x_1,\pi) - \sigma(x_2,\pi)| \leq C|x_1 - x_2|, \quad (5)$$

- $b(\cdot,\pi), \sigma(\cdot,\pi)$ satisfy a linear growth condition with respect to x:

$$\forall \pi \in A, \quad |b(x,\pi)| + |\sigma(x,\pi)| \leq C(1+|x|). \quad (6)$$

To ensure the boundedness of the objective function (2) we further assume:

- $f(t,\cdot,\pi), g(\cdot)$ satisfy a quadratic growth condition:

$$\forall t \in [0,T], \forall \pi \in A, \quad |g(x)| + |f(t,x,\pi)| \leq C(1+|x|^2), \quad \forall x \in \mathbb{R}^n. \quad (7)$$

Both proofs can be found in (Pham, 2009) chapter 3.2. For further proceedings we assume also that:

- b, σ, f, g are twice continuously differentiable with respect to x and π:

$$\forall t \in [0,T], \quad (b,\sigma,f,g)(t,\cdot,\cdot) \in C^{1,2}(\mathbb{R}^n, A), \quad (8)$$

- f, g are uniformly concave with respect to x and π.

To be explicit, we use a Markov Chain approximation in section 4.1. In order to calculate the Brownian increments for this Markov chain approximation in (36), σ needs to be invertible. This also means that $d = n$. Otherwise, we can use Quantization methods or Monte Carlo simulations to calculate expectations instead. This would make no difference to the general scheme.

3 THE STOCHASTIC MAXIMUM PRINCIPLE

3.1 Derivation of the FBSDE

Following (Pham, 2009), let us suppose there exists a unique solution $v \in C^{1,3}([0,T)\times\mathbb{R}^n) \cap C^0([0,T]\times\mathbb{R}^n)$ to (4) and an optimal control $\tilde{\pi}^* \in \mathbb{A}$ described in (3) with associated controlled diffusion $\hat{X}_t$ satisfying (1).

The adjoint equations can be derived in two basic steps, namely 1) derive the HJB equation at the optimal control with respect to x:

$$\partial_x\left(\partial_t v(t,\hat{X}_t) + G(t,\hat{X}_t,\pi_t^*,\nabla_x v(t,\hat{X}_t),\nabla_x^2 v(t,\hat{X}_t))\right) = 0, \quad (9)$$

where $G : [0,T]\times\mathbb{R}^n\times A\times\mathbb{R}^n\times\mathbb{R}^{n\times n} \to \mathbb{R}$ is given by:

$$G(t,x,\pi,p,M) := b(x,\pi)p + \tfrac{1}{2}\mathrm{tr}(\sigma\sigma'(x,\pi)M) + f(t,x,\pi). \quad (10)$$

2) Apply Ito's formula to $\nabla_x v(t,\hat{X}_t)$ and plug in the above relation (9). After a few calculations, the adjoint equations come out as:

$$
\begin{gathered}
-d\,\nabla_x v(t,\hat{X}_t) = \\
\nabla_x H(t,\hat{X}_t,\pi_t^*,\nabla_x v(t,\hat{X}_t),\nabla_x^2 v(t,\hat{X}_t)\sigma(\hat{X}_t,\pi_t^*))dt \\
-\nabla_x^2 v(t,\hat{X}_t)\sigma(\hat{X}_t,\pi_t^*)dW_t, \\
\nabla_x v(T,\hat{X}_T) = \nabla_x g(\hat{X}_t),
\end{gathered}
$$

where the so called Hamiltonian $H:[0,T]\times\mathbb{R}^n\times A\times\mathbb{R}^n\times\mathbb{R}^{n\times d}\to\mathbb{R}$ is defined by:

$$
H(t,x,\pi,y,z) := b(x,\pi)y + \mathrm{tr}[\sigma'(x,\pi)z] + f(t,x,\pi). \tag{11}
$$

Furthermore, since G is continuously differentiable with respect to π, we get:

$$
\begin{aligned}
0 &= \partial_\pi G(t,\hat{X}_t,\pi_t^*,\nabla_x v(t,\hat{X}_t),\nabla_x^2 v(t,\hat{X}_t)) \\
&= \partial_\pi H(t,\hat{X}_t,\pi_t^*,\nabla_x v(t,\hat{X}_t),\nabla_x^2 v(t,\hat{X}_t)\sigma(x,\pi)).
\end{aligned}
$$

Assuming concavity of H with respect to π, H must attain a maximum at π_t^*.

Let us summarize the above results. Under the above assumptions and H being concave with respect to π, the triple:

$$
(\hat{X}_t,\hat{Y}_t,\hat{Z}_t) := \left(\hat{X}_t,\nabla_x v(t,\hat{X}_t),\nabla_x^2 v(t,\hat{X}_t)\sigma(\hat{X}_t,\pi_t^*)\right)
$$

is the unique solution to the coupled FBSDE system:

$$
\begin{aligned}
X_t = &\; x + \int_0^t b(X_s,\pi_s^*)ds + \int_0^t \sigma(X_s,\pi_s^*)dW_s, \\
Y_t = &\; \nabla_x g(X_T) + \int_t^T \nabla_x H(s,X_s,\pi_s^*,Y_s,Z_s)ds \\
&- \int_t^T Z_s dW_s,
\end{aligned} \tag{12}
$$

such that the following optimality condition holds:

$$
\pi_t^* = \arg\max_{\pi_t\in A} H(t,\hat{X}_t,\pi_t,\hat{Y}_t,\hat{Z}_t). \tag{13}
$$

3.2 Verification theorem

Theorem 1. Suppose that there exists a unique solution $v\in C^{1,3}([0,T]\times\mathbb{R}^n,\mathbb{R})$ of the value function (4). Let $(\hat{X},\hat{Y}_t,\hat{Z}_t)$ and the control $\tilde{\pi}^* := \{\pi_t^*\}_{t\in[0,T]}$ be associated solutions to the FBSDE system (12) such that the optimality condition (13) holds. Additionally assume that $g(\cdot)$ and $H(t,\cdot,\cdot,\hat{Y}_t,\hat{Z}_t)$ are uniformly concave in (x,π). Then:

- $\tilde{\pi}^*$ is the optimal control of the stochastic control problem (4) and $\hat{X}_t$ is the solution of the associated controlled state process (1),
- $(\hat{Y}_t,\hat{Z}_t) = \left(\nabla_x v(t,\hat{X}_t),\nabla_x^2 v(t,\hat{X}_t)\sigma(\hat{X}_t,\pi_t^*)\right)$

Proof: The proof of the first statement is given in (Pham, 2009) Theorem 6.5.4.

For the second step we define a function $u(t,\hat{X}_t)$ via its first derivative $\nabla_x u(t,\hat{X}_t) := \hat{Y}_t$ and its terminal value $u(T,\hat{X}_T) := g(\hat{X}_T)$ and apply Ito's formula to $\nabla_x u$. Comparing the diffusion term with the backward SDE for $\hat{Y}_t$ we get $\hat{Z}_t = \nabla_{xx}^2 u(t,\hat{X}_t)\sigma(\hat{X}_t,\pi_t^*)$. Comparing the drift terms we get a third order PDE which is exactly the derivative of the HJB equation with respect to x. Since the solution of this PDE is unique, the verification is completed by using the verification theorem for the HJB equation.

The concavity of H is an important condition for the connection of the optimal control problem and the optimality condition for the FBSDE. This condition mainly specifies the problem class for applications. Recall that we already assumed f,g to be concave in section 2.1.

4 A numerical scheme for solving the FBSDE

Let us state the complete (coupled) FBSDE problem one more time:

$$
\begin{aligned}
X_t &= X_0 + \int_0^t b(X_s,\pi_s^*)ds + \int_0^t \sigma(X_s,\pi_s^*)dW_s, \\
Y_t &= Y_T + \int_t^T \nabla_x H(s,X_s,\pi_s^*,Y_s,Z_s)ds - \int_t^T Z_s dW_s, \\
\pi_t^* &= \arg\max_{\pi_t\in A} H(t,X_t,\pi_t,Y_t,Z_t),
\end{aligned} \tag{14}
$$

where $X_0 = x$ and $Y_T = \nabla_x g(X_T)$.

4.1 The discrete problem

Following (Kushner and Dupuis, 1992), let us define a fixed, scalar approximation parameter $h>0$. In the following, the superscript h denotes the dependency on this approximation parameter. Let $\Delta t_k^h > 0$, for $k=0,\ldots,N-1$, $N<\infty$, discretize the time interval $[0,T]$ by defining:

$$
t_0 := 0, \quad t_1 := \Delta t_0^h, \quad t_k := \sum_{i=0}^{k-1} \Delta t_i^h, \quad t_N := T.
$$

Suppose that $\Delta t_k^h \to 0$, as $h\to 0$. Let:

$$
\mathrm{C}_k := \left\{(\xi^j)_{j\in I_k}, I_k\subset\mathbb{N}\right\} \subset \mathbb{R}^n, \quad \forall k=0..N, \tag{15}
$$

be a spatial grid satisfying $\mathrm{C}_j \subset \mathrm{C}_i$, for all $j<i$. Furthermore, we denote the set of admissible, piecewise constant controls by $\mathbb{A}^h = \{\tilde{\pi}\in\mathbb{A}\,|\,\pi_t \text{ is constant over } [t_k,t_{k+1}), \forall k\le N-1\}$.

To calculate the propagation of the state process from time t_k to t_{k+1}, we choose the Markov chain approximation method here. As mentioned above, we can use different methods like Quantization. The Markov chain is defined by its transition probabilities:

$$
\begin{gathered}
\forall k=0,\ldots,N-1, \forall \xi^j\in\mathrm{C}_k, \forall \xi^l\in\mathrm{C}_{k+1}, \\
P(\xi^j,\xi^l\,|\,\pi_k(\xi^j)) = p_k^{jl}(\pi_k(\xi^j)).
\end{gathered} \tag{16}
$$

We denote $\Delta\xi_k = \xi_{k+1} - \xi_k$. The discrete Markov chain approximation ξ_k converges to the real state process (1) as $h \to 0$, if the following local consistency conditions hold:

$$\begin{aligned} &E_k\Delta\xi_k = b(\xi_k,\pi_k)\Delta t_k^h(\xi_k,\pi_k) + o(\Delta t_k^h(\xi_k,\pi_k)),\\ &\mathrm{Var}_k\Delta\xi_k^h = [\sigma\sigma'](\xi_k,\pi_k)\Delta t_k^h(\xi_k,\pi_k) + o(\Delta t_k^h(\xi_k,\pi_k)),\\ &\sup_{k,\omega}|\Delta\xi_k| \to 0, \text{ as } h \to 0, \end{aligned} \tag{17}$$

where E_k is the conditioned expectation given all information up to time k and Var_k is the variance accordingly. For methods to derive proper transition probabilities see (Kushner and Dupuis, 1992).

For any control $\tilde{\pi}^h \in \mathbb{A}^h$, we define the following approximation of the objective function (2):

$$\begin{aligned} &J^h(t_k,x,\tilde{\pi}^h) = \\ &E_k^{\tilde{\pi}^h}\left[\sum_{i=k}^{N-1} f(t_i,\xi_i,\pi_i^h)\Delta t_i^h + g(\xi_N)\,\middle|\, \xi_k = x\right], \end{aligned} \tag{18}$$

and the approximation of the value function (4) by:

$$V_k^h(x) = \max_{\tilde{\pi}^h\in\mathbb{A}^h} J^h(t_k,x,\tilde{\pi}^h). \tag{19}$$

4.2 A controlled forward-backward algorithm

Starting from final time T going backwards, let us suppose we have already calculated the approximations $Y_{k+1}^h(\cdot)$, $Z_{k+1}^h(\cdot)$ for (41). Now, let us use these functions as natural predictors for the still unknown $Y_k^h(\cdot)$ and $Z_k^h(\cdot)$ respectively. Then $\forall \xi^j \in \mathrm{C}_k$ we calculate:

$$\pi_{t_k}^*(\xi^j) = \arg\max_{\pi\in A} H(t_k,\xi^j,\pi,Y_{k+1}^h(\xi^j),Z_{k+1}^h(\xi^j)). \tag{20}$$

For notational simplicity we denote $\pi_{t_k}^*(\xi^j)$ by π_k^{j*}. If we think of the control π_k^* in (41) as being a function of t,X_t,Y_t,Z_t:

$$\pi_k^*(t,X_t,Y_t,Z_t) = \arg\max_{\pi\in A} H(t_k,X_t,\pi,Y_t,Z_t), \tag{21}$$

we can replace the control variable by this function in the FBSDE system (41) and receive a fully coupled FBSDE system without control.

To solve the coupled FBSDE system we make use of existing numerical schemes. In particular, we choose the algorithm for coupled FBSDE systems proposed by (Delarue and Menozzi, 2006) using the pre-calculated controls π_k^{j*} (20). The algorithm then reads as follows:

$$\begin{aligned} V_k^h(\xi^j) &= f(t_k,\xi^j,\pi_k^{j*})\Delta t_k^h \\ &+ \textstyle\sum_{\xi^l\in\mathrm{C}_{k+1}} p_k^{jl}(\pi_k^{j*})V_{k+1}^h(\xi^l), \end{aligned} \tag{22}$$

$$\begin{aligned} Y_k^h(\xi^j) &= \nabla_x H(t_k,\xi^j,\pi_k^{j*},Y_{k+1}^h(\xi^j),Z_{k+1}^h(\xi^j))\,\Delta t_k^h \\ &+ \textstyle\sum_{\xi^l\in\mathrm{C}_{k+1}} p_k^{jl}(\pi_k^{j*})Y_{k+1}^h(\xi^l), \end{aligned} \tag{23}$$

$$Z_k^h(\xi^j) = \frac{1}{\Delta t_k^h}\sum_{\xi^l\in\mathrm{C}_{k+1}} p_k^{jl}(\pi_k^{j*})Y_{k+1}^h(\xi^l)\Delta W_k^{jl}, \tag{24}$$

where ΔW_k^{jl} is calculated via the Euler relationship:

$$\Delta W_k^{jl} = \sigma^{-1}(\xi^j,\pi_k^{j*})(\xi^l - \xi^j - b(\xi^j,\pi_k^{j*})\Delta t_k^h). \tag{25}$$

The main contribution in our paper is the explicit pre-calculation of π^* via (20). This is the essential step in our scheme to solve an optimal control problem through a FBSDE representation. The significance is that the optimization is performed externally from the backward calculations for V_t,Y_t,Z_t in (22), (23), (24) and does not include the calculation of expectations. This will be outlined more precisely in the following sections.

4.3 Optimization

At first appearance, the difference of the above method to dynamic programming is that in the former method the optimization does not have to be performed over an expectation operator. Instead, optimization is performed over a known explicit function in (20) where H is given by:

$$b(\xi^j,\pi)Y_{k+1}^j + \mathrm{tr}[\sigma'(\xi^j,\pi)Z_{k+1}^j] + f(t_k,\xi^j,\pi), \tag{26}$$

where $Y_{k+1}^j := Y_{k+1}^h(\xi^j)$. To be explicit, we consider the Newton method in a line search algorithm, see (Nocedal and Wright, 2006) for details. For fixed points in space and time k,j we guess a starting point $\pi^0 \in A$ and perform the following iteration:

$$\pi^{i+1} = \pi^i + p^i, \text{ for } i = 1,..,m. \tag{27}$$

A careful reader might notice that the step length is 1 here, which is enough in Newton's method to hold the Wolf conditions. In the exact Newton method, the search directions $p^i \in \mathbb{R}^r$ are calculated by solving the linear system:

$$\nabla_\pi^2 H(\pi^i)p^i = \nabla_\pi H(\pi^i). \tag{28}$$

Since H is smooth enough and the Hessian $\nabla_\pi^2 H(\pi^i)$ is positive definite - see above assumptions - the method converges to a local minimum and the rate of convergence of $\{\pi^i\}$ is quadratic. For a proof see (Nocedal and Wright, 2006), Theorem 3.2 and Theorem 3.5.

In detail, to solve the optimization in (20) for one point (t_k,ξ^j) we denote:

$$\begin{aligned} &H(\pi) := -H(t_k,\xi^j,\pi,Y_{k+1}^h(\xi^j),Z_{k+1}^h(\xi^j)) = \\ &-b(\xi^j,\pi)Y_{k+1}^j - \mathrm{tr}[\sigma'(\xi^j,\pi)Z_{k+1}^j] - f(t_k,\xi^j,\pi), \end{aligned} \tag{29}$$

and calculate:

$$H_\pi = -b_\pi Y^h_{k+1} - \sum_{i,j} \sigma^{ij}_\pi Z^{ij}_{k+1} - f_\pi,$$

$$H_{\pi\pi} = -b_{\pi\pi} Y^h_{k+1} - \sum_{i,j} \sigma^{ij}_{\pi\pi} Z^{ij}_{k+1} - f_{\pi\pi},$$

where we denote the gradients for all functions $g = b,\sigma, f, F$ by $g_\pi = \nabla_\pi g(\pi)$ and the Hessians by $g_{\pi\pi}$ accordingly. Note that $b_\pi, b_{\pi\pi}, \sigma^{ij}_\pi, \sigma^{ij}_{\pi\pi}, f_\pi, f_{\pi\pi}$ are known continuous functions, given by the problem definition. Then we solve the system:

$$H_{\pi\pi}(\pi^i)p^i = H_\pi(\pi^i),$$

and repeat this procedure until the error is smaller than a certain specified ε^h. If b,σ, f are too complex to derive the first or second derivatives by hand, one can use methods of 'automatic differentiation' to calculate them. An introduction into these methods can be found for example in (Nocedal and Wright, 2006).

4.4 Full algorithm

Recall that the the space grid at time point k is indexed by I_k in (15) and G is defined by (10). Therefore, the full algorithm reads as follows:

$\forall j \in I_N$,

$$\pi^{j*}_N = \arg\max_{\pi\in A} G(T, \xi^j, \pi, \nabla_x g(\xi^j), \nabla^2_{xx} g(\xi^j)), \quad (30)$$

$$V^h_N(\xi^j) = g(\xi^j), \quad Y^h_N = \nabla_x g(\xi^j), \\ Z^h_N = \nabla^2 g(\xi^j)\sigma(\xi^j, \pi^{j*}_N). \quad (31)$$

$\forall k = N-1..0$,

$\forall j \in I_k$,

set $\pi^{j0}_k = \pi^{j*}_{k-1}$. For $i = 1..i_{\max}$ solve:

$$H_{\pi\pi}(\pi^{ji}_k)p^i = H_\pi(\pi^{ji}_k), \\ \pi^{j,i+1}_k = \pi^{ji}_k + p^i, \quad (32)$$

until $||p^i|| < c\varepsilon^h$. Set $\pi^{j*}_k = \pi^{j,i+1}_k$.

$$V^h_k(\xi^j) = f(t_k, \xi^j, \pi^{j*}_k)\Delta t^h_k + \textstyle\sum_{\xi^l \in C_{k+1}} p^{jl}_k(\pi^{j*}_k) V^h_{k+1}(\xi^l), \quad (33)$$

$$Y^h_k(\xi^j) = \nabla_x H(t_k, \xi^j, \pi^{j*}_k, Y^h_{k+1}(\xi^j), Z^h_{k+1}(\xi^j))\Delta t^h_k + \textstyle\sum_{\xi^l \in C_{k+1}} p^{jl}_k(\pi^{j*}_k) Y^h_{k+1}(\xi^l), \quad (34)$$

$$Z^h_k(\xi^j) = \frac{1}{\Delta t^h_k} \sum_{\xi^l \in C_{k+1}} p^{jl}_k(\pi^{j*}_k) Y^h_{k+1}(\xi^l)\Delta W^{jl}_k, \quad (35)$$

where ΔW^{jl}_k is calculated via the Euler relationship:

$$\Delta W^{jl}_k = \sigma^{-1}(\xi^j, \pi^{j*}_k)(\xi^l - \xi^j - b(\xi^j, \pi^{j*}_k)\Delta t^h_k). \quad (36)$$

4.5 Known convergence results

For a specific subclass of stochastic control problems, the convergence of the proposed FBSDE scheme can be rigorously proved. Suppose that only the drift is controlled, i.e. $\sigma(x,\cdot) = \sigma(x)$. We also drop the time-dependence in the coefficients. Furthermore, suppose that

- $\sigma\sigma'$ is positive definite,
- $\nabla_x H(x, \pi^*(x,y,z), y, z)$ and $\nabla_x g(x)$ are Lipschitz continuous in x,
- $\nabla_x g \in C^{2+\alpha}(\mathbb{R}^r)$, for $\alpha > 0$, and its norm in bounded.

Note that $H = by + \mathrm{tr}[\sigma' z] + f$, is linear and Lipschitz continuous in y, z by definition.

Also remember that we have written the optimal control variable π^* as a function of the state variables x,y,z. Then, according to (Delarue and Menozzi, 2006), the numerical FBSDE scheme converges. We remark that a general proof of convergence for controlled diffusion processes, i.e. $\sigma(x,\pi)$, would be analogous to proving the existence and uniqueness of a fully nonlinear PDE. Such a proof is beyond the scope of this paper.

5 APPLICATIONS

5.1 An investment-consumption model

As an example, let us consider an investment-consumption model with convex transaction costs. Convex transaction costs preserve the problem from the usual bang-bang control, see (Davis and Norman, 1990). It is a reasonable assumption in certain markets, e.g. scarce commodities.

Let $A_t \in \mathbb{R}$ denote the portfolio owner's monetary amount of assets and let $\alpha_t, \beta_t > 0$ denote his investment and deinvestment rate at time t respectively. Let the convex functions $f^\beta, f^\alpha \in C^2(\mathbb{R},\mathbb{R})$ determine the transaction costs, respectively. Furthermore, let $B_t \in \mathbb{R}$ denote the portfolio owner's bank account, which pays a deterministic interest rate $r \geq 0$. Let the dynamics of the states be given by:

$$\begin{aligned} dA_t = &\ (\mu A_t + \alpha_t - \beta_t)dt + \sigma A_t dW_t, & A_0 = a_0, \\ dB_t = &\ (rB_t - c_t - \alpha_t + \beta_t)dt \\ &-(f^\alpha(\alpha_t) + f^\beta(\beta_t))dt, & B_0 = b_0, \end{aligned} \quad (37)$$

where $\mu, \sigma > 0$. W_t denotes a standard Brownian motion and $c_t \geq 0$ denotes the portfolio owner's consumption at time t.

The goal of the risk averse portfolio owner is to maximize his expected, concave utility $u \in C^2(\mathbb{R},\mathbb{R})$ from consumption over a given time horizon T:

$$J(t, a_0, b_0, \alpha, \beta, c) = E_0\left[\int_0^T e^{-\delta t} u(c_t)dt + e^{-\delta T} u(A_T + B_T) \Big| a_0,\, b_0\right], \quad (38)$$

$$v(t,a,b) = \max_{(\alpha,\beta,c)\in\mathbb{A}} J(,a,b,\alpha,\beta,c), \quad (39)$$

where $\mathbb{A}$ is the set of all F_t-measurable control strategies, $\delta \geq 0$ denotes the owner's impatience to consume and E_0 denotes the expectation operator with information set at time zero.

Using this problem setup, the Hamiltonian $H(t,x,y,z,\alpha_t,\beta_t,c_t)$ in (11) takes the following form:

$$\begin{aligned} H = \; & [\alpha_t - \beta_t + \mu A_t] y^A + \\ & + \left[rB_t - c_t - \alpha_t + \beta_t - \left(f^\alpha(\alpha_t) + f^\beta(\beta_t)\right)\right] y^B \\ & + (\sigma A_t) z^{AA} + e^{-\delta t} u(c_t). \end{aligned} \quad (40)$$

For the example calculations below, we choose log-utility and quadratic transaction costs:

$$u(x) = \ln(x),$$

$$f^\alpha(a) = c\,a^2, \quad f^\beta(b) = c\,b^2, \quad \text{for } c > 0.$$

Then, the adjoint equations become:

$$\begin{aligned} -dY_t^A &= \left(\mu Y_t^A + \sigma Z_t^{AA}\right) dt - Z_t^{AA} dW_t, & Y_T^A &= \tfrac{e^{-\delta T}}{A_T + B_T} \\ -dY_t^B &= r Y_t^B dt, & Y_T^B &= \tfrac{e^{-\delta T}}{A_T + B_T}. \end{aligned} \quad (41)$$

There are similar problems that have a strictly concave Hamiltonian with respect to the controls. One example is when the trading activities α, β influence the asset drift $\mu(\alpha,\beta)$, which is called feedback control.

5.2 Numerical results

To produce the results below, we implement the numerical algorithm with Matlab using parallel processing on an eight core machine.

We implemented the forward-backward (FB) algorithm of section 4.4 in two ways: using a Markov chain approximation and using a Quantization method. Since the latter method attained better results, we present the results of this method here. Therefore, let ΔW_i, $i = 1,\ldots,L$ denote the precalculated Brownian increments and let p_i denote their probabilities. For simplicity, we used a log-scaled, time-constant grid C. Then, for a fixed time point t_k, we calculated for each grid point $(A_k^j, B_k^j) \in$ C its propagation $A_{k+1}^{j,i} = A_k^j + \Delta A_t^{j,i}$, for each $i = 1\ldots L$, where:

$$\Delta A_t^i = (\mu A_t + \alpha_t - \beta_t)\Delta t + \sigma A_t p_i \sqrt{\Delta t}\Delta W_i. \quad (42)$$

In order to evaluate the function V, Y^A, Y^B, Z^{AA} at the points $(A_{k+1}^{j,i}, B_{k+1}^j)$ we used a linear interpolation/extrapolation.

In the Figures below, we used $M^2 = |I| = 100^2$ space points, $N = 100$ time steps, $L = 20$ quantization points, $C \in [0.2, 2.2] \times [0.2, 2.2]$ as space grid and $[t_0, T] = [0, 2]$ as time interval. Moreover, we set $\mu = 5\%$, $r = 3\%$, $\delta = 2\%$, $\sigma = 0.2$ and $c = 0.1\%$.

Figure 1 shows the consumption c_t for five different space points (A,B) over time. Figure 2 shows the surface of the combined control $\alpha^* - \beta^*$ at time step $k = 98$. The average calculation time for one time step was 2.6 seconds.

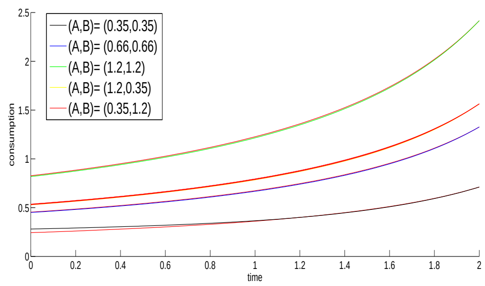

Figure 1: Optimal consumption c_t^* for five different space points, using the FB algorithm. Plots of the optimal consumption of the original Merton problem are added in red color.

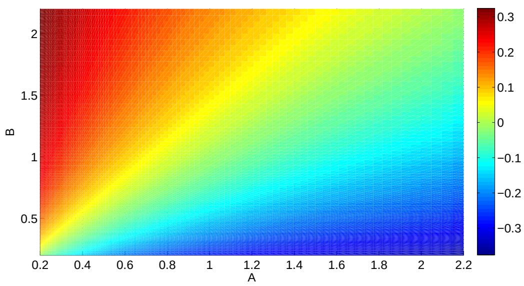

Figure 2: Optimal control $\alpha^* - \beta^*$ at time $t_k = 1.96$ using the FB algorithm.

The consumption in this problem is close to the consumption of the original Merton problem without transaction costs. A slight irregularity can be seen for $(A,B) = (0.35, 0.35)$ for small t. This shows the truncation error that propagates inside the space grid while going backward in time. The selected point is affected first since it is close to the lower space grid boundary (0.2,0.2). Since $\frac{\mu - r}{\sigma^2} = 0.5$, the optimal controls α^*, β^* should be zero whenever $A = B$, as in the original Merton problem.

To compare the results with the dynamic programming (DP) approach, we used the Matlab optimization toolbox. We obtained the best results with the provided Sequential Quadratic Programming (SQP) method, which is based on a line search quasi-Newton method. Details can be found at http://www.mathworks.com/help/toolbox/optim/. We needed to set $L = 40$ and used a spline interpolation method to get reasonable results at all.

Figure 2 show the surface of the combined, optimal control $\alpha^* - \beta^*$ at time step $k = 98$. The calculation time for one time step was 450 seconds.

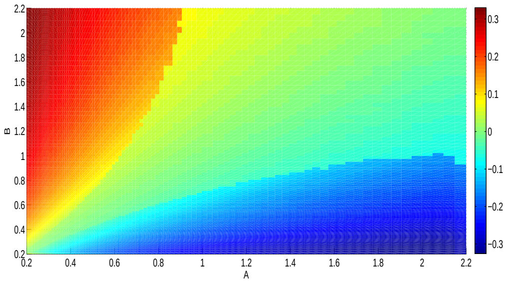

Figure 3: Optimal control $\alpha^* - \beta^*$ at time $t_k = 1.96$ using the DP method.

In this example, our FB algorithm is 170 times faster than the DP method. This is due to 1) the smaller amount of function evaluations and 2) the different interpolation method needed.

Moreover, the $\alpha^* - \beta^*$ surface of the FB method in Figure 2 is smooth, while the surface of the DP method in Figure 3 already has become rough at time step $k = 98$, indicating instability. One reason may be that in the DP method, the optimization is performed over the highly nonlinear value function V. In our FB algorithm, the optimization step depends on the functions Y_t and Z_t only.

5.3 Comparison of computational cost

We briefly compare our FB algorithm with the standard dynamic programming approach. Let $M^n = |I_l|$; that is, M is the number of grid points in each dimension of space where we have n dimensions. Also, let L be the number of calculated transition probabilities (or, alternatively, the number of quantization points/simulations if we use a Quantization/Monte Carlo method). N is the number of time-steps and let m be the number of iterations for the Newton-Raphson method to reach a given level of accuracy ε^h. Assuming that for a given level of accuracy, the same time grid and space grid may be used for both algorithms, the computational cost for the dynamic programming approach is $NM^nLm(1+2r^2)$ while the computational cost for the FB algorithm is $NM^n[L(1+n+nd)+mr]$. Assuming a large number of simulations L are needed for each evaluation, the FB algorithm is superior if:

$$2mr^2 > n(d+1)+1. \tag{43}$$

It is clear that the FB algorithm has a significantly lower computational cost if a nontrivial number of optimization iterations are required. For example, in one dimension the FB algorithm has $\frac{3}{2m}$ the computational cost of the dynamic programming approach. The advantage of the FB algorithm is that we do not need to optimize over the entire value function, which requires one to recalculate the expectation in the value function for at least m times. This is very computationally expensive. Instead, one only has to optimize the Hamiltonian, which is a much simpler procedure.

6 CONCLUSIONS

We have proposed a complete numerical algorithm to solve optimal control problems through the associated FBSDE system. By complete we mean that the algorithm explicitly includes the optimization step. Our numerical approach is an alternative to standard dynamic programming methods. A comparison of computational cost between the dynamic programming method and the FBSDE method illustrate the advantages of the FBSDE approach.

We included results of a numerical example that commonly appears in finance and economics. These results confirm the advantage in accuracy and computational efficiency of the FB algorithm compared to the dynamic programming method for certain problem classes.

A next step would be to analyze the convergence speed and the convergence error in theory and practice in detail.

REFERENCES

Davis, M. and Norman, A. (1990). Portfolio selection with transaction costs. In *Mathematics of Operation Research, Vol. 15, No. 4*. The lnstitute of Management Sciences/Operations Research Society of America.

Delarue, F. and Menozzi, S. (2006). A forward-backward stochastic algorithm for quasi-linear pdes. In *The Annals of Applied Probability, Vol. 16, No. 1, 140184*. Universite Paris.

Kushner, H. and Dupuis, P. (1992). *Numerical Methods for Stochastic Control Problems in Continuous Time*. Springer, London, 1st edition.

Nocedal, J. and Wright, S. (2006). *Numerical Optimization*. Springer Series in Operation Research and Financial Engineering, London, 2nd edition.

Peng, S. (1990). A general stochastic maximum principle for optimal control problems. In *Journal of Control and Optimization, Vol. 28, No. 4, pp.966-979*. SIAM.

Pham, H. (2009). Continuous-time stochastic control and optimization with financial applications. In *Stochastic Modeling and Applied Probability 61*. Springer-Verlag Berlin Heidelberg.