A Malliavin-based Monte-Carlo Approach for Numerical Solution of Stochastic Control Problems: Experiences from Merton's Problem

Simon Lysbjerg Hansen[1]

31st May 2005

[1]Department of Accounting and Finance, University of Southern Denmark, Campusvej 55, DK–5230 Odense M. E-mail: slh@sam.sdu.dk

Abstract

Using a simplified version of Merton's problem as a benchmark, a numerical procedure for solving stochastic control problems is developed. The algorithm involves the estimation of conditional expectations, which are conditioned on the controlled state process. Although Merton's problem can be reduced to not depend on the controlled state process the suggested method does not use this fact.

1 Introduction

The problem of choosing optimal investment and consumption strategies has been widely studied. In continuous time theory the pioneering work by Merton (1969) is a standard reference. In his work, Merton studied a continuous time economy with constant investment opportunities. Since then Merton's problem has been extended in many ways to capture empirically observed investment and consumption behavior. As more realism is incorporated into a model, the problem of optimal investment and consumption becomes harder to solve. Only rarely can analytical solutions be found, and only for problems possessing nice characteristics. To solve problems lacking analytical solutions we must apply numerical methods. Many realistic problems, however, are difficult to solve even numerically, due to their dimensionality. It was such a problem that motivated this paper. To be more specific, we were interested in the problem of choosing between fixed-rate mortgages and adjustable-rate mortgages. In a recent paper, Campbell and Cocco (2003) take a utility-based approach to analyze this problem in discrete time when inflation, labor income and interest rates are stochastic. Moreover, house prices and financial assets could be introduced to further complicate the problem, leaving a large number of state variables affecting the mortgage decision. This realistic problem lacks a closed-form solution and must be solved numerically.

Unfortunately, most existing numerical methods suffer from the curse of dimensionality, meaning that the time and space required to find a solution grow exponentially as the dimension increases. The so-called grid-based methods suffer from this curse. One such method is the Markov chain approximation approach developed by Kushner and described in the book by Kushner and Dupuis (2001). An application of the method to Merton's problem can be found in Munk (2003). The method approximates a continuous-time, continuous-state problem with a discrete-time, discrete-state Markov chain, which converges to the continuous-time problem as the discretization becomes finer. The state space is discretized with a grid, whose size grows exponentially as the number of state variables increase. Another grid-based method is the so-called quantization algorithm studied in Pagès, Pham, and Printems (2004) for multi-dimensional stochastic control problems. The idea in this method is to project a time-discretized version of a continuous-time stochastic process onto an "optimal" grid, in the sense that some error is minimized for the "optimal" grid. For many state variables these grids become large and computationally intractable.

Unlike the grid-based methods, the space and time required to solve a problem with Monte-Carlo methods only grows linearly in the number of state variables. Recently, Monte-Carlo methods have been introduced in the solution of stochastic control problems. The papers by Detemple, Garcia, and Rindisbacher (2003) and Cvitanić, Goukasian, and Zapatero (2003) exploit the martingale approach for a complete market to express the optimal investment strategies as (conditional) expectations which can be simulated. As noted earlier optimal mortgage choice depends on labor income, which cannot be hedged in the financial market. Hence markets are incomplete and these methods are inapplicable. Brandt, Goyal, Santa-Clara, and Stroud (2005) use the dynamic programming principle to solve an optimal portfolio problem recursively backwards. By approximating the value function with a fourth-order Taylor series expansion, they derive formulas which implicitly define the optimal controls. The formulas involve conditional expectations, which are estimated as suggested by Longstaff and Schwartz (2001) for the pricing of American options. Specifically, they regress ex-post simulated values of the stochastic variable in interest onto a set of basis functions, e.g. polynomials. Whereas the pricing of an American option involves a binary decision variable (exercise or not), optimal portfolio choice involves continuous decision variables. Because of this, an imprecise estimator has greater significance in the optimal portfolio choice problem, since it could still lead to the right choices in the binary choice problem. Longstaff (2001) applies the least-squares Monte-Carlo method to a continuous-time portfolio choice problem. However, it turns out that the control is binary, so an imprecise estimator could still lead to an optimal choice.

The purpose of this paper is to present a numerical procedure for solving high-dimensional stochastic control problems arising in the study of optimal portfolio choice. For expositional reasons we develop the algorithm in one dimension, but the mathematical results needed can be generalized to a multi-dimensional setting. The starting point of the algorithm is an initial guess about the agent's consumption strategy at all times and wealth levels. Given this guess it is possible to simulate the wealth process until the investment horizon of the agent. We exploit the dynamic programming principle to break the problem into a series of smaller one-period problems, which can be solved recursively backwards. To be specific we determine a first-order condition relating the optimal control to the value function in the next period. Starting from the end we now numerically solve this first-order condition for all simulated paths. Part of this computation involves the estimation of a conditional expectation, in which the wealth process is the conditioning variable. Therefore, these conditional expectations depend on the simulated distribution of wealth, which in turn depends on the initial guess about the consumption strategy. We can, however, use the consumption strategy resulting from the above backwards procedure to update the simulated wealth paths and repeat the procedure iteratively.

In an option pricing framework Fournié, Lasry, Lebuchoux, and Lions (2001) demonstrate how to compute conditional expectations with the use of Malliavin calculus. Bouchard, Ekeland, and Touzi (2004) generalize this result and discuss variance minimizing issues related to the Monte-Carlo simulated estimate of the conditional expectation. The idea in the papers is to express conditional expectations as a ratio between two unconditional expectations, which can be estimated by ordinary Monte-Carlo simulations of the conditioning variables. Strongly inspired by the latter paper, we use this approach to estimate the conditional expectations arising in the above mentioned problem. Bouchard, Ekeland, and Touzi also apply the approach to a dynamic portfolio problem, but the problem is reduced so the (controlled) wealth process is not a conditioning variable. For many interesting problems, like the optimal mortgage problem mentioned earlier, such simplifications cannot be made. This paper focuses on the issues related to conditioning on a controlled process. Specifically, the algorithm we suggest does not exploit the homogeneity property that exists in Merton's problem.

The numerical properties of the algorithm are analyzed by testing it on a simplified version of Merton's optimal portfolio choice problem. The reason for this is that the solution to Merton's problem is explicitly known and can therefore serve as a benchmark for the algorithm. Our results indicate that it is possible to obtain some sort of convergence in both the initial control and in the future distribution of the control. However, the results are obtained for a coarse time discretization, and numerical experiments indicate problems when the discretization is fine. Possible explanations of this problem are discussed and suggestions for improvement are made.

Bearing in mind that we intend to apply the algorithm to a multi-dimensional setting, we also consider the possible complications that might arise. However, the state variables added will in most cases be exogenous non-controllable processes, which does not complicate the optimization routine in the proposed algorithm. Problems with computer storage could arise, but they should be solvable with clever computer programming.

The rest of this paper is organized as follows. In section 2 we review Merton's problem. Section 3 discretize Merton's problem and describes the numerical method. In section 4 we apply the method to Merton's problem and discuss some issues related to the implementation. Finally, we conclude in section 5. Appendix A contains some result from Malliavin calculus that the algorithm in section 3 builds upon.

2 Merton's Problem Revisited

In this section we formulate Merton's problem and provide the closed-form solution. Readers familiar with Merton's problem can safely skip this section, perhaps after reviewing theorem 2.1. The problem consists in finding the optimal consumption and investment strategies for an economic agent with a finite investment horizon in a continuous time economy. This problem was first studied and solved in the pioneering paper by Merton (1969). In a discrete time economy Samuelson (1969) solved a similar problem. However, since the purpose of this paper is to test a numerical approach's ability to solve continuous time problems, we walk down the same road as Merton.

2.1 The Mathematical Problem

We consider an agent with initial wealth W_0 and investment horizon $[0,T]$, who wants to choose a consumption strategy, $(c_t)_{t\in[0,T]}$, and an investment strategy, $(\pi_t)_{t\in[0,T]}$, such that his expected lifetime utility is maximized. At time t, the agent consumes at the rate c_t and holds a fraction π_t of his wealth invested in a risky asset. The wealth dynamics of the agent is governed by the stochastic differential equation

$$\mathrm{d}W_t = W_t\left[r + \pi_t(\mu - r)\right]\mathrm{d}t - c_t\mathrm{d}t + W_t\pi_t\sigma\mathrm{d}B_t, \tag{2.1}$$

where r is the instantaneous risk free rate and μ and σ are the instantaneous drift rate and volatility of the risky asset price process, respectively. In Merton's problem these are all assumed constant, i.e. investment opportunities are constant. As a consequence, the risky asset price process is a geometric Brownian motion. Finally, $(B_t)_{t\in[0,T]}$ is a Brownian motion defined on a probability space $(\Omega, \mathcal{F}, P)$, i.e. $(B_t)_{t\in[0,T]}$ is a Brownian motion with respect to the probability measure P. We let the σ-algebra, $\mathcal{F}$, be the natural filtration, i.e. $\mathcal{F} \triangleq \mathcal{F}_T$ where $\mathcal{F}_t \triangleq \sigma(B_s \mid 0 \leq s \leq t)$ is the σ-algebra generated by the Brownian motion. The state space, Ω, will be specified in much more detail in appendix A.

To ensure that the integrals in (2.1) are well-defined the integrands (and hence the controls) must satisfy some integrability conditions, see e.g. Duffie (2001). Since we are going to apply the dynamic programming approach, the agent is restricted to choose the consumption and investment strategies from the class of Markov controls. Hence the consumption and investment strategies at time t only depend on the state of the system at that time, i.e. $c_t = C(W_t, t)$ and $\pi_t = \Pi(W_t, t)$ for some functions C and Π. Fortunately, this assumption is not crucial, since allowing the agent to choose from the larger class of $\mathcal{F}_t$-adapted controls does not lead to higher expected lifetime utility, as long as some regularity conditions are satisfied, cf. Øksendal (2000, Theorem 11.2.3). At last, for a consumption process to be admissible we will require it to assume positive values only, i.e. for fixed $t \in [0, T]$ the random variable $\omega \to c_t(\omega)$ must be a positive random variable. The investment process is allowed to take on any real values, i.e. short-sale of the risky asset is allowed.

The agent derives utility from intertemporal consumption and terminal wealth according to a utility function, u. In our examples u will always be a constant relative risk aversion (CRRA) utility function with constant relative risk aversion parameter γ, on the form

$$u(x) = \frac{x^{1-\gamma}}{1-\gamma}.$$

The agent chooses consumption and investment strategies in order to maximize his expected lifetime utility

$$\mathbb{E}\left\{\int_0^T e^{-\delta s} u(c_s)\mathrm{d}s + e^{-\delta T} u(W_T)\right\}.$$

Since Merton's problem is a dynamic problem we need to consider the agent's problem at a future time, t. To this purpose we define the indirect utility function (a.k.a. the value function) as

$$J(w,t) \triangleq \sup_{(c_s,\pi_s)_{s\in[t,T]}} \mathbb{E}\left\{\int_t^T e^{-\delta(s-t)} u(c_s)\mathrm{d}s + e^{-\delta(T-t)} u(W_T) \;\middle|\; W_t = w\right\}, \tag{2.2}$$

and the corresponding indirect utility process as $J_t \triangleq J(W_t, t)$. The indirect utility function is essential to the solution of Merton's problem as we shall see in the next section.

2.2 The Closed-Form Solution

A fundamental result from stochastic control theory states that the indirect utility function (2.2) must satisfy the Hamilton-Jacobi-Bellman (HJB) equation

$$\delta J(w,t) = \sup_{c\geq 0, \pi\in\mathbb{R}} \left\{ u(c) + \frac{\partial J}{\partial t}(w,t) + J_w(w,t)\left(w\left[r+\pi(\mu - r)\right] - c\right) + \frac{1}{2} J_{ww}(w,t) w^2 \pi^2 \sigma^2 \right\},$$

for all $(w,t) \in \mathbb{R}_+ \times [0,T)$ with the boundary condition $J(w,T) = u(w)$ for all $w \in \mathbb{R}_+$. Subscripts indicate partial derivatives. Essentially, this is a continuous time version of the dynamic programming principle, which states that an agent with wealth w at time t chooses a consumption rate c and a portfolio weight π, given optimal behavior at all future dates. We immediately get the following first-order conditions for the optimal strategies

$$\begin{aligned} 0 &= u'(c) - J_w(w,t) \\ 0 &= J_w(w,t) w(\mu - r) + J_{ww}(w,t) w^2 \pi \sigma^2. \end{aligned}$$

Isolating c and π, we get the following candidates for the optimal consumption and investment strategies

$$\begin{aligned} C(w,t) &= (u')^{-1}(J_w(w,t)) \\ \Pi(w,t) &= -\frac{J_w(w,t)}{J_{ww}(w,t)w} \frac{\mu - r}{\sigma^2}. \end{aligned}$$

Letting W_t^* denote the wealth process induced by following the optimal strategies, the optimal strategies at time t are

$$\begin{aligned} c_t^* &= C(W_t^*, t) \\ \pi_t^* &= \Pi(W_t^*, t). \end{aligned}$$

If the candidate optimal controls are substituted back into the HJB equation we need to solve a nonlinear partial differential equation in order to find J. Such equations are not easy to solve, but fortunately we are able to make a qualified guess and then verify that it actually solves the HJB equation.

Due to the linearity of the wealth dynamics, it seems reasonable to make the following conjecture. If an agent with time t wealth w optimally chooses to consume c_t^* and invest π_t^* in the risky asset and following these strategies imply a terminal wealth of W_T^* then an agent with time t wealth kw will optimally consume kc_t^* and invest π_t^* in the risky asset, implying a terminal wealth of kW_T^*. With CRRA utility we get

$$
\begin{aligned}
J(kw,t) &= \mathbb{E}\left\{\int_t^T e^{-\delta(s-t)}u(kc_s^*)\mathrm{d}s + e^{-\delta(T-t)}u(kW_T^*) \;\middle|\; W_t = kw\right\} \\
&= k^{1-\gamma}\mathbb{E}\left\{\int_t^T e^{-\delta(s-t)}u(c_s^*)\mathrm{d}s + e^{-\delta(T-t)}u(W_T^*) \;\middle|\; W_t = w\right\} \\
&= k^{1-\gamma}J(w,t),
\end{aligned}
$$

which for $k = \frac{1}{w}$ reads

$$
\begin{aligned}
J(w,t) &= J(1,t)w^{1-\gamma} && (2.3)\\
&= \frac{g(t)^{\gamma}w^{1-\gamma}}{1-\gamma}, && (2.4)
\end{aligned}
$$

where $g(t)^{\gamma} \triangleq (1-\gamma)J(1,t)$. In other words, $w \to J(w,t)$ is homogeneous of degree $1-\gamma$.

To verify our conjecture we check whether the HJB equation is satisfied or not. The relevant derivatives are

$$
\begin{aligned}
\frac{\partial J}{\partial t}(w,t) &= \frac{\gamma g(t)^{\gamma-1}g'(t)w^{1-\gamma}}{1-\gamma} \\
J_w(w,t) &= g(t)^{\gamma}w^{-\gamma} \\
J_{ww}(w,t) &= -\gamma g(t)^{\gamma}w^{-\gamma-1},
\end{aligned}
$$

which gives the following candidate consumption and investment strategies

$$
\begin{aligned}
C(w,t) &= (u')^{-1}(g(t)^{\gamma}w^{-\gamma}) \\
&= (u')^{-1}\left(u'\left(\frac{w}{g(t)}\right)\right) \\
&= \frac{w}{g(t)} \\
\Pi(w,t) &= \frac{g(t)^{\gamma}w^{-\gamma}}{\gamma g(t)^{\gamma}w^{-\gamma}}\frac{\mu-r}{\sigma^2} \\
&= \frac{1}{\gamma}\frac{\mu-r}{\sigma^2}.
\end{aligned}
$$

Substituting the candidate controls, the conjecture, and its derivatives into the HJB equation, we arrive at

$$\begin{aligned}
\delta\frac{g(t)^{\gamma}w^{1-\gamma}}{1-\gamma} &= \frac{g(t)^{\gamma-1}w^{1-\gamma}}{1-\gamma} + \frac{\gamma g(t)^{\gamma-1}g'(t)w^{1-\gamma}}{1-\gamma} \\
&\quad + g(t)^{\gamma}w^{1-\gamma}\left(\left[r + \frac{1}{\gamma}\frac{(\mu-r)^2}{\sigma^2}\right] - \frac{1}{g(t)}\right) - \frac{1}{2}\gamma g(t)^{\gamma}w^{1-\gamma}\frac{1}{\gamma^2}\frac{(\mu-r)^2}{\sigma^2} \\
&= \frac{g(t)^{\gamma-1}w^{1-\gamma}}{1-\gamma}\left(\gamma g'(t) + g(t)\left(r(1-\gamma) + \frac{1}{2}\frac{1-\gamma}{\gamma}\frac{(\mu-r)^2}{\sigma^2}\right) + \gamma\right).
\end{aligned}$$

Moving everything to the right-hand side and collecting terms with common factors, we obtain

$$0 = \frac{g(t)^{\gamma-1}w^{1-\gamma}}{1-\gamma}\left(\gamma g'(t) + g(t)\left(-\delta + r(1-\gamma) + \frac{1}{2}\frac{1-\gamma}{\gamma}\frac{(\mu-r)^2}{\sigma^2}\right) + \gamma\right).$$

For this equation to be satisfied for all t and w, the expression in parentheses must be equal to zero, i.e.

$$\gamma g'(t) = g(t)\left(\delta - r(1-\gamma) - \frac{1}{2}\frac{1-\gamma}{\gamma}\frac{(\mu-r)^2}{\sigma^2}\right) - \gamma.$$

Hence g must satisfy the ordinary differential equation

$$\begin{aligned}
g'(t) &= g(t)\left(\frac{\delta - r(1-\gamma)}{\gamma} - \frac{1}{2}\frac{1-\gamma}{\gamma^2}\frac{(\mu-r)^2}{\sigma^2}\right) - 1 \\
&= Ag(t) - 1,
\end{aligned}$$

where we have defined

$$A \triangleq \left(\frac{\delta - r(1-\gamma)}{\gamma} - \frac{1}{2}\frac{1-\gamma}{\gamma^2}\frac{(\mu-r)^2}{\sigma^2}\right). \tag{2.5}$$

Imposing the boundary condition $g(T) = [(1-\gamma)u(1)]^{1/\gamma} = 1$, the solution is

$$g(t) = A^{-1}\left(1 + [A-1]\,e^{-A(T-t)}\right).$$

With this g our conjecture in equation (2.4) satisfies the HJB equation. Furthermore, it satisfies some technical conditions for it to equal the indirect utility function. We have now justified the following theorem, which was first proved by Merton (1969).

Theorem 2.1. *With*

$$g(t) \triangleq A^{-1}\left(1 + [A-1]\,e^{-A(T-t)}\right)$$

the indirect utility function is given by

$$\begin{aligned} J(w,t) &= \frac{g(t)^\gamma w^{1-\gamma}}{1-\gamma} \\ &= g(t)^\gamma u(w). \end{aligned}$$

The optimal investment strategy is given by

$$\Pi(w,t) = \frac{1}{\gamma}\frac{\mu - r}{\sigma^2},$$

and the optimal consumption rate is given by

$$\begin{aligned} C(w,t) &= \frac{1}{g(t)} w \\ &= A\left(1 + [A-1]\, e^{-A(T-t)}\right)^{-1} w. \end{aligned}$$

Remark. The homogeneity property in equation (2.3) reduces the dimensionality of the problem by one: instead of solving a second-order partial differential equation, the problem is reduced to finding the solution of a much simpler ordinary differential equation. In some multi-dimensional models similar homogeneity properties lead to a similar reduction in the dimensionality of the problem. In general, however, the value function does not possess such homogeneity properties, and the approach we develop does not exploit such. Of course, this seems foolish, but bear in mind that the algorithm will be applied to a more general problem.

3 The Numerical Method

Merton's problem can be solved analytically, whereas multi-dimensional problems in general cannot. With this in mind we now consider a numerical approach to solving Merton's problem, which can be tested up against the explicitly known solution. Hopefully, the numerical procedure is applicable for multi-dimensional problems as well. To simplify the problem as much as possible, we only allow the agent to control the consumption rate, i.e. we fix the portfolio weight at the (constant) Merton solution. By doing this, future wealth is still stochastic and the control variable is both time and state dependent.

3.1 A Discrete-Time Approximation

Since computers in their nature work discretely, we cannot feed them with a continuous time problem like the one in equation (2.2) subject to the wealth dynamics in equation (2.1). The problem must be discretized. We therefore partition the time horizon $[0,T]$ into N intervals of equal length $\Delta t \triangleq \frac{T}{N}$ and put $t_n \triangleq n\Delta t$ for $n = 0, 1, \ldots, N$. Approximating the stochastic differential equation (2.1) with an Euler discretization

$$W_{t_{n+1}} = W_{t_n} + W_{t_n}\left[r + \pi(\mu - r)\right]\Delta t - c_{t_n}\Delta t + W_{t_n}\pi\sigma(B_{t_{n+1}} - B_{t_n}), \tag{3.1}$$

we now consider the discrete time problem of choosing a consumption strategy $(c_{t_n})_{n=0}^{N-1}$ such that the discretized objective

$$\mathbb{E}\left\{\sum_{n=0}^{N-1} e^{-\delta t_n} u(c_{t_n})\Delta t + e^{-\delta T} u(W_T)\right\} \tag{3.2}$$

is maximized. Instead of using an Euler discretization we could consider a true discrete time wealth dynamics. However, the purpose of this paper is to test a numerical approach's ability to solve continuous time problems, which should be possible with the Euler dynamics, since it converges to the continuous time wealth dynamics when $\Delta t \to 0$.

The maximization of the objective in equation (3.2) involves the entire consumption process. As the following computation demonstrates, the problem can be decomposed into a series of one-period maximization problems. The indirect utility process at time t_n is

$$\begin{aligned}
J_{t_n} &= \sup_{(c_{t_i})_{i=n}^{N-1}} \mathbb{E}\left\{\sum_{i=n}^{N-1} e^{-\delta(t_i - t_n)} u(c_{t_i})\Delta t + e^{-\delta(T - t_n)} u(W_T) \;\middle|\; \mathcal{F}_{t_n}\right\} \\
&= \sup_{(c_{t_i})_{i=n}^{N-1}} \mathbb{E}\left\{\sum_{i=n}^{N-1} e^{-\delta(i-n)\Delta t} u(c_{t_i})\Delta t + e^{-\delta(N-n)\Delta t} u(W_T) \;\middle|\; \mathcal{F}_{t_n}\right\} \\
&= \sup_{(c_{t_i})_{i=n}^{N-1}} \mathbb{E}\left\{u(c_{t_n})\Delta t + \sum_{i=n+1}^{N-1} e^{-\delta(i-n)\Delta t} u(c_{t_i})\Delta t + e^{-\delta(N-n)\Delta t} u(W_T) \;\middle|\; \mathcal{F}_{t_n}\right\} \\
&= \sup_{(c_{t_i})_{i=n}^{N-1}} \mathbb{E}\Bigg\{u(c_{t_n})\Delta t \\
&\qquad + \mathbb{E}\left\{\sum_{i=n+1}^{N-1} e^{-\delta(i-n)\Delta t} u(c_{t_i})\Delta t + e^{-\delta(N-n)\Delta t} u(W_T) \;\middle|\; \mathcal{F}_{t_{n+1}}\right\} \;\Bigg|\; \mathcal{F}_{t_n}\Bigg\} \\
&= \sup_{c_{t_n}} \mathbb{E}\left\{u(c_{t_n})\Delta t + e^{-\delta\Delta t} J_{t_{n+1}} \mid \mathcal{F}_{t_n}\right\} \\
&= \sup_{c_{t_n}} \left\{u(c_{t_n})\Delta t + e^{-\delta\Delta t}\mathbb{E}\left\{J_{t_{n+1}} \mid \mathcal{F}_{t_n}\right\}\right\},
\end{aligned}$$

and the associated indirect utility function is

$$J(w, t_n) = \sup_{c_{t_n}} \left\{u(c_{t_n})\Delta t + e^{-\delta\Delta t}\mathbb{E}\left\{J(W_{t_{n+1}}, t_{n+1}) \mid W_{t_n} = w\right\}\right\}. \tag{3.3}$$

This is the so-called Bellman equation, which is the building block in backward recursive solutions of dynamic programming problems. The HJB equation can also be seen as a limit of the Bellman equation when $\Delta t \to 0$.

3.2 First-Order Conditions

In this section we derive the formulas that we will use in the numerical approach. At each time t_n we need to find the optimal consumption strategy. Inserting the Euler dynamics (3.1) in the Bellman equation (3.3) we have

$$
\begin{aligned}
J(w, t_n) &= \sup_{c_{t_n}} \Big\{ u(c_{t_n})\Delta t + e^{-\delta \Delta t} \mathbb{E} \left\{ J(W_{t_{n+1}}, t_{n+1}) \mid W_{t_n} = w \right\} \Big\} \\
&= \sup_{c_{t_n}} \Big\{ u(c_{t_n})\Delta t \\
&\qquad + e^{-\delta \Delta t} \mathbb{E} \left\{ J(w + w\left[r + \pi(\mu - r)\right]\Delta t - c_{t_n}\Delta t + w\pi\sigma(B_{t_{n+1}} - B_{t_n}), t_{n+1}) \right\} \Big\}
\end{aligned}
$$

The first-order condition with respect to c_{t_n} is

$$
\begin{aligned}
0 &= u'(c_{t_n})\Delta t \\
&\quad + e^{-\delta \Delta t} \mathbb{E} \left\{ J_w(w + w\left[r + \pi(\mu - r)\right]\Delta t - c_{t_n}\Delta t + w\pi\sigma(B_{t_{n+1}} - B_{t_n}), t_{n+1})(-\Delta t) \right\} \\
&= u'(c_{t_n})\Delta t - e^{-\delta \Delta t} \mathbb{E} \left\{ J_w(W_{t_{n+1}}, t_{n+1}) \mid W_{t_n} = w \right\} \Delta t \\
&= \Big[u'(c_{t_n}) - e^{-\delta \Delta t} \mathbb{E} \left\{ J_w(W_{t_{n+1}}, t_{n+1}) \mid W_{t_n} = w \right\} \Big] \Delta t,
\end{aligned}
$$

i.e.

$$
u'(c_{t_n}) = e^{-\delta \Delta t} \mathbb{E} \left\{ J_w(W_{t_{n+1}}, t_{n+1}) \mid W_{t_n} = w \right\}. \tag{3.4}
$$

It is equation (3.4) that will be the main driver in our numerical approach. Since the indirect utility function is known at time $t_N = T$, the problem can be solved recursively backwards by the dynamic programming approach. Because $W_{t_{n+1}}$ depends on c_{t_n} equation (3.4) only defines the consumption rate implicitly. Later we shall discuss how to find the optimal consumption rate, but for the moment we just assume that an optimal consumption strategy exists and denote it by $c^*_{t_n}$. Inserting the optimal consumption rate into the Bellman equation (3.3), the indirect utility at time t_n is

$$
J(w, t_n) = u(c^*_{t_n})\Delta t + e^{-\delta \Delta t} \mathbb{E} \Big\{ J(W^*_{t_{n+1}}, t_{n+1}) \mid W_{t_n} = w \Big\},
$$

where

$$
W^*_{t_{n+1}} = W_{t_n} + W_{t_n}\left[r + \pi(\mu - r)\right]\Delta t - c^*_{t_n}\Delta t + W_{t_n}\pi\sigma(B_{t_{n+1}} - B_{t_n}).
$$

In order to find the optimal consumption strategy at time t_{n-1} we need the partial derivative of the indirect utility function with respect to wealth at time t_n, which is

$$
\begin{aligned}
J_w(w, t_n) &= u'(c^*_{t_n})\frac{\partial c^*_{t_n}}{\partial w}\Delta t + e^{-\delta\Delta t}\mathbb{E}\left\{ J_w(W^*_{t_{n+1}}, t_{n+1})\frac{\partial W^*_{t_{n+1}}}{\partial w} \;\middle|\; W_{t_n} = w \right\} \\
&= u'(c^*_{t_n})\frac{\partial c^*_{t_n}}{\partial w}\Delta t + e^{-\delta\Delta t}\mathbb{E}\Bigg\{ J_w(W^*_{t_{n+1}}, t_{n+1})\Bigg(1 + [r + \pi(\mu - r)]\,\Delta t \\
&\qquad\qquad\qquad\qquad - \frac{\partial c^*_{t_n}}{\partial w}\Delta t + \pi\sigma(B_{t_{n+1}} - B_{t_n})\Bigg) \;\Bigg|\; W_{t_n} = w \Bigg\} \\
&= e^{-\delta\Delta t}\mathbb{E}\left\{ J_w(W^*_{t_{n+1}}, t_{n+1})\left(1 + [r + \pi(\mu - r)]\,\Delta t + \pi\sigma(B_{t_{n+1}} - B_{t_n})\right) \;\middle|\; W_{t_n} = w \right\} \\
&= e^{-\delta\Delta t}\mathbb{E}\left\{ J_w(W^*_{t_{n+1}}, t_{n+1})\frac{W^*_{t_{n+1}} + c^*_{t_n}\Delta t}{W_{t_n}} \;\middle|\; W_{t_n} = w \right\}
\end{aligned}
$$

where the third equality follows since $c^*_{t_n}$ satisfies the first-order condition in equation (3.4). Multiplying both sides by w we see that the marginal indirect utility function, J_w, must satisfy

$$
\begin{aligned}
wJ_w(w, t_n) &= e^{-\delta\Delta t}\mathbb{E}\left\{ J_w(W^*_{t_{n+1}}, t_{n+1})W^*_{t_{n+1}} \mid W_{t_n} = w \right\} \\
&\quad + e^{-\delta\Delta t}\mathbb{E}\left\{ J_w(W^*_{t_{n+1}}, t_{n+1}) \mid W_{t_n} = w \right\} c^*_{t_n}\Delta t \\
&= e^{-\delta\Delta t}\mathbb{E}\left\{ J_w(W^*_{t_{n+1}}, t_{n+1})W^*_{t_{n+1}} \mid W_{t_n} = w \right\} + u'(c^*_{t_n})c^*_{t_n}\Delta t. \qquad (3.5)
\end{aligned}
$$

Equations (3.4) and (3.5) are the main ingredients in the backward recursive algorithm presented in section 3.4.

3.3 Computing Conditional Expectations

In order to implement the formulas we need a way to compute conditional expectations for all points $(w, t) \in \mathbb{R}_+ \times \{t_0, t_1, \ldots, t_N\}$. Over one period such expectations could be estimated with Monte-Carlo simulation, i.e. given M simulations, $W^1_{t_{n+1}}, W^2_{t_{n+1}}, \ldots, W^M_{t_{n+1}}$, of the random variable $W_{t_{n+1}}$ we have

$$
\mathbb{E}\left\{ g(W_{t_{n+1}}) \mid W_{t_n} = w \right\} \approx \frac{1}{M}\sum_{m=1}^{M} g(W^m_{t_{n+1}}).
$$

In a multi-period model where we need to know such conditional expectations for $n = 0, 1, \ldots, N-1$, we could try to do the same. Initially, we simulate M values of the random variable $W^w_{t_1}$, where the superscript indicates the value of the process at the beginning of the period. For each of these simulations we simulate M values of the random variable $W^{W^m_{t_1}}_{t_2}$, and hence have M^2 simulations of W_{t_2}. With N periods there will be M^N simulations of W_T. Not much imagination is needed to see that this requires many computations. Hence this type of Monte-Carlo simulation seems to be a computationally infeasible task. The following result emanating from Malliavin calculus provides an alternative representation of conditional expectations, which at all points, $W^m_{t_n}$, can be estimated from one set of simulations.

Theorem 3.1 (Bouchard, Ekeland, and Touzi (2004), Corollary 3.1)**.** *Let h_t^n satisfy*

$$\int_0^{t_{n+1}} (D_t W_{t_n}) h_t^n \mathrm{d}t = 1 \quad \textit{and} \quad \int_0^{t_{n+1}} (D_t W_{t_{n+1}}) h_t^n \mathrm{d}t = 0, \tag{3.6}$$

where D_t is the Malliavin derivative. Further, let $\varphi : \mathbb{R} \to \mathbb{R}$ be a smooth localizing function *in the sense that φ and φ' are continuous and bounded mappings and $\varphi(0) = 1$. Then*

$$\mathbb{E}\left\{g(W_{t_{n+1}}) \mid W_{t_n} = w\right\} = \frac{\mathbb{E}\left\{H_w(W_{t_n}) g(W_{t_{n+1}}) S^{h^n}(\varphi(W_{t_n} - w))\right\}}{\mathbb{E}\left\{H_w(W_{t_n}) S^{h^n}(\varphi(W_{t_n} - w))\right\}}, \tag{3.7}$$

where $H_w(W_{t_n}) = \mathbf{1}_{[w,\infty)}(W_{t_n})$ is the indicator function and $S^{h^n}(\varphi(W_{t_n} - w))$ is the Skorohod integral $\int_0^{t_{n+1}} \varphi(W_{t_n} - w) h_t^n \delta B_t$.

Proof. See appendix A.4. ☐

Remark. The localizing functions in the numerator and denominator of (3.7) need not be the same, which we will use in our implementation to minimize the var.

3.3.1 Monte-Carlo Simulation of Conditional Expectations

It is theorem 3.1 that tells us how to compute conditional expectations. Given M simulated paths of $(W_{t_n})_{n=0}^N$ all starting at the same initial point, W_0, we have

$$\begin{aligned} \mathbb{E}\left\{g(W_{t_{n+1}}) \mid W_{t_n} = w\right\} &= \frac{\mathbb{E}\left\{H_w(W_{t_n}) g(W_{t_{n+1}}) S^{h^n}(\varphi(W_{t_n} - w))\right\}}{\mathbb{E}\left\{H_w(W_{t_n}) S^{h^n}(\varphi(W_{t_n} - w))\right\}} \\ &\approx \frac{\frac{1}{M}\sum_{m=1}^M H_w(W_{t_n}^m) g(W_{t_{n+1}}^m) S^{h^n}(\varphi(W_{t_n}^m - w))}{\frac{1}{M}\sum_{m=1}^M H_w(W_{t_n}^m) S^{h^n}(\varphi(W_{t_n}^m - w))}. \end{aligned} \tag{3.8}$$

Remark. Notice that the Skorohod integral does not depend on g. Hence we only need to compute them once for each point (n, m) in our simulation, even though we need to compute several conditional expectations for different g's.

3.3.2 Optimal Localizing Functions

A problem related to Monte-Carlo simulations is the relatively high number of simulations required to get a precise estimate, i.e. small confidence intervals. In fact, the accuracy of a simulation is only increased by a factor $\sqrt{k}$ when the number of simulations is k-doubled. Another way to improve Monte-Carlo simulation is to use so-called variance-reduction techniques, which can speed up the simulation as they require fewer simulations. Examples of such variance reduction techniques are to use antithetic variables and control variates. As in Bouchard, Ekeland, and Touzi (2004) we here consider another type of variance reduction: localizing functions. The idea of such functions is that paths closer to the conditioning value, w, are weighted heavier than paths far away from w.

The variance of the Monte-Carlo estimator

$$\mathbb{E}\left\{H_w(W_{t_n})f(W_{t_{n+1}})S^{h^n}(\varphi(W_{t_n}-w))\right\} \approx \frac{1}{M}\sum_{m=1}^{M} H_w(W_{t_n}^m)f(W_{t_{n+1}}^m)S^{h^n}(\varphi(W_{t_n}^m-w))$$

is

$$\begin{aligned}
&\mathbb{V}\left\{\frac{1}{M}\sum_{m=1}^{M} H_w(W_{t_n}^m)f(W_{t_{n+1}}^m)S^{h^n}(\varphi(W_{t_n}^m-w))\right\}\\
&= \frac{1}{M^2}\sum_{m=1}^{M}\mathbb{V}\left\{H_w(W_{t_n}^m)f(W_{t_{n+1}}^m)S^{h^n}(\varphi(W_{t_n}^m-w))\right\}\\
&= \frac{1}{M}\mathbb{V}\left\{H_w(W_{t_n})f(W_{t_{n+1}})S^{h^n}(\varphi(W_{t_n}-w))\right\}\\
&= \frac{1}{M}\mathbb{E}\left\{H_w(W_{t_n})f(W_{t_{n+1}})^2S^{h^n}(\varphi(W_{t_n}-w))^2\right\}\\
&\quad -\frac{1}{M}\mathbb{E}\left\{H_w(W_{t_n})f(W_{t_{n+1}})S^{h^n}(\varphi(W_{t_n}-w))\right\}^2
\end{aligned}$$

since the $W_{t_n}^m$'s are independent and identically distributed random variables.

We now consider the problem of minimizing the mean square error

$$I^{h^n}[f](\varphi) \triangleq \int_{\mathbb{R}}\mathbb{E}\left\{H_w(W_{t_n})f(W_{t_{n+1}})^2S^{h^n}(\varphi(W_{t_n}-w))^2\right\}\mathrm{d}w \tag{3.9}$$

for exponential localizing functions $\varphi(x)=e^{-\eta x}$, $\eta \geq 0$.

Remark. Due to the indicator function, we are only considering $W_{t_n} \geq w$. On this set $I^{h^n}[f]$ is convex as the following computation shows

$$\begin{aligned}
&I^{h^n}[f](\lambda\varphi+(1-\lambda)\psi)\\
&\quad= \int_{\mathbb{R}}\mathbb{E}\left\{H_w(W_{t_n})f(W_{t_{n+1}})^2\left(\lambda S^{h^n}(\varphi(W_{t_n}-w))+(1-\lambda)S^{h^n}(\psi(W_{t_n}-w))\right)^2\right\}\mathrm{d}w\\
&\quad\leq \int_{\mathbb{R}}\mathbb{E}\left\{H_w(W_{t_n})f(W_{t_{n+1}})^2\lambda S^{h^n}(\varphi(W_{t_n}-w))^2\right\}\mathrm{d}w\\
&\qquad+\int_{\mathbb{R}}\mathbb{E}\left\{H_w(W_{t_n})f(W_{t_{n+1}})^2(1-\lambda)S^{h^n}(\psi(W_{t_n}-w))^2\right\}\mathrm{d}w\\
&\quad= \lambda I^{h^n}[f](\varphi)+(1-\lambda)I^{h^n}[f](\psi),
\end{aligned}$$

since the Skorohod integral is linear and $x \to x^2$ is convex.

Remark. Bouchard, Ekeland, and Touzi (2004) show that the optimal localizing function in the class of *separable* localizing function will be of the exponential form in a d-dimensional setting. In general they prove an existence and uniqueness result, and give a partial differential equation characterization in the 2-dimensional case.

Since φ is parameterized by η, minimizing the mean square error in equation (3.9) simply reduces to an ordinary minimization problem over η

$$\inf_{\eta\geq 0}\int_{\mathbb{R}}\mathbb{E}\left\{H_w(W_{t_n})f(W_{t_{n+1}})^2S^h(\varphi(W_{t_n}-w))^2\right\}\mathrm{d}w.$$

Since φ is also convex, the mean square error is convex in the parameter η. Moreover,

$$\begin{aligned}S^h(\varphi(W_{t_n}-w)) &= \varphi(W_{t_n}-w)S^h(1)-\varphi'(W_{t_n}-w)\\ &= \varphi(W_{t_n}-w)\left[S^h(1)+\eta\right],\end{aligned}$$

so we get

$$\begin{aligned}&\int_{\mathbb{R}}H_w(W_{t_n})f(W_{t_{n+1}})^2S^h(\varphi(W_{t_n}-w))^2\mathrm{d}w\\ &\quad=\int_{-\infty}^{W_{t_n}}f(W_{t_{n+1}})^2\varphi(W_{t_n}-w)^2\left[S^h(1)+\eta\right]^2\mathrm{d}w\\ &\quad=f(W_{t_{n+1}})^2\int_{-\infty}^{W_{t_n}}\varphi(W_{t_n}-w)^2\mathrm{d}w\left[S^h(1)+\eta\right]^2\\ &\quad=f(W_{t_{n+1}})^2\int_{-\infty}^{W_{t_n}}e^{-2\eta(W_{t_n}-w)}\mathrm{d}w\left[S^h(1)+\eta\right]^2\\ &\quad=f(W_{t_{n+1}})^2\left[\frac{1}{2\eta}e^{-2\eta(W_{t_n}-w)}\right]_{-\infty}^{W_{t_n}}\left[S^h(1)+\eta\right]^2\\ &\quad=f(W_{t_{n+1}})^2\frac{1}{2\eta}\left[S^h(1)+\eta\right]^2\\ &\quad=f(W_{t_{n+1}})^2\left[\frac{S^h(1)^2}{2\eta}+S^h(1)+\frac{\eta}{2}\right].\end{aligned}$$

Taking expectations on both sides, and differentiating with respect to η we get the following first-order condition for a minimum

$$\mathbb{E}\left\{f(W_{t_{n+1}})^2\left[-\frac{S^h(1)^2}{2\eta^2}+\frac{1}{2}\right]\right\}=0,$$

i.e.

$$\mathbb{E}\left\{f(W_{t_{n+1}})^2\left[-S^h(1)^2+\eta^2\right]\right\}=0,$$

which reduces to

$$\eta^2=\frac{\mathbb{E}\left\{f(W_{t_{n+1}})^2S^h(1)^2\right\}}{\mathbb{E}\left\{f(W_{t_{n+1}})^2\right\}}.$$

Remark. As noted earlier we can choose different localizing functions in the nominator and denominator of

$$\mathbb{E}\left\{g(W_{t_{n+1}}) \mid W_{t_n} = w\right\} = \frac{\mathbb{E}\left\{H_w(W_{t_n})g(W_{t_{n+1}})S^h(\varphi(W_{t_n} - w))\right\}}{\mathbb{E}\left\{H_w(W_{t_n})S^h(\varphi(W_{t_n} - w))\right\}}.$$

With exponential localizing functions this means that we can choose two η's to minimize the mean square error of both expectations on the right hand side.

3.3.3 Control Variate

A common way to reduce the error induced by a numerical method is to adjust the estimate with the error the method gives on a similar problem, for which the solution is known. This is called a *control variate* technique. By definition of the Skorohod integral we have

$$\begin{aligned}
\mathbb{E}\left\{g(W_{t_{n+1}})S^{h^n}(\varphi(W_{t_n} - w))\right\} &= \mathbb{E}\left\{\int_0^T (D_t g(W_{t_{n+1}}))\varphi(W_{t_n} - w)h_t^n \mathrm{d}t\right\} \\
&= \mathbb{E}\left\{g'(W_{t_{n+1}})\varphi(W_{t_n} - w)\int_0^T (D_t W_{t_{n+1}})h_t^n \mathrm{d}t\right\} \\
&= 0
\end{aligned}$$

since $h^n \in \mathbb{H}_n$. Hence, instead of estimating

$$\mathbb{E}\left\{H_w(W_{t_n})g(W_{t_{n+1}})S^{h^n}(\varphi(W_{t_n} - w))\right\},$$

we might as well estimate

$$\begin{aligned}
&\mathbb{E}\left\{H_w(W_{t_n})g(W_{t_{n+1}})S^{h^n}(\varphi(W_{t_n} - w))\right\} - 0 \\
&\quad = \mathbb{E}\left\{(H_w(W_{t_n}) - c)g(W_{t_{n+1}})S^{h^n}(\varphi(W_{t_n} - w))\right\},
\end{aligned}$$

for some $c \in \mathbb{R}$. If we choose c in order to minimize the variance of the Monte-Carlo estimator, we get

$$c = \frac{\mathbb{E}\left\{H_w(W_{t_n})g(W_{t_{n+1}})^2 S^{h^n}(\varphi(W_{t_n} - w))^2\right\}}{\mathbb{E}\left\{g(W_{t_{n+1}})^2 S^{h^n}(\varphi(W_{t_n} - w))^2\right\}}.$$

Remark. Like Bouchard, Ekeland, and Touzi (2004) our numerical tests indicate that the control variate has little effect on our results, and we will therefore not use it in our implementation.

3.4 The Algorithm

The basic ingredient in the algorithm is M simulated paths of the wealth process. To begin with we assume that we can find an initial feasible consumption strategy, $(c_{t_n}^{m,0})_{n=0}^{N-1}$, for all simulated paths $m = 1, 2, \ldots, M$. Later we will discuss how to initialize this solution. Given this strategy we are now able to simulate wealth according to the Euler dynamics in equation (3.1), i.e.

$$\begin{aligned} W_0 &= 0 \\ W_{t_{n+1}} &= W_{t_n} + W_{t_n}\left[r + \pi(\mu - r)\right]\Delta t - c_{t_n}\Delta t + W_{t_n}\pi\sigma(B_{t_{n+1}} - B_{t_n}) \\ &= W_{t_n} + W_{t_n}\left[r + \pi(\mu - r)\right]\Delta t - c_{t_n}\Delta t + W_{t_n}\pi\sigma\sqrt{\Delta t}Z_{t_{n+1}}, \end{aligned}$$

where $Z_{t_{n+1}} \sim \mathcal{N}(0,1)$ is a standard normally distributed random variable.

We will use the notation $W_{t_n}^m$ to denote wealth at time t_n on the m'th simulated path. The algorithm proceeds as follows:

- At time $t_N = T$ we impose the boundary condition $J_T(w) = u(w)$ to calculate the derivative $J'_T(W_T^m) = u'(W_T^m)$ for all $m = 1, 2, \ldots, M$.

- At time t_n, $n = N-1, N-2, \ldots, 1$ we iteratively solve the first-order conditions estimating the conditional expectations with the Malliavin Monte-Carlo approach, i.e. for all $m = 1, 2, \ldots, M$ we update the control iteratively until the consumption rate in the i'th iteration satisfies

$$\begin{aligned} &u'(c_{t_n}^{m,i}) \\ &\quad= e^{-\delta\Delta t}\mathbb{E}\left\{J_w(W_{t_{n+1}}, t_{n+1}) \mid W_{t_n} = W_{t_n}^m\right\} \\ &\quad= e^{-\delta\Delta t}\mathbb{E}\left\{J_w(W_{t_n}^m + W_{t_n}^m[r + \pi(\mu - r)]\Delta t - c_{t_n}^{m,i-1}\Delta t + W_{t_n}^m\pi\sigma(B_{t_{n+1}} - B_{t_n}), t_{n+1})\right\} \\ &\quad\approx e^{-\delta\Delta t}\frac{\sum_{l=1}^M H_w(W_{t_n}^l)J_w(W_{t_{n+1}}^l, t_{n+1})S^{h^{n,l}}(\varphi(W_{t_n}^l - W_{t_n}^m))}{\sum_{l=1}^M H_w(W_{t_n}^l)S^{h^{n,l}}(\varphi(W_{t_n}^l - W_{t_n}^m))} \end{aligned}$$

  at some pre-specified tolerance level. If we keep wealth at time t_n sorted in some way, say $W_{t_n}^1 \leq W_{t_n}^2 \leq \cdots \leq W_{t_n}^M$, this can be done iteratively as follows: starting at the M'th path, the sums only contain one term involving the M'th path, so one iteration looks like this:

  - update the control
  - update wealth $W_{t_{n+1}}^M$
  - update $J_w(W_{t_{n+1}}^M, t_{n+1})$
  - update the Skorohod integral $S^{h^{n,M}}$

  We iterate until some pre-specified tolerance level is satisfied. When the M'th path is done we move onto the $(M-1)$'th path. Now the sums involve two terms, one for the M'th path and one for the $(M-1)$'th path. However, we have solved for the M'th path so given this solution we can now solve for the $(M-1)$'th path as above. We continue with this until we have solved for all paths.

  When the optimal strategy has been found for all paths, we need to find $J_w(W_{t_n}^m, t_n)$ for all paths $m = 1, \ldots, M$, since this must be used at the next time step. This is done using equation (3.5)

$$W_{t_n}^m J_w(W_{t_n}^m, t_n) = e^{-\delta\Delta t}\mathbb{E}\left\{J_w(W_{t_{n+1}}^*, t_{n+1})W_{t_{n+1}}^* \mid W_{t_n} = W_{t_n}^m\right\} + u'(c_{t_n}^*)c_{t_n}^*\Delta t,$$

where the conditional expectation is estimated with the Malliavin Monte-Carlo approach.

- At time $t_0 = 0$ we only have one wealth level, so we use ordinary Monte-Carlo simulation to solve the first-order condition numerically.

- Now we have estimated the optimal consumption strategy for wealth levels $W_{t_n}^m$, for $n = 0, 1, \ldots, N$ and $m = 1, \ldots, M$. We now estimate a new candidate control process $(c_{t_n}^{m,0})_{n=0}^{N-1}$, by simulating the wealth process (with the same random numbers), but along each path we choose the consumption rate by interpolating between the consumption rates found in the previous iteration.

- This procedure could be repeated until the candidate control does not change much from one iteration to the next or until a given number of iterations have been completed. To compare the convergence properties when the algorithm is started with different initial guesses, we implement the latter.

- When simulating the wealth process we simultaneously compute the indirect utility from following the current candidate consumption strategy.

3.5 Initialization

In this section we discuss how important the initial candidate consumption strategy is. Over one period the dynamic programming principle states that the optimal strategy only depends on the current state and not on the past. However, seen from time 0 the initial guess does matter since the future wealth distribution depends on this guess. Intuitively, the simulated distribution would be better the closer the guess is to the optimal strategy. Brandt, Goyal, Santa-Clara, and Stroud (2005) also note this, but their approach is somewhat different: instead of simulating the wealth process they choose a grid of wealth levels according to some distributional assumptions.

In the next section we will analyze the algorithm for three different initial consumption strategies

- The *exact* continuous time solution.

- A *good* guess.

- A *bad* guess.

For the bad guess, the consumption rate is fixed at a level $c_t = 0.1$ for all $t \in [0, T]$. For the good guess, we choose the control $c_t = cW_t$, where c is a constant, such that the indirect utility is maximized. The problem is to

$$\sup_{c} \quad \mathbb{E}\left\{\int_0^T e^{-\delta s}u(cW_s)\mathrm{d}s + e^{-\delta T}u(W_T)\right\} \tag{3.10}$$

$$\text{s.t.} \quad \mathrm{d}W_t = W_t\left[r+\pi(\mu-r)\right]\mathrm{d}t - cW_t\mathrm{d}t + W_t\pi\sigma\mathrm{d}B_t \tag{3.11}$$

$$W_0 = w.$$

Since all coefficient in the wealth dynamics in equation (3.11) are constant, the wealth process is a geometric Brownian motion. Hence, future wealth is log-normally distributed

$$W_t = w\exp\left(\left[r+\pi(\mu-r)-c-\frac{1}{2}\pi^2\sigma^2\right]t+\pi\sigma B_t\right).$$

Since we are considering a CRRA utility function, $u(x)=\frac{x^{1-\gamma}}{1-\gamma}$, we observe from equation (3.10) that we need to compute

$$\begin{aligned}
\mathbb{E}\left\{W_t^{1-\gamma}\right\} &= \mathbb{E}\left\{w^{1-\gamma}\exp\left((1-\gamma)\left[r+\pi(\mu-r)-c-\frac{1}{2}\pi^2\sigma^2\right]t+(1-\gamma)\pi\sigma B_t\right)\right\} \\
&= w^{1-\gamma}\exp\left((1-\gamma)\left[r+\pi(\mu-r)-c-\frac{1}{2}\pi^2\sigma^2+\frac{1}{2}(1-\gamma)\pi^2\sigma^2\right]t\right) \\
&= w^{1-\gamma}\exp\left((1-\gamma)\left[r+\pi(\mu-r)-c-\frac{1}{2}\gamma\pi^2\sigma^2\right]t\right),
\end{aligned}$$

so the expectation in equation (3.10) becomes

$$\begin{aligned}
&\mathbb{E}\left\{\int_0^T e^{-\delta s}u(cW_s)\mathrm{d}s + e^{-\delta T}u(W_T)\right\} \\
&= \int_0^T e^{-\delta s}\mathbb{E}\left\{u(cW_s)\right\}\mathrm{d}s + e^{-\delta T}\mathbb{E}\left\{u(W_T)\right\} \\
&= \frac{c^{1-\gamma}}{1-\gamma}\int_0^T e^{-\delta s}\mathbb{E}\left\{W_s^{1-\gamma}\right\}\mathrm{d}s + \frac{1}{1-\gamma}e^{-\delta T}\mathbb{E}\left\{W_T^{1-\gamma}\right\} \\
&= \frac{(cw)^{1-\gamma}}{1-\gamma}\int_0^T e^{-\delta s}\exp\left((1-\gamma)\left[r+\pi(\mu-r)-c-\frac{1}{2}\gamma\pi^2\sigma^2\right]s\right)\mathrm{d}s \\
&\quad + \frac{w^{1-\gamma}}{1-\gamma}e^{-\delta T}\exp\left((1-\gamma)\left[r+\pi(\mu-r)-c-\frac{1}{2}\gamma\pi^2\sigma^2\right]T\right) \\
&= \frac{(cw)^{1-\gamma}}{1-\gamma}\int_0^T \exp\left((1-\gamma)\left[r+\pi(\mu-r)-c-\frac{1}{2}\gamma\pi^2\sigma^2-\frac{\delta}{1-\gamma}\right]s\right)\mathrm{d}s \\
&\quad + \frac{w^{1-\gamma}}{1-\gamma}\exp\left((1-\gamma)\left[r+\pi(\mu-r)-c-\frac{1}{2}\gamma\pi^2\sigma^2-\frac{\delta}{1-\gamma}\right]T\right) \\
&= \frac{(cw)^{1-\gamma}}{1-\gamma}\int_0^T \exp\left(Cs\right)\mathrm{d}s + \frac{w^{1-\gamma}}{1-\gamma}\exp\left(CT\right) \\
&= \frac{(cw)^{1-\gamma}}{1-\gamma}\frac{1}{C}\left(\exp\left(CT\right)-1\right) + \frac{w^{1-\gamma}}{1-\gamma}\exp\left(CT\right),
\end{aligned} \tag{3.12}$$

where we have defined

$$C \triangleq (1-\gamma)\left[r + \pi(\mu - r) - c - \frac{1}{2}\gamma\pi^2\sigma^2 - \frac{\delta}{1-\gamma}\right].$$

The good guess can now be found numerically by maximizing the expression in equation (3.12) with respect to c.

4 Experiences from Merton's Problem

In this section we solve Merton's problem as described in section 2 by the algorithm suggested in section 3. As seen earlier we need to estimate conditional expectations on the form

$$\mathbb{E}\left\{g(W_{t_{n+1}}) \mid W_{t_n} = w\right\},$$

for some function g, where $W_{t_{n+1}}$ is described by the Euler dynamics given in equation (3.1)

$$\begin{aligned} W_{t_0} &= W_0 \\ W_{t_{n+1}} &= W_{t_n} + W_{t_n}[r + \pi(\mu - r)]\Delta t - c_{t_n}\Delta t + W_{t_n}\pi\sigma(B_{t_{n+1}} - B_{t_n}). \end{aligned}$$

As described in section 3.3 such conditional expectations can be calculated as

$$\mathbb{E}\left\{g(W_{t_{n+1}}) \mid W_{t_n} = w\right\} = \frac{\mathbb{E}\left\{H_w(W_{t_n})g(W_{t_{n+1}})S^{h^n}(\varphi(W_{t_n} - w))\right\}}{\mathbb{E}\left\{H_w(W_{t_n})S^{h^n}(\varphi(W_{t_n} - w))\right\}}.$$

4.1 Derivation of h_t^n and the Skorohod Integral

In this section we derive the process h_t^n and compute the Skorohod integral used in equation (3.7). According to equation (3.6) we look for a process h_t^n satisfying

$$\int_0^{t_{n+1}} (D_t W_{t_n}) h_t^n \mathrm{d}t = 1 \quad \text{and} \quad \int_0^{t_{n+1}} (D_t W_{t_{n+1}}) h_t^n \mathrm{d}t = 0.$$

The Malliavin derivative of the wealth process can be determined recursively as

$$\begin{aligned} D_t W_{t_1} &= W_0 \pi\sigma \mathbf{1}_{[0,t_1]}(t) \\ D_t W_{t_{n+1}} &= D_t W_{t_n} + D_t W_{t_n}[r + \pi(\mu - r)]\Delta t - D_t W_{t_n}\frac{\partial c_{t_n}}{\partial W_{t_n}}\Delta t \\ &\quad + D_t W_{t_n}\pi\sigma(B_{t_{n+1}} - B_{t_n}) + W_{t_n}\pi\sigma\mathbf{1}_{(t_n,t_{n+1}]}(t), \end{aligned}$$

where we have used the chain rule and product rule for Malliavin derivatives, see appendix A. Assuming that h_t^n satisfies the first condition in (3.6), the second condition reduces to

$$
\begin{aligned}
&\int_0^{t_{n+1}} (D_t W_{t_{n+1}}) h_t^n \mathrm{d}t \\
&\quad = \int_0^{t_{n+1}} \bigg(D_t W_{t_n} + D_t W_{t_n}[r+\pi(\mu-r)]\Delta t - D_t W_{t_n} \frac{\partial c_{t_n}}{\partial W_{t_n}} \Delta t \\
&\qquad\qquad + D_t W_{t_n} \pi\sigma(B_{t_{n+1}} - B_{t_n}) + W_{t_n}\pi\sigma \mathbf{1}_{(t_n,t_{n+1}]}(t) \bigg) h_t^n \mathrm{d}t \\
&\quad = 1 + [r+\pi(\mu-r)]\Delta t - \frac{\partial c_{t_n}}{\partial W_{t_n}} \Delta t + \pi\sigma(B_{t_{n+1}} - B_{t_n}) + W_{t_n}\pi\sigma \int_{t_n}^{t_{n+1}} h_t^n \mathrm{d}t. \qquad (4.1)
\end{aligned}
$$

From this we see that if we define

$$
h_t^n \triangleq -(\Delta t)^{-1}(W_{t_n}\pi\sigma)^{-1} \left(1 + [r+\pi(\mu-r)]\Delta t - \frac{\partial c_{t_n}}{\partial W_{t_n}} \Delta t + \pi\sigma(B_{t_{n+1}} - B_{t_n}) \right)
$$

for $t \in (t_n, t_{n+1}]$, the second condition is satisfied as long as the first condition is satisfied. Since $D_t W_{t_n} = 0$ for $t \in (t_n, t_{n+1}]$ cf. proposition A.5, the first condition reduces to

$$
\begin{aligned}
\int_0^{t_{n+1}} (D_t W_{t_n}) h_t^n \mathrm{d}t &= \int_0^{t_n} (D_t W_{t_n}) h_t^n \mathrm{d}t \\
&= 1.
\end{aligned}
$$

Thus, if we define

$$
\begin{aligned}
h_t^n &\triangleq (\Delta t)^{-1}(D_t W_{t_n})^{-1} \\
&= (\Delta t)^{-1}(W_{t_{n-1}}\pi\sigma)^{-1}
\end{aligned}
$$

for $t \in (t_{n-1}, t_n]$ the first condition is satisfied. To summarize,

$$
h_t^n \triangleq \begin{cases} 0 & \text{for } t \in [0, t_{n-1}] \\ \dfrac{1}{W_{t_{n-1}}\pi\sigma\Delta t} & \text{for } t \in (t_{n-1}, t_n] \\ -\dfrac{1 + [r+\pi(\mu-r)]\Delta t - \frac{\partial c_{t_n}}{\partial W_{t_n}}\Delta t + \pi\sigma(B_{t_{n+1}} - B_{t_n})}{W_{t_n}\pi\sigma\Delta t} & \text{for } t \in (t_n, t_{n+1}] \end{cases}
$$

satisfies (3.6). To compute the Skorohod integrals in equation (3.7) it is sufficient to compute the Skorohod integral of h_t^n since equation (A.9) gives

$$
\begin{aligned}
S^{h^n}(\varphi(W_{t_n} - w)) &= \varphi(W_{t_n} - w) S^{h^n}(1) - \varphi'(W_{t_n} - w) \\
&= \varphi(W_{t_n} - w) \left[S^{h^n}(1) + \eta \right],
\end{aligned}
$$

where the last equality holds if $\varphi(x) = e^{-\eta x}$ for some $\eta > 0$. We obtain

$$
\begin{aligned}
\int_0^{t_{n+1}} h_t^n \delta B_t &= \int_0^{t_{n-1}} h_t^n \delta B_t + \int_{t_{n-1}}^{t_n} h_t^n \delta B_t + \int_{t_n}^{t_{n+1}} h_t^n \delta B_t \\
&= \int_{t_{n-1}}^{t_n} h_t^n \delta B_t + \int_{t_n}^{t_{n+1}} h_t^n \delta B_t \\
&= \frac{B_{t_n} - B_{t_{n-1}}}{W_{t_{n-1}} \pi\sigma\Delta t} + \int_{t_n}^{t_{n+1}} h_t^n \delta B_t \\
&= \frac{B_{t_n} - B_{t_{n-1}}}{W_{t_{n-1}} \pi\sigma\Delta t} - \frac{\left(1 + [r + \pi(\mu - r)]\Delta t - \frac{\partial c_{t_n}}{\partial W_{t_n}}\Delta t\right)(B_{t_{n+1}} - B_{t_n})}{W_{t_n}\pi\sigma\Delta t} \\
&\quad - \frac{\pi\sigma}{W_{t_n}\pi\sigma\Delta t} \int_{t_n}^{t_{n+1}} (B_{t_{n+1}} - B_{t_n}) \delta B_t \\
&= \frac{B_{t_n} - B_{t_{n-1}}}{W_{t_{n-1}} \pi\sigma\Delta t} - \frac{\left(1 + [r + \pi(\mu - r)]\Delta t - \frac{\partial c_{t_n}}{\partial W_{t_n}}\Delta t\right)(B_{t_{n+1}} - B_{t_n})}{W_{t_n}\pi\sigma\Delta t} \\
&\quad - \frac{(B_{t_{n+1}} - B_{t_n})^2 - \Delta t}{W_{t_n}\Delta t},
\end{aligned}
\tag{4.2}
$$

where the last Skorohod integral has been computed in example A.19.

4.2 The Solutions

In this section the numerical properties of the suggested algorithm will be presented and analyzed. The exact solution to the continuous time Merton problem will be used as a benchmark. Of course, we would only expect our algorithm to converge to the continuous time solution as the number of time steps $N \to \infty$. On the other hand, we could compare our solution to the exact solution of a discrete time portfolio problem, but since we use an Euler approximation of the continuous time wealth process, we would not expect the algorithm to converge to the discrete time solution either.

4.2.1 The Exact Continuous Time Solution

Since we try to solve Merton's problem it seems reasonable to compare our results with the exact solution. As parameter values we have chosen

$$\mu = 0.08, \quad \sigma = 0.20, \quad r = 0.04, \quad \delta = 0.02, \quad \gamma = 4, \quad T = 1, \quad \text{and} \quad W_0 = 1.$$

Using the result stated in theorem 2.1, we get

$$\begin{aligned}
A &= \frac{\delta - r(1-\gamma)}{\gamma} - \frac{1}{2}\frac{1-\gamma}{\gamma^2}\frac{(\mu-r)^2}{\sigma^2} \\
&= \frac{0.02+0.12}{4} + \frac{1}{2}\frac{3}{16}\frac{0.0016}{0.04} \\
&= 0.03875 \\
g(0) &= A^{-1}\left(1 + [A-1]e^{-AT}\right) \\
&= 1.9429,
\end{aligned}$$

such that the investor should hold a fraction of

$$\pi_t^* = \frac{1}{\gamma}\frac{\mu - r}{\sigma^2} = \frac{1}{4}\frac{0.04}{0.04} = 0.25$$

invested in the risky asset, and initially consume at a rate

$$c_0^* = \frac{1}{g(0)} W_0 = 0.5147.$$

Following the optimal strategies the agent yields an indirect utility of

$$J(W_0, 0) = g(0)^4 \frac{W_0^{-3}}{-3} = -4.7495.$$

4.2.2 The Numerical Solution

In this subsection we will analyze the convergence properties of the algorithm. In the first test, the consumption rate is initialized to the exact continuous time solution. For a given number of simulations we then run the algorithm 100 times to obtain an estimate of the initial consumption rate and indirect utility. The algorithm is run with one outer iteration, meaning that we perform only one backward recursion and update the wealth process once. The results are shown in table 1. We observe that not much precision is gained when increasing the number of simulations above 5000. Hence we use a maximum of 5000 simulations in the following tests.

Note that the estimated indirect utility is higher than the exact continuous time solution, which seems counterintuitive since the discrete time consumption strategy is feasible in a continuous time setting. The peculiar difference occurs because we use an Euler discretization, which we only expect to yield good results when $\Delta t = \frac{T}{N}$ is small. It would be more fair to compare with the indirect utility obtained after zero iterations because we would expect this solution to converge to the exact solution since we initialize the control to the exact continuous time consumption rate. These values are also higher than the exact continuous time solution.

We now consider how many iterations are needed to obtain reasonable results with various initial consumption strategies. The results are given in table 2 and figure 1 for $M = 2400$ simulations and in table 3 and figure 2 for $M = 5000$ simulations. We observe that the initial candidate control

Simulations	Consumption rate Estimate	Consumption rate Standard deviation	Indirect utility Estimate	Indirect utility Standard deviation
1000	0.5129	0.0013	−4.7426	0.0020
2000	0.5132	0.0008	−4.7426	0.0021
3000	0.5135	0.0008	−4.7422	0.0016
4000	0.5138	0.0006	−4.7421	0.0011
5000	0.5136	0.0004	−4.7421	0.0011
6000	0.5135	0.0005	−4.7423	0.0011
7000	0.5136	0.0004	−4.7421	0.0009
8000	0.5136	0.0004	−4.7420	0.0007
9000	0.5137	0.0004	−4.7422	0.0008
10000	0.5137	0.0004	−4.7420	0.0007

Table 1: Dependence on the number of simulations for $N = 12$ time steps, 1 outer iteration and 100 batches. For both the consumption rate column and indirect utility column the estimate is the average over the 100 runs and the standard deviation is the standard deviation of this estimate.

process plays a significant role in how fast the algorithm converges. This was expected, since a bad guess gives a poor estimated wealth distribution and hence poor estimates for the conditional expectations. We also note that the algorithm alters the exact guess, because this guess is not optimal in the discretized problem.

So far we have only tested how well the algorithm finds the optimal solution at time 0. To illustrate how well it finds the entire optimal consumption process, we now consider the future distribution of the consumption process and the associated wealth process. First, we derive the expected value and standard derivation of the continuous time optimal wealth process. Inserting the optimal controls in the wealth dynamics in equation (2.1) we get

$$\mathrm{d}W_t = W_t\left[r + \pi(\mu - r)\right]\mathrm{d}t - \frac{W_t}{g(t)}\mathrm{d}t + W_t\pi\sigma\mathrm{d}B_t,$$

which has the solution

$$W_t = W_0 \exp\left(\left[r + \pi(\mu - r) - \frac{1}{2}\sigma^2\pi^2\right]t - \int_0^t \frac{1}{g(s)}\mathrm{d}s + \sigma\pi B_t\right). \tag{4.3}$$

Iteration	Exact solution		Good guess		Bad guess	
i	c_0^i	J_0^i	c_0^i	J_0^i	c_0^i	J_0^i
0	0.5147	−4.7406	0.6996	−5.3205	0.1000	−330.6860
1	0.5140	−4.7406	0.5037	−4.7509	0.7926	−56.2071
2	0.5141	−4.7406	0.5088	−4.7424	0.4545	−5.0308
3	0.5141	−4.7406	0.5163	−4.7407	0.5382	−4.7490
4	0.5141	−4.7406	0.5135	−4.7407	0.5088	−4.7414
5	0.5141	−4.7406	0.5142	−4.7406	0.5151	−4.7407
6	0.5141	−4.7406	0.5141	−4.7406	0.5138	−4.7409
7	0.5141	−4.7406	0.5141	−4.7406	0.5141	−4.7406
8	0.5141	−4.7406	0.5141	−4.7407	0.5141	−4.7406
9	0.5141	−4.7406	0.5141	−4.7406	0.5141	−4.7406
10	0.5140	−4.7406	0.5141	−4.7406	0.5140	−4.7407

Table 2: Convergence results for $N = 12$ time steps and $M = 2400$ simulations. For the three different initializations the initial consumption rate and indirect utility resulting from the first 10 iterations are shown.

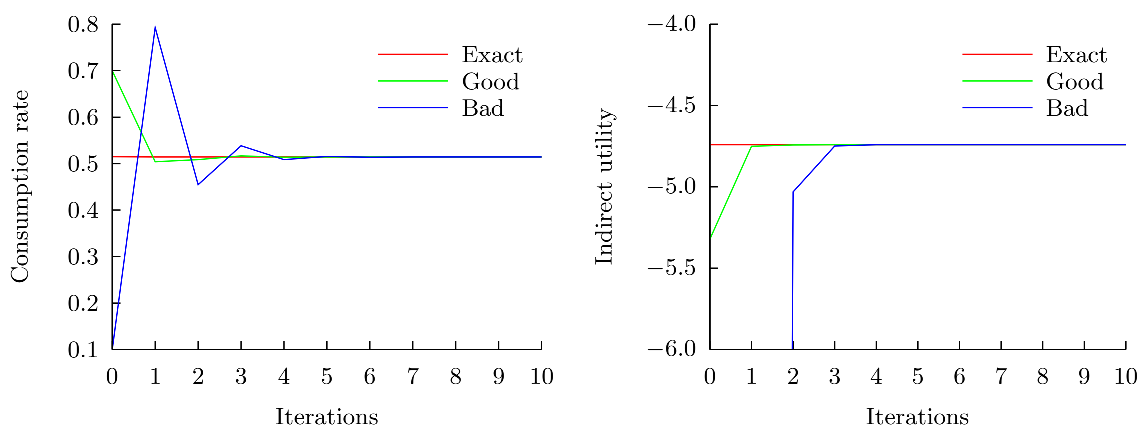

Figure 1: Convergence results for $N = 12$ time steps and $M = 2400$ simulations. For the three different initializations the initial consumption rate and indirect utility resulting from the first 10 iterations are shown.

Iteration	Exact solution		Good guess		Bad guess	
i	c_0^i	J_0^i	c_0^i	J_0^i	c_0^i	J_0^i
0	0.5147	−4.7413	0.6996	−5.3213	0.1000	−330.6861
1	0.5128	−4.7416	0.5023	−4.7507	0.7907	−202.2616
2	0.5133	−4.7417	0.5082	−4.7436	0.5829	−4.9191
3	0.5132	−4.7419	0.5153	−4.7416	0.4882	−4.7578
4	0.5132	−4.7420	0.5122	−4.7421	0.5195	−4.7420
5	0.5130	−4.7421	0.5132	−4.7420	0.5117	−4.7419
6	0.5130	−4.7422	0.5132	−4.7418	0.5136	−4.7416
7	0.5131	−4.7431	0.5130	−4.7419	0.5130	−4.7419
8	0.5131	−4.7425	0.5130	−4.7420	0.5130	−4.7420
9	0.5131	−4.7424	0.5131	−4.7422	0.5130	−4.7420
10	0.5131	−4.7423	0.5131	−4.7422	0.5131	−4.7421

Table 3: Convergence results for $N = 12$ time steps and $M = 5000$ simulations. For the three different initializations the initial consumption rate and indirect utility resulting from the first 10 iterations are shown.

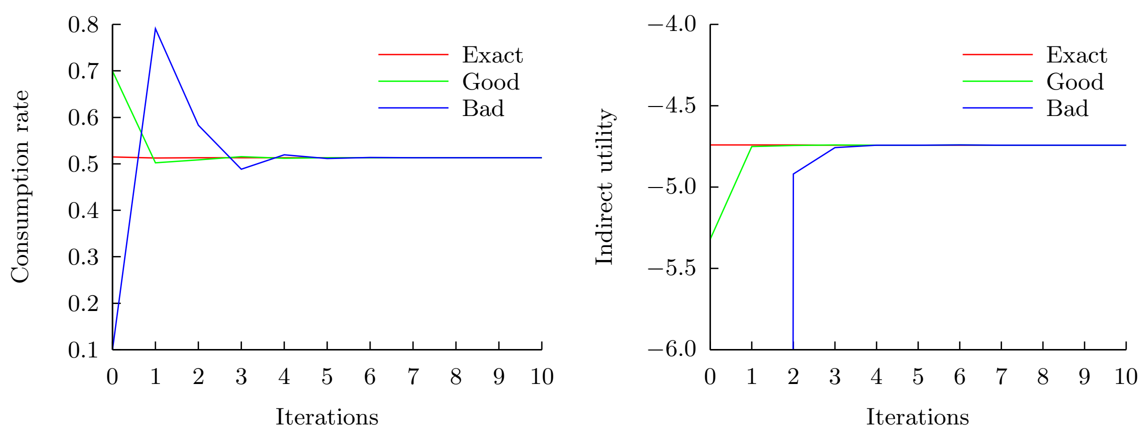

Figure 2: Convergence results for $N = 12$ time steps and $M = 5000$ simulations. For the three different initializations the initial consumption rate and indirect utility resulting from the first 10 iterations are shown.

In this equation

$$
\begin{aligned}
\int_0^t \frac{1}{g(s)}\mathrm{d}s &= \int_0^t \frac{A}{1+(A-1)e^{-A(T-s)}}\mathrm{d}s \\
&= \int_{-AT}^{-A(T-t)} \frac{A}{1+(A-1)e^x}\frac{1}{A}\mathrm{d}x \\
&= \int_{-AT}^{-A(T-t)} \frac{1}{1+(A-1)e^x}\mathrm{d}x \\
&= \left[x - \log\left(1+(A-1)e^x\right)\right]_{-AT}^{-A(T-t)} \\
&= At - \log\left(\frac{1+(A-1)e^{-A(T-t)}}{1+(A-1)e^{-AT}}\right),
\end{aligned}
$$

where A is defined as in equation (2.5). Inserting into equation (4.3) we get

$$
W_t = W_0 \exp\left(\left[r+\pi(\mu-r)-\frac{1}{2}\sigma^2\pi^2 - A\right]t + \sigma\pi B_t\right)\left(\frac{1+(A-1)e^{-A(T-t)}}{1+(A-1)e^{-AT}}\right).
$$

Since W_t is log-normally distributed the first two moments are

$$
\begin{aligned}
\mathbb{E}\left\{W_t\right\} &= W_0 \exp\left(\left[r+\pi(\mu-r)-A\right]t\right)\left(\frac{1+(A-1)e^{-A(T-t)}}{1+(A-1)e^{-AT}}\right) \\
\mathbb{E}\left\{W_t^2\right\} &= W_0^2 \exp\left(2\left[r+\pi(\mu-r)+\frac{1}{2}\sigma^2\pi^2 - A\right]t\right)\left(\frac{1+(A-1)e^{-A(T-t)}}{1+(A-1)e^{-AT}}\right)^2 \\
&= \mathbb{E}\left\{W_t\right\}^2 \exp\left(\sigma^2\pi^2 t\right).
\end{aligned}
$$

Therefore, the variance of future expected wealth is

$$
\begin{aligned}
\mathbb{V}\left\{W_t\right\} &= \mathbb{E}\left\{W_t^2\right\} - \mathbb{E}\left\{W_t\right\}^2 \\
&= \mathbb{E}\left\{W_t\right\}^2 \exp\left(\sigma^2\pi^2 t\right) - \mathbb{E}\left\{W_t\right\}^2 \\
&= \mathbb{E}\left\{W_t\right\}^2 \left[\exp\left(\sigma^2\pi^2 t\right) - 1\right].
\end{aligned}
$$

Likewise, the expected future consumption rate and its variance are given by

$$
\begin{aligned}
\mathbb{E}\left\{c_t\right\} &= \mathbb{E}\left\{\frac{W_t}{g(t)}\right\} \\
&= \frac{1}{g(t)}\mathbb{E}\left\{W_t\right\} \\
\mathbb{V}\left\{c_t\right\} &= \mathbb{V}\left\{\frac{W_t}{g(t)}\right\} \\
&= \frac{1}{g(t)^2}\mathbb{V}\left\{W_t\right\}.
\end{aligned}
$$

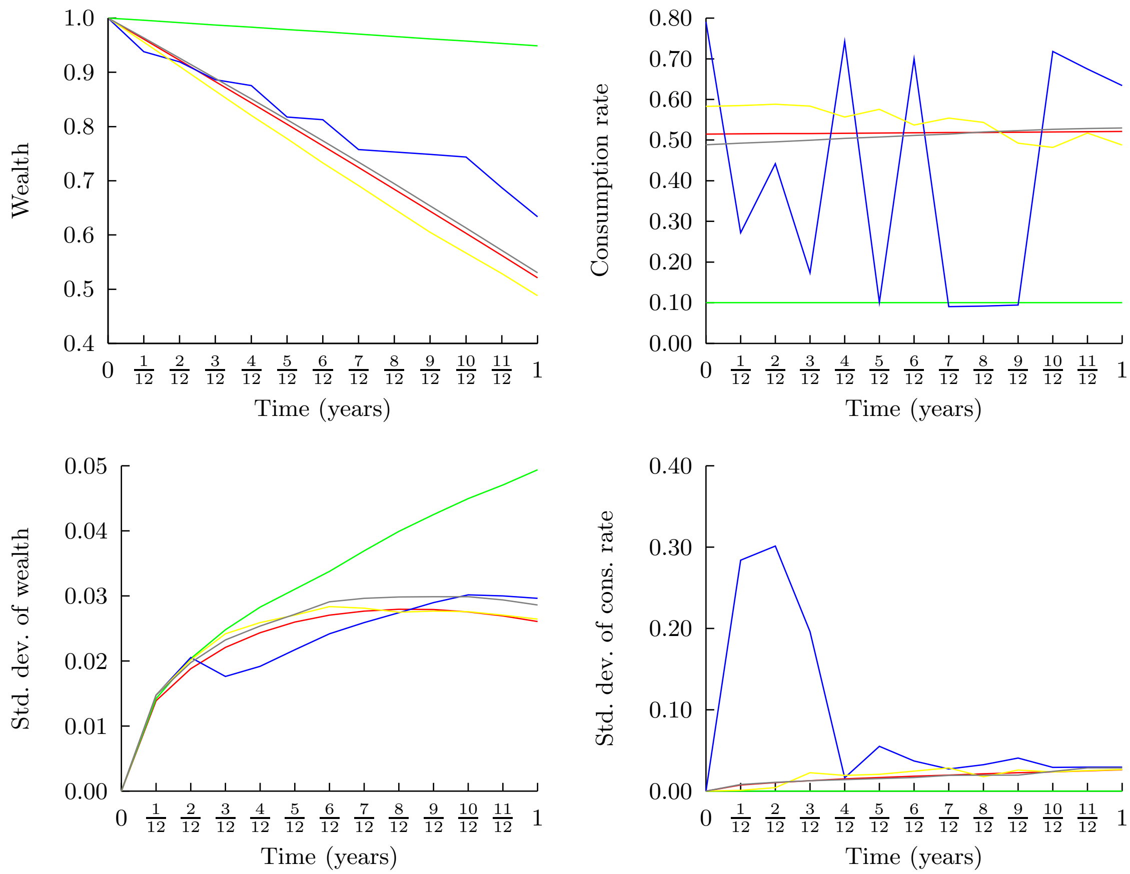

Figure 3: Future wealth and consumption rate when the initial consumption rate is $c_t = 0.1$ for all $t \in [0, T]$. The red line (—) indicates the exact expected wealth and consumption rate from Merton's problem. In the two upper subfigures the green line (—), blue line (—), and yellow line (—) indicate the average of the future simulated wealth levels and consumption rates after iterations 0, 1, and 2, respectively. In the two lower subfigures the corresponding standard deviations are depicted.

Inserting our parameter values in the above formulas, we can now compare the true wealth and consumption rate distributions with the simulated distributions.

In figure 3 the average future consumption rate and its standard deviation are plotted for different iterations when we initialize the algorithm with the bad guess. Also the exact expected wealth and consumption rate and their standard deviations are depicted. Likewise, figure 4 illustrates the same when the algorithm is initialized with the good guess. Again, we observe that the initialization of the consumption strategy is important for the convergence speed. For the initial exact continuous time guess the distribution is (nearly) correct from the beginning so a similar figure will serve no purpose.

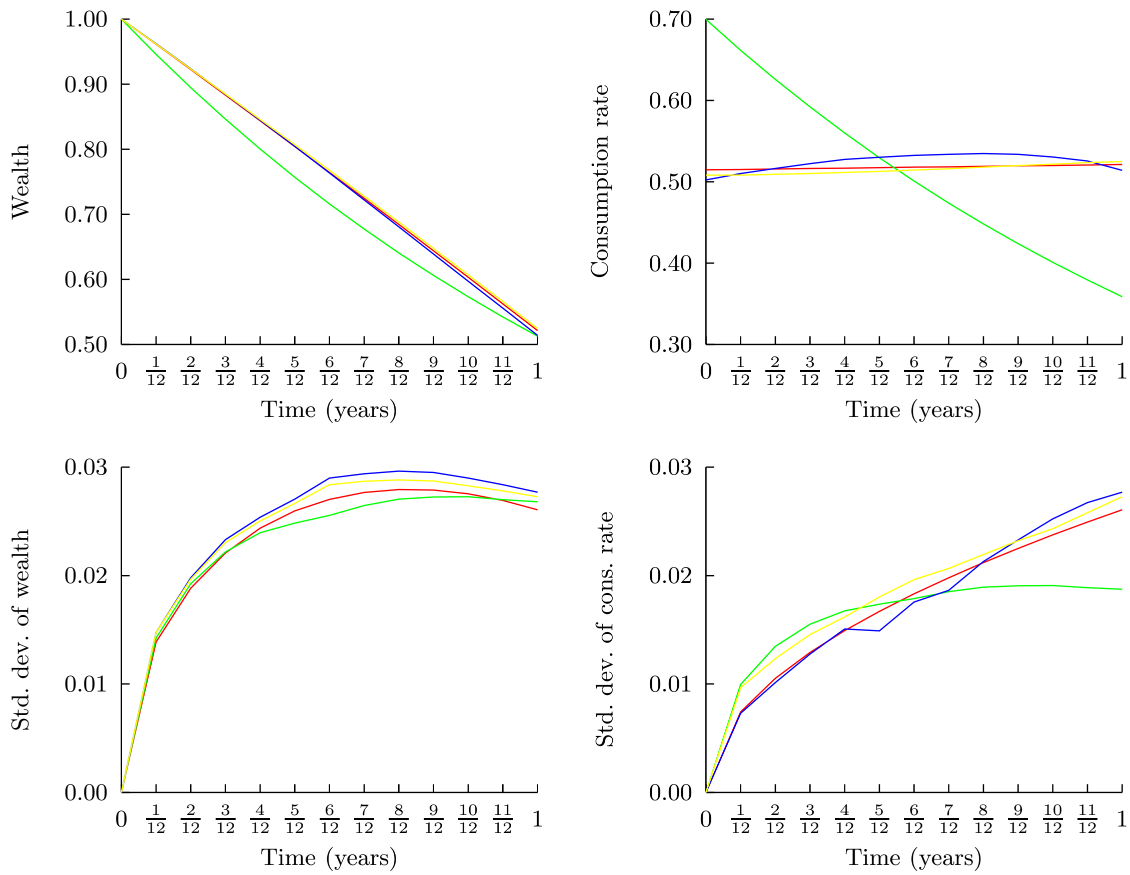

Figure 4: Future wealth and consumption rate when the initial consumption rate is $c_t = cW_t$, where c is chosen to maximize the expression in equation (3.12). The red line (—) indicates the exact expected wealth and consumption rate from Merton's problem. In the two upper subfigures the green line (—), blue line (—), and yellow line (—) indicate the average of the future simulated wealth levels and consumption rates after iterations 0, 1, and 2, respectively. In the two lower subfigures the corresponding standard deviations are depicted.

4.3 The Numerical Problems

Until now the convergence properties have only been analyzed for a coarse time discretization.[1] In this section we will discuss some numerical issues that we encountered during our tests for finer time discretizations. Since we test the algorithm on a continuous time problem, the ultimate test is whether the numerical solution converges to the exact solution when $\Delta t \to 0$.

When we increase the number of time steps, however, numerical problems occur. For the exact and good guesses, the first iterations behave nicely, but at some point the consumption rate process diverges. Since we let $\Delta t \to 0$, the explanation should be found here. Numerical experiments indicate that the problem is due to the Skorohod integral, which becomes large when $\Delta t \to 0$, as can be seen from equation (4.2). A possible solution to this problem would be to find a different $h_t^n \in \mathbb{H}_n$ which behaves more nicely. Also, when the function g in

$$\mathbb{E}\left\{g(W_{t_{n+1}}) \mid W_{t_n} = w\right\} = \frac{\mathbb{E}\left\{H_w(W_{t_n})g(W_{t_{n+1}})S^{h^n}(\varphi(W_{t_n} - w))\right\}}{\mathbb{E}\left\{H_w(W_{t_n})S^{h^n}(\varphi(W_{t_n} - w))\right\}}$$

is differentiable, it is possible to relax the restrictions on h_t^n, as shown by Fournié, Lasry, Lebuchoux, and Lions (2001). To be specific we could drop the restriction that

$$\int_0^{t_{n+1}} (D_t W_{t_{n+1}}) h_t^n \mathrm{d}t = 0$$

in equation (A.8) and adjust the proof of lemma A.21 to obtain

$$\begin{aligned}\mathbb{E}\left\{g(W_{t_{n+1}}) \mid W_{t_n} = w\right\} = {} & \frac{\mathbb{E}\left\{H_w(W_{t_n})g(W_{t_{n+1}})S^{h^n}(\varphi(W_{t_n} - w))\right\}}{\mathbb{E}\left\{H_w(W_{t_n})S^{h^n}(\varphi(W_{t_n} - w))\right\}} \\ & - \frac{\mathbb{E}\left\{H_w(W_{t_n})g'(W_{t_{n+1}})\varphi(W_{t_n} - w)\int_0^{t_{n+1}}(D_t W_{t_{n+1}})h_t^n \mathrm{d}t\right\}}{\mathbb{E}\left\{H_w(W_{t_n})S^{h^n}(\varphi(W_{t_n} - w))\right\}}.\end{aligned}$$

However, to do this we must keep track of the Malliavin derivative of the wealth process and the integral of it multiplied by h_t^n. If, for example, we put $h_t^n = 0$ for $t \in (t_n, t_{n+1}]$, this is not as difficult as it first might seem, since the last term in equation (4.1) disappears. However, as can be seen from equation (4.2), it is not this restriction on h_t^n that makes the Skorohod integral explode when $\Delta t \to 0$, so more effort should be made to choose h_t^n carefully for $t \in [0, t_n]$.

4.4 Multi-Dimensional Problems

We will now consider how the algorithm can be extended to handle multi-dimensional problems. As proved in Bouchard, Ekeland, and Touzi (2004), theorem 3.1 can be generalized to a multi-dimensional Markov diffusion. Also, as example 2.2 in Bouchard, Ekeland, and Touzi demonstrates, it is possible to find h_t^n such that equation (A.8) is satisfied. Note, however, that the Malliavin derivative of a multi-dimensional state variable is a matrix, in which the ij'th element in some

[1] The portfolio problem analyzed by Bouchard, Ekeland, and Touzi (2004) also has a coarse time discretization.

sense is the Malliavin derivative of the i'th state variable with respect to the j'th Brownian motion. Hence, h_t^n also becomes a matrix. However, the Skorohod integral involved in the alternative representation of conditional expectations in equation (3.7) becomes a multiple stochastic integral, which numerically is a challenging task to handle.

5 Conclusion

In this paper we have proposed an algorithm to solve stochastic control problems numerically. The algorithm is based on the dynamic programming approach to the solution of stochastic control problems. We focus on a classical stochastic control problem in finance, namely Merton's optimal portfolio choice problem. Given an initial guess about the agent's consumption strategy, his wealth process can be simulated. With a recursive backwards procedure the consumption strategy is updated such that it satisfies a first-order condition at all simulated wealth paths. This procedure requires the estimation of many conditional expectations. Malliavin calculus provides an alternative representation of such conditional expectations, which can be estimated with Monte-Carlo simulation.

Our numerical tests indicate promising convergence results when the number of time steps is small, so the algorithm seems to be applicable in the solution of discrete time stochastic control problems. When the number of time steps is increased the algorithm improves the control and indirect utility in the first iterations, but at some point it diverges from the optimal solution. Hence, further research must be conducted before applying the algorithm to the solution of continuous stochastic control problems.

Extending the algorithm to a multi-dimensional setting should be relatively unproblematic, since the mathematical foundation has been developed in a multi-dimensional setting. Also, the extension will usually consist in adding exogenous state variables, which does not complicate the demanding control updating procedure much. However, problem specific issues could arise in finding the process h_t^n, which is a matrix in a multi-dimensional setting, and also in computing the Skorohod integrals, if they cannot be solved analytically. These problems are subject to further research.

A Some Results from Malliavin Calculus

This appendix will give a brief introduction to Malliavin calculus and provide the results needed to develop the algorithm in section 3. To keep things simple the results are developed in one dimension, and to some extent only heuristically. For a more thorough treatment of the subject we refer the interested reader to Nualart (1995) or Øksendal (1997). Friz (2005) also gives a good introduction to the topic.

Since the usual probability space $(\Omega, \mathcal{F}, P)$ is very abstract, it is not possible to define the concept of differentiation per se. What we need is some kind of structure. For our needs the Wiener space turns out to have the desired properties. We begin this appendix with a justification of why the Wiener space possess the necessary properties. We then define the concepts of differentiation and integration on the Wiener Space, and demonstrate the concepts with some examples. We end this appendix with a proof of theorem 3.1.

A.1 The Wiener Space

Definition A.1 (The Wiener Space)**.** We let $\Omega \triangleq C_0([0,T])$ denote the space of real continuous functions on $[0,T]$ with value 0 at 0, i.e.

$$\Omega \triangleq C_0([0,T]) \triangleq \{\omega : [0,T] \to \mathbb{R} \mid \omega \text{ continuous}, \ \omega(0) = 0\}.$$

This space is called the *Wiener space* for reasons obvious by the end of this section.

Consider a probability space $(\Xi, \mathcal{G}, \nu)$ and let $\boldsymbol{\beta} = (\beta_t)_{t\in[0,T]}$ be a Brownian motion with respect to the probability measure ν, and let the σ-algebra, $\mathcal{G}$, be the natural filtration generated by this Brownian motion, i.e. $\mathcal{G} = \mathcal{G}_T$, where $\mathcal{G}_t \triangleq \sigma(\beta_s \mid 0 \le s \le t)$. Since Brownian motion is continuous, it can be regarded as a mapping from Ξ into Ω, namely the mapping of ξ from Ξ to the continuous function $t \to \beta_t(\xi)$.

We equip Ω with the σ-algebra $\mathcal{F}$ generated by the finite-dimensional cylinder sets

$$\{\omega \mid \omega(t_1) \in A_1, \dots, \omega(t_n) \in A_n\}, \quad 0 \le t_1 < t_2 < \cdots < t_n \le T, \quad A_1, A_2, \dots, A_n \in \mathcal{B},$$

where $\mathcal{B}$ is the Borel σ-algebra on $\mathbb{R}$. Now the Brownian motion, $\xi \to \boldsymbol{\beta}(\xi)$, can be regarded as a measurable mapping from $(\Xi, \mathcal{G}, \nu)$ to $(\Omega, \mathcal{F})$, and therefore induces a probability measure P on $(\Omega, \mathcal{F})$ given by

$$P\left(\{\omega \mid \omega(t_1) \in A_1, \dots, \omega(t_n) \in A_n\}\right) = \nu\left(\beta_{t_1} \in A_1, \dots, \beta_{t_n} \in A_n\right).$$

This measure is called the *Wiener measure.* Defining the coordinate mapping process, $B_t : \Omega \to \mathbb{R}$, on the Wiener space by

$$B_t(\omega) \triangleq \omega(t),$$

we now note that the process $\boldsymbol{B} = (B_t)_{t\in[0,T]}$ has the same distribution under P as $\boldsymbol{\beta} = (\beta_t)_{t\in[0,T]}$ has under ν. Hence $\boldsymbol{B} = (B_t)_{t\in[0,T]}$ is a Brownian motion on $(\Omega, \mathcal{F}, P)$. Moreover $\mathcal{F} = \mathcal{F}_T$, where $\mathcal{F}_t \triangleq \sigma\left(B_s \mid 0 \le s \le t\right)$ is the σ-algebra generated by the Brownian motion $\boldsymbol{B} = (B_t)_{t\in[0,T]}$.

We have now demonstrated that the coordinate mapping on the Wiener space becomes a Brownian motion under the Wiener measure. Hence, when we need to work with a Brownian motion, we can work with the coordinate mapping on Wiener space instead of some abstract probability space.

A.2 Differentiation on Wiener Space

It is well-known that Brownian motion is nowhere differentiable with respect to time. However, it is possible to define the concept of differentiation of random variables with respect to perturbations in the underlying Brownian motion, as we shall see in this section.

Definition A.2 (Øksendal (1997), Definitions 4.6–4.7)**.** Let $h \in L^2([0,T])$ be a deterministic function and consider directions on the form

$$\gamma(t) = \int_0^t h(s)\mathrm{d}s.$$

Notice that $t \to \gamma(t)$ is continuous on $[0,T]$ and $\gamma(0) = 0$. Therefore, $\gamma \in \Omega$ and is a valid direction.

For a random variable $F : \Omega \to \mathbb{R}$ we define the *directional derivative* of F at the point ω in the direction γ by

$$D_\gamma F(\omega) \triangleq \lim_{\varepsilon \to 0} \frac{F(\omega + \varepsilon\gamma) - F(\omega)}{\varepsilon}$$

if the limit exists in $L^2(\Omega)$. Further, if there exists $\psi(t,\omega) \in L^2([0,T] \times \Omega)$ such that

$$D_\gamma F(\omega) = \int_0^T \psi(t,\omega)h(t)\mathrm{d}t,$$

we say that F is *differentiable* and define the *derivative* of F to be the random variable

$$D_t F(\omega) \triangleq \psi(t,\omega).$$

Finally, we let $\mathcal{D}_{1,2}$ denote the set of all differentiable random variables.

Remark. As noted in Øksendal (1997), it is not clear whether $\mathcal{D}_{1,2}$ is closed under the norm

$$\|F\|_{1,2} \triangleq \|F\|_{L^2(\Omega)} + \|D_t F\|_{L^2([0,T]\times\Omega)}, \quad F \in \mathcal{D}_{1,2}.$$

To overcome this problem we introduce the family, $\mathbb{P}$, of *Wiener polynomials*, i.e. random variables $F : \Omega \to \mathbb{R}$ of the form

$$F(\omega) = p\left(\int_0^T h^1(t)\mathrm{d}B_t(\omega), \int_0^T h^2(t)\mathrm{d}B_t(\omega), \ldots, \int_0^T h^n(t)\mathrm{d}B_t(\omega)\right),$$

where p is a polynomial of degree n and $h^1, \ldots, h^n \in L^2([0,T])$. Due to the chain rule (A.1) and example A.8 the family of Wiener polynomials is differentiable, i.e. $\mathbb{P} \subset \mathcal{D}_{1,2}$. The closure of $\mathbb{P}$ with respect to the norm $\|F\|_{1,2}$ is the space, $\mathbb{D}_{1,2}$, containing all $F \in L^2(\Omega)$ for which there exists

$F_n \in \mathbb{P}$ such that

$$\begin{aligned} F_n &\to F \quad \text{in } L^2(\Omega) \\ (D_t F_n)_{n=1}^{\infty} & \quad \text{is convergent in } L^2([0,T]\times\Omega). \end{aligned}$$

By definition $\mathbb{P} \subseteq \mathbb{D}_{1,2} \subseteq L^2(\Omega)$. Noting that the closure of $\mathbb{P}$ with respect to the norm $\|F\|_{L^2(\Omega)}$ is equal to $L^2(\Omega)$, one can be tempted to think that $\mathbb{D}_{1,2} = L^2(\Omega)$ as well. As we shall see in example A.11, however, $\mathbb{D}_{1,2} \subsetneqq L^2(\Omega)$ so the two norms are not equivalent.

For elements in $\mathbb{D}_{1,2}$ we can now define a derivative as the limit of $D_t F_n$. This is the so-called *Malliavin* derivative. Since the two derivatives coincide for $F \in \mathcal{D}_{1,2} \cap \mathbb{D}_{1,2}$, we will use the notation $D_\gamma F$ for the directional derivative and $D_t F$ for the derivative of such random variables.

A.2.1 Differentiation Rules

In this section we present some useful results, which make the computation of the Malliavin derivative easier. The results are developed on a heuristic level and are analogous to similar results from ordinary calculus.

Proposition A.3 (Chain rule). *Let $F \in \mathcal{D}_{1,2}$ and let $f : \mathbb{R} \to \mathbb{R}$ be a differentiable function. Then $f(F(\omega)) \in \mathcal{D}_{1,2}$ and*

$$D_t f(F(\omega)) = f'(F(\omega)) D_t F(\omega). \tag{A.1}$$

Proof. From definition A.2 we immediately get

$$\begin{aligned} D_\gamma f(F(\omega)) &= \lim_{\varepsilon\to 0} \frac{1}{\varepsilon} \left[f(F(\omega+\varepsilon\gamma)) - f(F(\omega)) \right] \\ &= \lim_{\varepsilon\to 0} \left[\frac{f(F(\omega+\varepsilon\gamma)) - f(F(\omega))}{F(\omega+\varepsilon\gamma) - F(\omega)} \frac{F(\omega+\varepsilon\gamma) - F(\omega)}{\varepsilon} \right] \\ &= f'(F(\omega)) D_\gamma F(\omega) \end{aligned}$$

and since $F \in \mathcal{D}_{1,2}$ we know that $D_t F(\omega)$ exists in $L^2([0,T]\times\Omega)$ and that

$$\begin{aligned} D_\gamma f(F(\omega)) &= f'(F(\omega)) D_\gamma F(\omega) \\ &= f'(F(\omega)) \int_0^T D_t F(\omega) h(t) \mathrm{d}t \\ &= \int_0^T f'(F(\omega)) D_t F(\omega) h(t) \mathrm{d}t. \end{aligned}$$

Hence $f(F) \in \mathcal{D}_{1,2}$ and

$$D_t f(F(\omega)) = f'(F(\omega)) D_t F(\omega). \qquad \square$$

Proposition A.4 (Product rule)**.** *If $F, G \in \mathcal{D}_{1,2}$, then $FG \in \mathcal{D}_{1,2}$ and*

$$D_t(F(\omega)G(\omega)) = (D_tF(\omega))G(\omega) + F(\omega)(D_tG(\omega)). \tag{A.2}$$

Proof. From definition A.2 we immediately get

$$\begin{aligned}
D_\gamma\left(F(\omega)G(\omega)\right) &= \lim_{\varepsilon\to 0}\frac{F(\omega+\varepsilon\gamma)G(\omega+\varepsilon\gamma)-F(\omega)G(\omega)}{\varepsilon}\\
&= \lim_{\varepsilon\to 0}\frac{F(\omega+\varepsilon\gamma)G(\omega+\varepsilon\gamma)-F(\omega)G(\omega+\varepsilon\gamma)+F(\omega)G(\omega+\varepsilon\gamma)-F(\omega)G(\omega)}{\varepsilon}\\
&= \lim_{\varepsilon\to 0}\frac{\left[F(\omega+\varepsilon\gamma)-F(\omega)\right]G(\omega+\varepsilon\gamma)+F(\omega)\left[G(\omega+\varepsilon\gamma)-G(\omega)\right]}{\varepsilon}\\
&= \lim_{\varepsilon\to 0}\left(\frac{F(\omega+\varepsilon\gamma)-F(\omega)}{\varepsilon}G(\omega+\varepsilon\gamma)+F(\omega)\frac{G(\omega+\varepsilon\gamma)-G(\omega)}{\varepsilon}\right)\\
&= \left(D_\gamma F(\omega)\right)G(\omega)+F(\omega)\left(D_\gamma G(\omega)\right).
\end{aligned}$$

Also, since both $F, G \in \mathcal{D}_{1,2}$ we get

$$\begin{aligned}
D_\gamma\left(F(\omega)G(\omega)\right) &= \left(D_\gamma F(\omega)\right)G(\omega)+F(\omega)D_\gamma G(\omega)\\
&= \int_0^T (D_tF(\omega))h(t)\mathrm{d}tG(\omega)+F(\omega)\int_0^T (D_tG(\omega))h(t)\mathrm{d}t\\
&= \int_0^T \left[\left(D_tF(\omega)\right)G(\omega)+F(\omega)(D_tG(\omega))\right]h(t)\mathrm{d}t
\end{aligned}$$

so

$$D_t(F(\omega)G(\omega)) = \left(D_tF(\omega)\right)G(\omega) + F(\omega)(D_tG(\omega)).$$

□

Proposition A.5. *Let $F \in \mathbb{D}_{1,2}$ be $\mathcal{F}_s$-adapted. Then D_tF will be $\mathcal{F}_s$-adapted and for $t > s$ we have*

$$D_tF = 0.$$

Proof. We will only prove the result in a special case. A thorough proof builds on Wiener-Itô chaos expansions, which are outside the scope of this paper. The interested reader is referred to Nualart (1995) or Øksendal (1997). Consider a random variable on the form

$$F(\omega) = \exp\left(\int_0^T h(u)\mathrm{d}B_u(\omega) - \frac{1}{2}\int_0^T h(u)^2\mathrm{d}u\right), \tag{A.3}$$

where $h \in L^2([0,T])$ is deterministic. Note, that Novikov's condition is satisfied, so F is an exponential martingale. We then have

$$D_tF = Fh(t)$$

by the chain rule (A.1) and example A.8. Therefore

$$
\begin{aligned}
D_t \mathbb{E}\{F \mid \mathcal{F}_s\} &= D_t \exp\left(\int_0^s h(u)\mathrm{d}B_u(\omega) - \frac{1}{2}\int_0^s h(u)^2 \mathrm{d}u\right) \\
&= D_t \exp\left(\int_0^T h(u)\mathbf{1}_{[0,s]}(u)\mathrm{d}B_u(\omega) - \frac{1}{2}\int_0^s h(u)^2 \mathrm{d}u\right) \\
&= \exp\left(\int_0^s h(u)\mathrm{d}B_u(\omega) - \frac{1}{2}\int_0^s h(u)^2 \mathrm{d}u\right) h(t)\mathbf{1}_{[0,s]}(t) \\
&= \mathbb{E}\{F \mid \mathcal{F}_s\} h(t)\mathbf{1}_{[0,s]}(t) \\
&= \mathbb{E}\{F h(t) \mid \mathcal{F}_s\} \mathbf{1}_{[0,s]}(t) \\
&= \mathbb{E}\{D_t F \mid \mathcal{F}_s\} \mathbf{1}_{[0,s]}(t),
\end{aligned}
$$

where we have used that F is a martingale, the chain rule (A.1) and example A.8. The above computation extends to random variables in the linear span of random variables on the form (A.3). Since this linear span is dense in $L^2(\Omega)$ it seems reasonable that the result also holds for more general random variables. Of course, the result does not hold for all $F \in L^2(\Omega)$, since it involves the Malliavin derivative of F, which does not exist for all $F \in L^2(\Omega)$. The result can, however, be proved for all $F \in \mathbb{D}_{1,2}$, as shown in Nualart (1995) or Øksendal (1997).

In particular, if $F \in \mathbb{D}_{1,2}$ is $\mathcal{F}_s$-adapted, we get

$$
\begin{aligned}
D_t F &= D_t \mathbb{E}\{F \mid \mathcal{F}_s\} \\
&= \mathbb{E}\{D_t F \mid \mathcal{F}_s\} \mathbf{1}_{[0,s]}(t).
\end{aligned}
$$

Hence $D_t F$ is $\mathcal{F}_s$-adapted and

$$
D_t F = 0
$$

if $t > s$. □

We end this section with an important result that gives a representation of the integrand from Itô's representation theorem, see e.g. Øksendal (2000, Theorem 4.3.3). This theorem is the cornerstone of the martingale approach to optimal portfolio choice, where the integrand represents the optimal investment strategy. Since the representation theorem only gives the existence of such investment strategy, the following proposition is an important financial application of Malliavin calculus. Even though we will only use this result in example A.11, it has been used in Cvitanić, Goukasian, and Zapatero (2003) to solve optimal portfolio problems in complete markets with Monte-Carlo simulation.

Proposition A.6 (The Clark-Ocone formula)**.** *Let $F \in \mathbb{D}_{1,2}$ be $\mathcal{F}_T$-adapted. Then*

$$
F(\omega) = \mathbb{E}\{F\} + \int_0^T \mathbb{E}\{D_t F \mid \mathcal{F}_t\}(\omega)\mathrm{d}\omega(t). \tag{A.4}
$$

Proof. As in proposition A.5 we just sketch the proof by working with the exponential martingale in equation (A.3). Define the stochastic process F_t by

$$F_t(\omega) \triangleq \exp\left(\int_0^t h(u)\mathrm{d}B_u(\omega) - \frac{1}{2}\int_0^t h(u)^2\mathrm{d}u\right),$$

and let $F \triangleq F_T$. Introducing the auxiliary process $Z_t \triangleq \int_0^t h(u)\mathrm{d}B_u(\omega) - \frac{1}{2}\int_0^t h(u)^2\mathrm{d}u$ the dynamics of F_t is given by Itô's lemma as

$$\begin{aligned}
\mathrm{d}F_t &= F_t\mathrm{d}Z_t + \frac{1}{2}F_t(\mathrm{d}Z_t)^2 \\
&= F_t\left(h(t)\mathrm{d}B_t - \frac{1}{2}h(t)^2\mathrm{d}t\right) + \frac{1}{2}F_t h(t)^2\mathrm{d}t \\
&= F_t h(t)\mathrm{d}B_t.
\end{aligned}$$

Moreover

$$\begin{aligned}
\mathbb{E}\left\{D_t F \mid \mathcal{F}_t\right\} &= \mathbb{E}\left\{F h(t) \mid \mathcal{F}_t\right\} \\
&= \mathbb{E}\left\{F \mid \mathcal{F}_t\right\} h(t) \\
&= F_t h(t),
\end{aligned}$$

so writing the dynamics of F on integral form we obtain

$$\begin{aligned}
F &= \mathbb{E}\left\{F\right\} + \int_0^T F_t h(t)\mathrm{d}B_t \\
&= \mathbb{E}\left\{F\right\} + \int_0^T \mathbb{E}\left\{D_t F \mid \mathcal{F}_t\right\}\mathrm{d}B_t
\end{aligned}$$

since $F_0 = 1 = \mathbb{E}\left\{F\right\}$. We have now proved the result in a special case. Again the result extends to the linear span of random variables on the exponential form and can be extended to $F \in \mathbb{D}_{1,2}$. ☐

A.2.2 Examples

Example A.7. Let $F(\omega) = B_t(\omega) = \omega(t)$. We then have

$$F(\omega + \varepsilon\gamma) = \omega(t) + \varepsilon\gamma(t),$$

and hence

$$
\begin{aligned}
D_\gamma F(\omega) &= \lim_{\varepsilon \to 0} \frac{F(\omega + \varepsilon\gamma) - F(\omega)}{\varepsilon} \\
&= \gamma(t) \\
&= \int_0^t h(s)\mathrm{d}s \\
&= \int_0^T \mathbf{1}_{[0,t]}(s)h(s)\mathrm{d}s.
\end{aligned}
$$

We conclude from definition A.2 that

$$D_s B_t = \mathbf{1}_{[0,t]}(s).$$

If we think of the Malliavin derivative as a perturbation of the underlying Brownian motion, this result also makes sense: when the Brownian path changes at time $s \leq t$, the entire future path will change as well. If we change the Brownian path at time $s > t$, nothing happens at time t, which explains the use of the indicator function.

Example A.8. Let $F(\omega) = \int_0^T f(s)\mathrm{d}B_s(\omega) = \int_0^T f(s)\mathrm{d}\omega(s)$, where $f \in L^2([0,T])$ is a deterministic function. We then have

$$F(\omega + \varepsilon\gamma) = \int_0^T f(s)\mathrm{d}(\omega(s) + \varepsilon\gamma(s)),$$

and hence

$$
\begin{aligned}
D_\gamma F(\omega) &= \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \left[F(\omega + \varepsilon\gamma) - F(\omega)\right] \\
&= \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \left[\int_0^T f(s)\mathrm{d}(\omega(s) + \varepsilon\gamma(s)) - \int_0^T f(s)\mathrm{d}\omega(s)\right] \\
&= \int_0^T f(s)\mathrm{d}\gamma(s) \\
&= \int_0^T f(s)h(s)\mathrm{d}s.
\end{aligned}
$$

We conclude from definition A.2 that

$$D_s F(\omega) = f(s).$$

Choosing $f(s) = \mathbf{1}_{[0,t]}(s)$, we get the previous example.

Example A.9. Let $F(\omega) = f(B_t(\omega)) = f(\omega(t))$ where f is differentiable. By the chain rule (A.1) and example A.7 we then have

$$D_s F(\omega) = f'(B_t(\omega))\mathbf{1}_{[0,s]}(t).$$

Example A.10. Let $F(\omega) = \int_0^T f(B_t(\omega)) \mathrm{d}B_t(\omega) = \int_0^T f(\omega(t)) \mathrm{d}\omega(t)$. We then have

$$\begin{aligned} F(\omega + \varepsilon\gamma) &= \int_0^T f(\omega(t) + \varepsilon\gamma(t)) \mathrm{d}(\omega(t) + \varepsilon\gamma(t)) \\ &= \int_0^T f(\omega(t) + \varepsilon\gamma(t)) \mathrm{d}\omega(t) + \varepsilon \int_0^T f(\omega(t) + \varepsilon\gamma(t)) \mathrm{d}\gamma(t), \end{aligned}$$

and hence

$$\begin{aligned} D_\gamma F(\omega) &= \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \left[F(\omega + \varepsilon\gamma) - F(\omega) \right] \\ &= \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \left[\int_0^T f(\omega(t) + \varepsilon\gamma(t)) \mathrm{d}\omega(t) + \varepsilon \int_0^T f(\omega(t) + \varepsilon\gamma(t)) \mathrm{d}\gamma(t) - \int_0^T f(\omega(t)) \mathrm{d}\omega(t) \right] \\ &= \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \left[\int_0^T \left[f(\omega(t) + \varepsilon\gamma(t)) - f(\omega(t)) \right] \mathrm{d}\omega(t) + \varepsilon \int_0^T f(\omega(t) + \varepsilon\gamma(t)) \mathrm{d}\gamma(t) \right] \\ &= \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \left[\int_0^T \left[f(\omega(t) + \varepsilon\gamma(t)) - f(\omega(t)) \right] \mathrm{d}\omega(t) \right] + \int_0^T f(\omega(t)) \mathrm{d}\gamma(t) \\ &= \int_0^T f'(\omega(t)) \gamma(t) \mathrm{d}\omega(t) + \int_0^T f(\omega(t)) \mathrm{d}\gamma(t) \\ &= \int_0^T f'(\omega(t)) \left(\int_0^t h(s) \mathrm{d}s \right) \mathrm{d}\omega(t) + \int_0^T f(\omega(t)) h(t) \mathrm{d}t \\ &= \int_0^T \left(\int_s^T f'(\omega(t)) \mathrm{d}\omega(t) \right) h(s) \mathrm{d}s + \int_0^T f(\omega(s)) h(s) \mathrm{d}s \\ &= \int_0^T \left(\int_s^T f'(\omega(t)) \mathrm{d}\omega(t) + f(\omega(s)) \right) h(s) \mathrm{d}s \\ &= \int_0^T \left(\int_s^T f'(B_t(\omega)) \mathrm{d}B_t(\omega) + f(B_s(\omega)) \right) h(s) \mathrm{d}s. \end{aligned}$$

We conclude from definition A.2 that

$$D_s \int_0^T f(B_t(\omega)) \mathrm{d}B_t(\omega) = \int_s^T f'(B_t(\omega)) \mathrm{d}B_t(\omega) + f(B_s(\omega)).$$

Example A.11. For $F \in \mathcal{F}$, the indicator function $\mathbf{1}_F \in \mathbb{D}_{1,2}$ if and only if $P(F) \in \{0, 1\}$. First assume that $P(F) \in \{0, 1\}$. Then $\mathbf{1}_F(\omega)$ is constant almost surely, and hence the Malliavin derivative is zero, i.e. $\mathbf{1}_F \in \mathbb{D}_{1,2}$ and $D_t \mathbf{1}_F(\omega) = 0$.

Now assume that $\mathbf{1}_F \in \mathbb{D}_{1,2}$. We need to show that this can only be true if $P(F) \in \{0, 1\}$. To see this remember that

$$\mathbf{1}_F(\omega) = (\mathbf{1}_F(\omega))^2 = \begin{cases} 1, & \text{if } \omega \in F \\ 0, & \text{if } \omega \notin F. \end{cases}$$

Since the mapping $x \to x^2$ is differentiable the chain rule (A.1) yields

$$D_t \mathbf{1}_F(\omega) = D_t \left(\mathbf{1}_F(\omega)\right)^2 = 2\mathbf{1}_F(\omega) D_t \mathbf{1}_F(\omega).$$

For $\omega \in F^c$, we have $\mathbf{1}_F(\omega) = 0$ so

$$D_t \mathbf{1}_F(\omega) = 2 \cdot 0 \cdot D_t \mathbf{1}_F(\omega) = 0,$$

and for $\omega \in F$ we get

$$D_t \mathbf{1}_F(\omega) = 2 D_t \mathbf{1}_F(\omega),$$

which can only be satisfied if $D_t \mathbf{1}_F(\omega) = 0$. Therefore, $D_t \mathbf{1}_F(\omega) = 0$ for all $\omega \in \Omega$. Since, by assumption, $\mathbf{1}_F \in \mathbb{D}_{1,2}$ the Clark-Ocone formula (A.4) gives

$$\begin{aligned} 1_F(\omega) &= \mathbb{E}\left\{1_F\right\} + \int_0^T \mathbb{E}\left\{D_t 1_F \mid \mathcal{F}_t\right\}(\omega) \mathrm{d}B_t \\ &= P(F) + \int_0^T \mathbb{E}\left\{0 \mid \mathcal{F}_t\right\}(\omega) \mathrm{d}B_t \\ &= P(F), \end{aligned}$$

which can only be true if $P(F) \in \{0, 1\}$. Note, that we have also proved that $\mathbb{D}_{1,2} \subsetneqq L^2(\Omega)$.

A.3 Integration

Is is well-known that the Itô integral is well-defined only for a limited class of processes. The processes are for example required to be adapted to a filtration, which the Brownian motion is a martingale with respect to. In this section we introduce the Skorohod integral, which allows us to integrate more general processes. In order to distinguish the two stochastic integrals, the Itô integral will be denoted $\int_0^T h_t \mathrm{d}B_t$ as usual whereas the Skorohod integral will be denoted $S(h) \triangleq \int_0^T h_t \delta B_t$. Sometimes we will also use the notation $S^h(F)$ for the Skorohod integral $S(Fh)$.

A.3.1 Integration by Parts on the Wiener Space

Proposition A.12 (Integration by parts). *Let $F, G \in \mathcal{D}_{1,2}$ and define $\gamma(t) = \int_0^t h_s \mathrm{d}s$ for $h \in L^2([0,T])$. Then*

$$\mathbb{E}\left\{\int_0^T (D_t F) h_t \mathrm{d}t\right\} = \mathbb{E}\left\{F \int_0^T h_t \mathrm{d}B_t\right\}. \tag{A.5}$$

Proof. By definition of the directional derivative we get

$$
\begin{aligned}
\mathbb{E}\left\{\int_0^T (D_tF)h_t\mathrm{d}t\right\} &= \int_\Omega D_\gamma F(\omega)\mathrm{d}P(\omega) \\
&= \int_\Omega \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\left[F(\omega+\varepsilon\gamma)-F(\omega)\right]\mathrm{d}P(\omega) \\
&= \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\int_\Omega \left[F(\omega+\varepsilon\gamma)-F(\omega)\right]\mathrm{d}P(\omega) \\
&= \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\left[\int_\Omega F(\omega+\varepsilon\gamma)\mathrm{d}P(\omega)-\int_\Omega F(\omega)\mathrm{d}P(\omega)\right]. \qquad \text{(A.6)}
\end{aligned}
$$

Since $h\in L^2([0,T])$, εh satisfies Novikov's condition which ensures that

$$
M_t \triangleq \exp\left(-\varepsilon\int_0^t h_s\mathrm{d}B_s-\frac{1}{2}\varepsilon^2\int_0^t h_s^2\mathrm{d}s\right)
$$

is a P-martingale. By the Girsanov theorem $\tilde{B}_t \triangleq B_t+\varepsilon\int_0^t h_s\mathrm{d}s$ is a Brownian motion under the measure $\tilde{P}$ defined by

$$
\begin{aligned}
\frac{\mathrm{d}\tilde{P}}{\mathrm{d}P} &= M_T \\
&= \exp\left(-\varepsilon\int_0^T h_s\mathrm{d}B_s-\frac{1}{2}\varepsilon^2\int_0^T h_s^2\mathrm{d}s\right) \\
&= \exp\left(-\varepsilon\int_0^T h_s\mathrm{d}\tilde{B}_s+\frac{1}{2}\varepsilon^2\int_0^T h_s^2\mathrm{d}s\right).
\end{aligned}
$$

We can now write the first integral in equation (A.6) as

$$
\begin{aligned}
\int_\Omega F(\omega+\varepsilon\gamma)\mathrm{d}P(\omega) &= \int_\Omega F(\tilde{\omega})\exp\left(\varepsilon\int_0^T h_s\mathrm{d}\tilde{B}_s(\omega)-\frac{1}{2}\varepsilon^2\int_0^T h_s^2\mathrm{d}s\right)\mathrm{d}\tilde{P}(\omega) \\
&= \int_\Omega F(\omega)\exp\left(\varepsilon\int_0^T h_s\mathrm{d}B_s(\omega)-\frac{1}{2}\varepsilon^2\int_0^T h_s^2\mathrm{d}s\right)\mathrm{d}P(\omega),
\end{aligned}
$$

so we get

$$
\begin{aligned}
\mathbb{E}\left\{\int_0^T (D_tF)h_t\mathrm{d}t\right\} &= \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\left[\int_\Omega F(\omega+\varepsilon\gamma)\mathrm{d}P(\omega) - \int_\Omega F(\omega)\mathrm{d}P(\omega)\right] \\
&= \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\int_\Omega F(\omega)\left[\exp\left(\varepsilon\int_0^T h_s\mathrm{d}B_s(\omega) - \frac{1}{2}\varepsilon^2\int_0^T h_s^2\mathrm{d}s\right) - 1\right]\mathrm{d}P(\omega) \\
&= \int_\Omega F(\omega)\lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\left[\exp\left(\varepsilon\int_0^T h_s\mathrm{d}B_s(\omega) - \frac{1}{2}\varepsilon^2\int_0^T h_s^2\mathrm{d}s\right) - 1\right]\mathrm{d}P(\omega) \\
&= \int_\Omega F(\omega)\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\left[\exp\left(\varepsilon\int_0^T h_s\mathrm{d}B_s(\omega) - \frac{1}{2}\varepsilon^2\int_0^T h_s^2\mathrm{d}s\right)\right]_{\varepsilon=0}\mathrm{d}P(\omega) \\
&= \int_\Omega \left[F(\omega)\int_0^T h_s\mathrm{d}B_s(\omega)\right]\mathrm{d}P(\omega) \\
&= \mathbb{E}\left\{F\int_0^T h_t\mathrm{d}B_t\right\}. \qquad \square
\end{aligned}
$$

Corollary A.13. *Let $F, G \in \mathcal{D}_{1,2}$ and define $\gamma(t) = \int_0^t h_s\mathrm{d}s$ for $h \in L^2([0,T])$. Then*

$$
\mathbb{E}\left\{G\int_0^T (D_tF)h_t\mathrm{d}t\right\} = \mathbb{E}\left\{FG\int_0^T h_t\mathrm{d}B_t\right\} - \mathbb{E}\left\{F\int_0^T (D_tG)h_t\mathrm{d}t\right\}.
$$

Proof. By the product rule (A.2) we have that $FG \in \mathcal{D}_{1,2}$ and that

$$
D_t(FG) = FD_tG + GD_tF.
$$

Using the above proposition with "$F = FG$", we get

$$
\begin{aligned}
\mathbb{E}\left\{FG\int_0^T h_t\mathrm{d}B_t\right\} &= \mathbb{E}\left\{\int_0^T (D_tFG)h_t\mathrm{d}t\right\} \\
&= \mathbb{E}\left\{F\int_0^T (D_tG)h_t\mathrm{d}t\right\} + \mathbb{E}\left\{G\int_0^T (D_tF)h_t\mathrm{d}t\right\}. \qquad \square
\end{aligned}
$$

A.3.2 The Skorohod Integral and its Properties

As shown in e.g. Øksendal (1997), the Skorohod integral of a stochastic process can be constructed from the Wiener-Itô chaos expansion. However, it turns out that the Skorohod integral coincides with the adjoint operator of the Malliavin differential operator, i.e. for a Skorohod integrable process h_t and a Malliavin differentiable random variable F the Skorohod integral, $S(h)$, is defined as

$$
\langle F, S(h)\rangle_{L^2(\Omega)} = \langle D_tF, h\rangle_{L^2([0,T]\times\Omega)},
$$

where $\langle\cdot,\cdot\rangle$ denotes inner product. Hence we arrive at the following definition, in which we omitted the technical condition for a process to be Skorohod integrable.

Definition A.14 (Nualart (1995), Definition 1.3.1)**.** If h is Skorohod integrable, we define the

Skorohod integral of h, as the element $S(h) \in L^2(\Omega)$ that satisfies

$$\mathbb{E}\{FS(h)\} = \mathbb{E}\left\{\int_0^T (D_tF)h_t\mathrm{d}t\right\} \tag{A.7}$$

for all $F \in \mathbb{D}_{1,2}$.

Proposition A.15. *If h_t is $\mathcal{F}_t$-adapted, the Skorohod integral coincides with the Itô integral when defined, i.e.*

$$\int_0^T h_t\delta B_t = \int_0^T h_t\mathrm{d}B_t.$$

Proof. We will only prove the result for deterministic $L^2([0,T])$ functions, h_t. Let $F, G \in \mathbb{D}_{1,2}$ be two Malliavin differentiable random variables. Using the definition of the Skorohod integral and the integration by parts property in corollary A.13, we get

$$\begin{aligned}
\mathbb{E}\{GS(Fh)\} &= \mathbb{E}\left\{\int_0^T (D_tG)Fh_t\mathrm{d}t\right\} \\
&= \mathbb{E}\left\{F\int_0^T (D_tG)h_t\mathrm{d}t\right\} \\
&= \mathbb{E}\left\{GF\int_0^T h_t\mathrm{d}B_t\right\} - \mathbb{E}\left\{G\int_0^T (D_tF)h_t\mathrm{d}t\right\} \\
&= \mathbb{E}\left\{G\left(F\int_0^T h_t\mathrm{d}B_t - \int_0^T (D_tF)h_t\mathrm{d}t\right)\right\}.
\end{aligned}$$

Since this must hold for all $G \in \mathbb{D}_{1,2}$, an inner product argument gives

$$S(Fh) = F\int_0^T h_t\mathrm{d}B_t - \int_0^T (D_tF)h_t\mathrm{d}t.$$

With $F = 1$ we see that the Itô integral and the Skorohod integral coincide for deterministic $L^2([0,T])$ functions. If we take F to be $\mathcal{F}_s$-measurable and $h_t = \mathbf{1}_{(s,u]}(t)$ we obtain

$$\begin{aligned}
S(Fh) &= F\int_0^T h_t\mathrm{d}B_t - \int_0^T (D_tF)h_t\mathrm{d}t \\
&= F\int_s^u \mathrm{d}B_t - \int_s^u (D_tF)\mathrm{d}t \\
&= F(B_u - B_s) - 0,
\end{aligned}$$

where we have used proposition A.5 to get that $D_tF = 0$ for $t > s$ since F is $\mathcal{F}_s$-measurable. Since we can write

$$Fh_t = 0 \cdot \mathbf{1}_{[0,s]}(t) + F\mathbf{1}_{(s,u]}(t) + 0 \cdot \mathbf{1}_{(u,T]}(t),$$

we see that Fh_t is an elementary process and hence

$$\begin{aligned}\int_0^T Fh_t \mathrm{d}B_t &= 0 \cdot (B_s - B_0) + F(B_u - B_s) + 0 \cdot (B_T - B_u) \\ &= F(B_u - B_s).\end{aligned}$$

Therefore the Skorohod integral and Itô integral also coincide in this case, i.e.

$$S(Fh) = \int_0^T Fh_t \mathrm{d}B_t.$$

To show the result for any $\mathcal{F}_t$-adapted process $h_t \in L_2([0,T] \times \Omega)$, one can use an approximation argument. □

Proposition A.16. *Let F be a Malliavin differentiable random variable. Then*

$$\int_0^T Fh_t \delta B_t = F \int_0^T h_t \delta B_t - \int_0^T (D_t F) h_t \mathrm{d}t,$$

where h_t is Skorohod integrable. If h_t is $\mathcal{F}_t$-adapted we get

$$\int_0^T Fh_t \delta B_t = F \int_0^T h_t \mathrm{d}B_t - \int_0^T (D_t F) h_t \mathrm{d}t.$$

Proof. Let G be a Malliavin differentiable random variable. Using the product rule (A.2) and integration by parts (A.5) we obtain

$$\begin{aligned}\mathbb{E}\left\{\int_0^T (D_t G) F h_t \mathrm{d}t\right\} &= \mathbb{E}\left\{\int_0^T \left[(D_t GF) - G(D_t F)\right] h_t \mathrm{d}t\right\} \\ &= \mathbb{E}\left\{GF \int_0^T h_t \delta B_t - G \int_0^T (D_t F) h_t \mathrm{d}t\right\} \\ &= \mathbb{E}\left\{G\left[F \int_0^T h_t \delta B_t - \int_0^T (D_t F) h_t \mathrm{d}t\right]\right\}.\end{aligned}$$

Applying the definition of the Skorohod integral in equation (A.7) to the left hand side we have

$$\mathbb{E}\left\{G \int_0^T F h_t \mathrm{d}B_t\right\} = \mathbb{E}\left\{G\left[F \int_0^T h_t \delta B_t - \int_0^T (D_t F) h_t \mathrm{d}t\right]\right\},$$

and since this should be true for any choice of G, the result follows by an inner product argument. The last statement follows from the first statement and proposition A.15. □

A.3.3 Examples

Example A.17. With $h_t = B_t$ we have

$$\int_0^T B_t \delta B_t = \int_0^T B_t \mathrm{d}B_t$$
$$= \frac{1}{2}B_T^2 - \frac{1}{2}T.$$

since B_t is $\mathcal{F}_t$-adapted and the Skorohod integral coincides with the Itô integral cf. proposition A.15.

Example A.18. With $h_t = 1$ and $F = B_T$ in proposition A.16 we have

$$\begin{aligned}\int_0^T B_T \delta B_t &= B_T \int_0^T \delta B_t - \int_0^T (D_t B_T)\mathrm{d}t \\ &= B_T(B_T - B_0) - \int_0^T \mathbf{1}_{[0,T]}(t)\mathrm{d}t \\ &= B_T^2 - T,\end{aligned}$$

where we have used example A.7 to compute the Malliavin derivative of B_T.

Example A.19. With $h_t = 1$ and $F = B_{t_{n+1}} - B_{t_n}$ in proposition A.16 we have

$$\begin{aligned}\int_{t_n}^{t_{n+1}} (B_{t_{n+1}} - B_{t_n})\delta B_t &= (B_{t_{n+1}} - B_{t_n}) \int_{t_n}^{t_{n+1}} \delta B_t - \int_{t_n}^{t_{n+1}} (D_t(B_{t_{n+1}} - B_{t_n}))\mathrm{d}t \\ &= (B_{t_{n+1}} - B_{t_n})^2 - \int_{t_n}^{t_{n+1}} \mathbf{1}_{(t_n,t_{n+1}]}\mathrm{d}t \\ &= (B_{t_{n+1}} - B_{t_n})^2 - (t_{n+1} - t_n),\end{aligned}$$

where we have used example A.7 to determine the Malliavin derivative of $B_{t_{n+1}} - B_{t_n}$.

A.4 Conditional Expectations

In this section we will show how a conditional expectation can be computed from unconditional expectations. In the following we let $\mathbb{H}_n$ denote the set of processes h_t^n satisfying

$$\int_0^{t_{n+1}} (D_t W_{t_n}) h_t^n \mathrm{d}t = 1 \quad \text{and} \quad \int_0^{t_{n+1}} (D_t W_{t_{n+1}}) h_t^n \mathrm{d}t = 0. \tag{A.8}$$

Definition A.20 (Localizing functions)**.** Let $\varphi : \mathbb{R} \to \mathbb{R}$ be a continuous and bounded mapping. If $\varphi(0) = 1$ and φ' is continuous and bounded, we say that φ is a *smooth localizing function.* The set of such localizing functions is denoted by $\mathcal{L}$.

Remark. Since φ is smooth, $\varphi(W_{t_n}-w)$ is Malliavin differentiable. The chain rule (A.1) gives

$$D_t\varphi(W_{t_n}-w)=\varphi'(W_{t_n}-w)D_tW_{t_n}.$$

Furthermore, proposition A.16 with $F=\varphi(W_{t_n}-w)$ gives

$$\begin{aligned}S^{h^n}(\varphi(W_{t_n}-w)) &= \int_0^T \varphi(W_{t_n}-w)h_t^n\delta B_t\\ &= \varphi(W_{t_n}-w)\int_0^T h_t^n\delta B_t-\int_0^T(D_t\varphi(W_{t_n}-w))h_t^n\mathrm{d}t\\ &= \varphi(W_{t_n}-w)S^{h^n}(1)-\int_0^T\varphi'(W_{t_n}-w)(D_tW_{t_n})h_t^n\mathrm{d}t\\ &= \varphi(W_{t_n}-w)S^{h^n}(1)-\varphi'(W_{t_n}-w) \end{aligned}\tag{A.9}$$

since $\varphi'(W_{t_n}-w)$ does not depend on t, $h_t^n\in\mathbb{H}_n$ and $D_tW_{t_n}=0$ for $t>t_{n+1}$ cf. proposition A.5.

The main theorem 3.1 is based on the following lemma.

Lemma A.21 (Bouchard, Ekeland, and Touzi (2004), Theorem 3.1)**.** *Let f be a mapping from $\mathbb{R}$ into $\mathbb{R}$ with $f(W_{t_{n+1}})\in L^2(\Omega)$, and $A\in\mathcal{B}$ a Borel subset of $\mathbb{R}$. Then, for all $h^n\in\mathbb{H}_n$, and $\varphi\in\mathcal{L}$*

$$\mathbb{E}\left\{\mathbf{1}_A(W_{t_n})f(W_{t_{n+1}})\right\}=\int_A\mathbb{E}\left\{H_w(W_{t_n})f(W_{t_{n+1}})S^h(\varphi(W_{t_n}-w))\right\}\mathrm{d}w$$

where $H_w(W_{t_n})=\mathbf{1}_{[w,\infty)}(W_{t_n})$ is the indicator function.

Proof. Since the expectation on the right-hand side involves a Skorohod integral, we are tempted to use the Malliavin integration by parts formula at some point in the proof. This requires that we compute the Malliavin derivative of what is multiplied with the Skorohod integral. As shown in example A.11, the indicator function is not Malliavin differentiable unless it is equal to zero or one almost surely.[2] To overcome this problem we define the function

$$F(w_1,w_2)\triangleq\int_{-\infty}^{w_1}\mathbf{1}_A(w)f(w_2)\varphi(w_1-w)\mathrm{d}w,$$

and consider the random variable

$$F(W_{t_n},W_{t_{n+1}})=\int_{-\infty}^{W_{t_n}}\mathbf{1}_A(w)f(W_{t_{n+1}})\varphi(W_{t_n}-w)\mathrm{d}w.$$

The Malliavin derivative of F is

$$D_tF=\frac{\partial F(W_{t_n},W_{t_{n+1}})}{\partial w_1}(D_tW_{t_n})+\frac{\partial F(W_{t_n},W_{t_{n+1}})}{\partial w_2}(D_tW_{t_{n+1}}).$$

[2] Actually, the indicator function in example A.11 is defined for sets $F\in\mathcal{F}$ whereas the indicator function in the lemma is defined for intervals $I\in\mathcal{B}$. However, since $\mathbf{1}_I(W_{t_n}(\omega))=\mathbf{1}_{W_{t_n}^{-1}(I)}(\omega)$, the problem is the same as in example A.11.

Multiplying both sides with $h_t^n \in \mathbb{H}_n$ and integrating from 0 to T we get

$$\begin{aligned}\int_0^T (D_t F) h_t^n \mathrm{d}t &= \frac{\partial F(W_{t_n}, W_{t_{n+1}})}{\partial w_1} \int_0^T (D_t W_{t_n}) h_t^n \mathrm{d}t + \frac{\partial F(W_{t_n}, W_{t_{n+1}})}{\partial w_2} \int_0^T (D_t W_{t_{n+1}}) h_t^n \mathrm{d}t \\ &= \frac{\partial F(W_{t_n}, W_{t_{n+1}})}{\partial w_1} \\ &= \mathbf{1}_A(W_{t_n}) f(W_{t_{n+1}}) + \int_{-\infty}^{W_{t_n}} \mathbf{1}_A(w) f(W_{t_{n+1}}) \varphi'(W_{t_n} - w) \mathrm{d}w,\end{aligned}$$

since the partial derivatives of F with respect to w_1 is

$$\begin{aligned}\frac{\partial F(w_1, w_2)}{\partial w_1} &= \mathbf{1}_A(w_1) f(w_2) \varphi(w_1 - w_1) + \int_{-\infty}^{w_1} \mathbf{1}_A(w) f(w_2) \varphi'(w_1 - w) \mathrm{d}w \\ &= \mathbf{1}_A(w_1) f(w_2) + \int_{-\infty}^{w_1} \mathbf{1}_A(w) f(w_2) \varphi'(w_1 - w) \mathrm{d}w\end{aligned}$$

where the first equality follows from Leibnitz' rule.[3]

If we take expectations on both sides and rearrange the terms, we are left with

$$\begin{aligned}\mathbb{E}\left\{\mathbf{1}_A(W_{t_n}) f(W_{t_{n+1}})\right\} &= \mathbb{E}\left\{\int_0^T (D_t F) h_t^n \mathrm{d}t\right\} \\ &\quad - \mathbb{E}\left\{\int_{-\infty}^{W_{t_n}} \mathbf{1}_A(w) f(W_{t_{n+1}}) \varphi'(W_{t_n} - w) \mathrm{d}w\right\}.\end{aligned}$$

If we apply the Malliavin integration by parts formula to the first term on the right-hand side, we get

$$\begin{aligned}\mathbb{E}\left\{\int_0^T (D_t F) h_t^n \mathrm{d}t\right\} &= \mathbb{E}\left\{F \int_0^T h_t^n \delta B_t\right\} \\ &= \mathbb{E}\left\{F S^h(1)\right\} \\ &= \mathbb{E}\left\{\left(\int_{-\infty}^{W_{t_n}} \mathbf{1}_A(w) f(W_{t_{n+1}}) \varphi(W_{t_n} - w) \mathrm{d}w\right) S^h(1)\right\} \\ &= \mathbb{E}\left\{\int_{-\infty}^{W_{t_n}} \mathbf{1}_A(w) f(W_{t_{n+1}}) \varphi(W_{t_n} - w) S^h(1) \mathrm{d}w\right\}.\end{aligned}$$

[3]Leibnitz' rule: $\frac{\partial}{\partial x} \int_{f(x)}^{g(x)} h(x,y) \mathrm{d}y = h(x, g(x)) \frac{\partial g(x)}{\partial x} - h(x, f(x)) \frac{\partial f(x)}{\partial x} + \int_{f(x)}^{g(x)} \frac{\partial h(x,y)}{\partial x} \mathrm{d}y$.

Substituting this expression into the former we get

$$
\begin{aligned}
\mathbb{E}\left\{\mathbf{1}_A(W_{t_n})f(W_{t_{n+1}})\right\} &= \mathbb{E}\left\{\int_{-\infty}^{W_{t_n}} \mathbf{1}_A(w)f(W_{t_{n+1}})\varphi(W_{t_n}-w)S^h(1)\mathrm{d}w\right\} \\
&\quad - \mathbb{E}\left\{\int_{-\infty}^{W_{t_n}} \mathbf{1}_A(w)f(W_{t_{n+1}})\varphi'(W_{t_n}-w)\mathrm{d}w\right\} \\
&= \mathbb{E}\left\{\int_{-\infty}^{W_{t_n}} \mathbf{1}_A(w)f(W_{t_{n+1}})\left[\varphi(W_{t_n}-w)S^h(1)-\varphi'(W_{t_n}-w)\right]\mathrm{d}w\right\} \\
&= \mathbb{E}\left\{\int_{-\infty}^{W_{t_n}} \mathbf{1}_A(w)f(W_{t_{n+1}})S^h(\varphi(W_{t_n}-w))\mathrm{d}w\right\} \\
&= \mathbb{E}\left\{\int_{\mathbb{R}} \mathbf{1}_A(w)\mathbf{1}_{(-\infty,W_{t_n}]}(w)f(W_{t_{n+1}})S^h(\varphi(W_{t_n}-w))\mathrm{d}w\right\} \\
&= \mathbb{E}\left\{\int_{A} \mathbf{1}_{(-\infty,W_{t_n}]}(w)f(W_{t_{n+1}})S^h(\varphi(W_{t_n}-w))\mathrm{d}w\right\} \\
&= \mathbb{E}\left\{\int_{A} H_w(W_{t_n})f(W_{t_{n+1}})S^h(\varphi(W_{t_n}-w))\mathrm{d}w\right\} \\
&= \int_A \mathbb{E}\left\{H_w(W_{t_n})f(W_{t_{n+1}})S^h(\varphi(W_{t_n}-w))\right\}\mathrm{d}w,
\end{aligned}
$$

where we have used equation (A.9) in the third equality. □

As a direct consequence of lemma A.21 we can now prove theorem 3.1.

Proof of theorem 3.1. Consider the conditional expectation

$$\mathbb{E}\left\{g(W_{t_{n+1}}) \mid \mathcal{M}_{t_n}\right\},$$

where $\mathcal{M}_t \triangleq \sigma(W_s \mid 0 \le s \le t)$ is the σ-algebra generated by $(W_s)_{s\in[0,t]}$. By construction this conditional expectation is $\mathcal{M}_{t_n}$-measurable. Hence the Doob-Dynkin lemma ensures the existence of a Borel measurable function $G : \mathbb{R} \to \mathbb{R}$ such that

$$\mathbb{E}\left\{g(W_{t_{n+1}}) \mid \mathcal{M}_{t_n}\right\} = G(W_{t_n}).$$

The function $w \to G(w)$ is actually the conditional expectation $\mathbb{E}\left\{g(W_{t_{n+1}}) \mid W_{t_n} = w\right\}$ we are looking for. Let $A \in \mathcal{B}$ be a Borel measurable set. Then $(W_{t_n} \in A) \in \mathcal{M}_{t_n}$ and by definition of conditional expectations

$$
\begin{aligned}
\int_{(W_{t_n}\in A)} g(W_{t_{n+1}})\mathrm{d}P &= \int_{(W_{t_n}\in A)} G(W_{t_n})\mathrm{d}P \\
&= \int_A G(w)\mathrm{d}\hat{P}(w) \\
&= \int_A G(w)p(w)\mathrm{d}w,
\end{aligned}
$$

by the integral transformation theorem, where $\hat{P} = W_{t_n}(P)$ is the distribution of W_{t_n} and p is the density of W_{t_n}. The last equality follows since W_{t_n} has a density function.

From lemma A.21 we have

$$\begin{aligned}\int_{(W_{t_n}\in A)} g(W_{t_{n+1}})\mathrm{d}P &= \mathbb{E}\left\{\mathbf{1}_A(W_{t_n})g(W_{t_{n+1}})\right\} \\ &= \int_A \mathbb{E}\left\{H_w(W_{t_n})g(W_{t_{n+1}})S^h(\varphi(W_{t_n}-w))\right\}\mathrm{d}w\end{aligned}$$

and therefore

$$\int_A G(w)p(w)\mathrm{d}w = \int_A \mathbb{E}\left\{H_w(W_{t_n})g(W_{t_{n+1}})S^h(\varphi(W_{t_n}-w))\right\}\mathrm{d}w$$

for all $A \in \mathcal{B}$. We therefore conclude that

$$\begin{aligned}G(w) &= \frac{\mathbb{E}\left\{H_w(W_{t_n})g(W_{t_{n+1}})S^h(\varphi(W_{t_n}-w))\right\}}{p(w)} \\ &= \frac{\mathbb{E}\left\{H_w(W_{t_n})g(W_{t_{n+1}})S^h(\varphi(W_{t_n}-w))\right\}}{\mathbb{E}\left\{H_w(W_{t_n})S^h(\varphi(W_{t_n}-w))\right\}},\end{aligned}$$

where we have used lemma A.21 with $f = 1$ to find the density of W_{t_n}. $\square$

References

Bouchard, B., I. Ekeland, and N. Touzi (2004). On the Malliavin approach to Monte Carlo approximation of conditional expectations. *Finance and Stochastics 8*(1), 45–71.

Brandt, M. W., A. Goyal, P. Santa-Clara, and J. R. Stroud (2005). A Simulation Approach to Dynamic Portfolio Choice with an Application to Learning About Return Predictability. *Review of Financial Studies*.

Campbell, J. Y. and J. F. Cocco (2003). Household Risk Management and Optimal Mortgage Choice. *The Quarterly Journal of Economics 118*(4), 1449–1494.

Cvitanić, J., L. Goukasian, and F. Zapatero (2003). Monte Carlo computation of optimal portfolios in complete markets. *Journal of Economic Dynamics and Control 27*(6), 971–986.

Detemple, J. B., R. Garcia, and M. Rindisbacher (2003). A Monte-Carlo Method for Optimal Portfolios. *The Journal of Finance 58*(1), 401–446.

Duffie, D. (2001). *Dynamic Asset Pricing Theory* (Third ed.). Princeton University Press.

Fournié, E., J.-M. Lasry, J. Lebuchoux, and P.-L. Lions (2001). Applications of Malliavin calculus to Monte Carlo methods in finance. II. *Finance and Stochastics 5*(2), 201–236.

Friz, P. K. (2005). An Introduction to Malliavin Calculus. Lecture notes, `http://www.statslab.cam.ac.uk/~peter/malliavin/Malliavin2005/mall.pdf`.

Kushner, H. J. and P. Dupuis (2001). *Numerical Methods for Stochastic Control Problems in Continuous Time* (Second ed.). Springer.

Longstaff, F. A. (2001). Optimal portfolio choice and the valuation of illiquid securities. *The Review of Financial Studies 14*(2), 407–431.

Longstaff, F. A. and E. S. Schwartz (2001). Valuing American options by simulation: a simple least-squares approach. *The Review of Financial Studies 14*(1), 113–147.

Merton, R. C. (1969). Lifetime Portfolio Selection under Uncertainty: The Continuous-Time Case. *The Review of Economics and Statistics 51*(3), 247–257.

Munk, C. (2003). The Markov chain approximation approach for numerical solution of stochastic control problems: experiences from Merton's problem. *Applied Mathematics and Computation 136*(1), 47–77.

Nualart, D. (1995). *The Malliavin Calculus and Related Topics*. Springer-Verlag.

Pagès, G., H. Pham, and J. Printems (2004). An Optimal Markovian Quantization Algorithm for Multidimensional Stochastic Control Problems. *Stochastics and Dynamics 4*(4), 501–545.

Samuelson, P. A. (1969). Lifetime Portfolio Selection By Dynamic Stochastic Programming. *The Review of Economics and Statistics 51*(3), 239–246.

Øksendal, B. (1997). An Introduction to Malliavin Calculus with Applications to Economics. Lecture notes, `http://www.nhh.no/for/dp/1996/wp0396.pdf`.

Øksendal, B. (2000). *Stochastic Differential Equations. An Introduction with Applications* (Fifth ed.). Springer-Verlag. Corrected Second Printing.