# Information and Memory in Stochastic Optimal Control

Kristina Rognlien Dahl

Thesis presented for the degree of Philosophiae Doctor

Department of Mathematics
University of Oslo

2016

## Acknowledgements

First and foremost, I would like to thank my supervisor Bernt Øksendal. Bernt is the kindest and most humble mathematician I have met, and I am so grateful for his help and support. He is a source of inspiration to me and the rest of the mathematics community. Always positive and encouraging, motivating me to use my creativity, Bernt is a wonderful guide to mathematics.

Thanks to Fred Espen Benth, my co-supervisor, for help whenever I have needed it and for the opportunity to teach his class and review papers from time to time. Also, thanks to the stochastic analysis group at the Department of Mathematics for interesting seminars.

I would also like to thank the Department of Mathematics, University of Oslo, for funding this thesis. Thanks to the administration and IT-team for all their assistance, and thanks for the best office view anyone could hope for!

Thanks to Inger Christin Borge, Nils Voie Johansen and Tom Lindstrøm for including me in lots of fun teaching! To all of my students the past four years: Thank you for your questions, your $\epsilon$-$\delta$-frustrations and your kind words.

A huge thanks goes to my fellow PhDs and post-docs who have helped make these four years enjoyable. A special thanks to Espen, Torkel and Johanna, who have given me a good laugh whenever I needed it!

Thanks to all my friends for their love and support. In particular, I would like to thank the smart, witty and truly inspirational girls from my undergrad mathematics classes: Rebekka, Rebecca, Karen, Marianne, Caroline and Rucha. Without you, I would never have come this far.

I would like to thank my family: my mom Heidi, my dad Geir (a special thanks for the mathematical discussions), my grandmother Berit and my brother Eirik. Also, thanks to my bonus-family: Silje, Anna, Ola and Vilde. Thank you all for your everlasting support. For those of you who have asked from time to time "but what are you actually doing?": Here it is!

Finally, to the person who deals with me crying over proofs that collapse and unreadable papers, who listens to math jokes on a regular basis and who always has my back: My dearest Lars, thank you so much. Perhaps I could have done this without you, but it has been so much better with you by my side.

**Tusen takk!**

# Contents

Chapter 1

# Introduction

The purpose of this introduction is to provide an overview of the topics of the thesis and some relevant stochastic analysis theory. Our presentation is brief and non-technical, but focuses on central concepts and mathematical approaches. Moreover, in Section 1.1, we present the different papers of the thesis in some detail.

The world is an uncertain place. All sciences aim to describe and understand our universe in some way. Mathematics has proven to be an essential tool in describing our world. However, for a long time, the mathematical foundation of fields such as physics, economics and biology consisted of deterministic models. During the last decades, researchers have started incorporating stochasticity into their models, due to the uncertain nature of the world. This growth in applications of stochastic models is due to the realization that deterministic modeling is not always enough, but also due to the vast development there has been in the fields of stochastic analysis, mathematical theory and computational tools over the last five-six decades.

Roughly speaking, *stochastic analysis* is the mathematical study of uncertain processes developing in time. More precisely, we study processes $X(t,\omega)$ where $t$ is the time and $\omega$ is some potential *scenario* (or outcome) in a *scenario space* $\Omega$. These processes may be studied for discrete or continuous time, and finite or arbitrary scenario space $\Omega$. A simple example of a discrete stochastic process is tossing a coin three times, each toss resulting in either heads (h) or tails (ta). Then, the time $t \in \{1,2,3\}$ and $\Omega$ is the set of all combinations of outcomes of the three throws, i.e.,

$$\begin{aligned}\Omega &= \{(h,h,h),(h,h,ta),(h,ta,h),(h,ta,ta),\\ &\quad (ta,h,h),(ta,h,ta),(ta,ta,h),(ta,ta,ta)\}.\end{aligned}$$

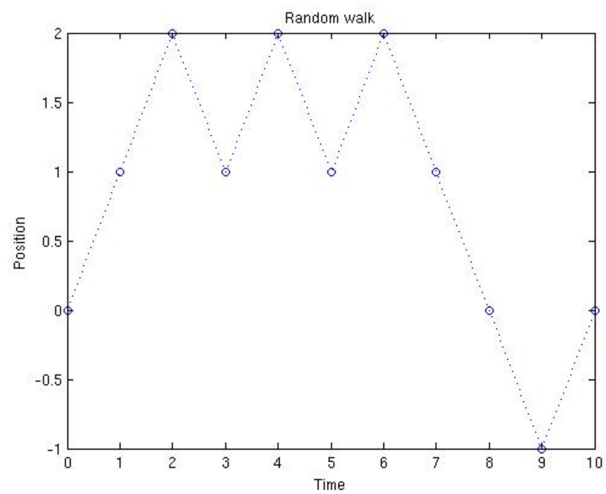


Figure 1.1: A random walk

A basic, and important, example of a continuous time stochastic process is the *Brownian motion* (also called the Wiener process), usually denoted $B(t,\omega)$ (or $W(t,\omega)$). This process has some very attractive properties: The increments $B(t)-B(s)$ and $B(s)-B(u)$, for all $u < s < t$, are independent, normally distributed random variables with expectation 0 and variance $t-s$ and $s-u$, respectively. Roughly, the Brownian motion is the limit as time steps converge towards 0 of a random walk (i.e., a discrete stochastic process where there is a 50/50 chance of going up or down an amount depending on the step size at each time step, see Figure 1.1). For an illustration of a path of a Brownian motion (i.e., the Brownian motion viewed as a function of time for one particular scenario $\omega in \Omega$), see Figure 1.2. One can prove the existence of Brownian motion using the Kolmogorov extension theorem, and one can also prove that Brownian motion has a continuous version (from Kolmogorov's continuity theorem).

It is possible to define a stochastic integral of a function $f(t,\omega)$ (satisfying certain measurability- and boundedness conditions) with respect to the Brownian motion:

$$\int_S^T f(t,\omega)dB(t,\omega).$$

This integral, called the *Itô integral*, is first defined for stochastic step-functions (simple functions) and then extended to all functions $f$ which are adapted, measurable (wrt. the filtration generated by the Brownian motion) and in $L^2$. The Itô integral has nice properties such as linearity, additivity and measurability, as well as having expectation 0. One can also prove the *Itô isometry*

$$E[(\int_S^T f(t,\omega)dB(t,\omega))^2] = E[\int_S^T f^2(t,\omega)dt]$$

and the very important *Itô formula*, which is a sort of chain rule for Itô integrals. Based on this, one can study *stochastic differential equations* (SDEs) of the form:

$$X(T) = x + \int_S^T b(t,X(t))dt + \int_S^T \sigma(t,X(t))dB(t)$$

or in differential form

$$\begin{aligned} dX(t) &= b(t,X(t))dt + \sigma(t,X(t))dB(t),\ t \in [S,T] \\ X(S) &= x \end{aligned} \tag{1.1}$$

where $x \in \mathbb{R}$ and the $\omega$'s have been suppressed from the notation for readability. Such equations can, in some cases, be solved analytically for the stochastic solution process $X(t)$, and there are results on the existence and uniqueness of solution of such equations. However, in some (actually, most) cases, SDEs cannot be solved analytically, but one can still find an approximate solution numerically. Numerical methods for SDE's are similar to numerical methods for ordinary differential equations. For instance, one has (pathwise) versions of Euler and Runge Kutta methods, which can be combined with many Monte Carlo simulations of the sample paths. See Chapter 8 for an Euler method for a special type of SDE with so-called noisy memory.

During the last few years, and in particular after the 2008 financial crisis, the

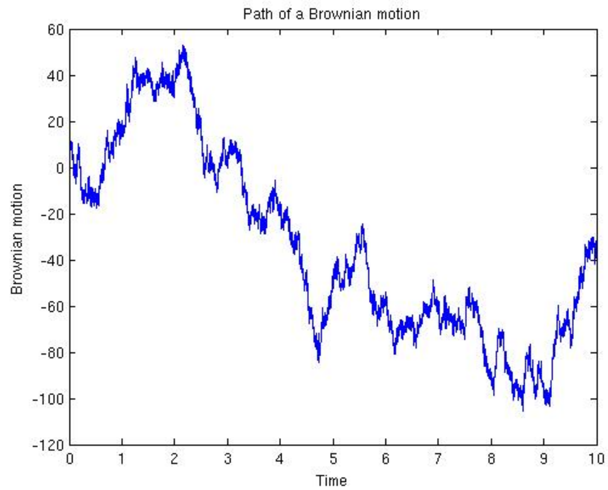


Figure 1.2: Path of a Brownian motion

need for stochastic models that allow for jumps (e.g. collapses in the market) has increased. Therefore, many new developments in stochastic analysis are now done using so-called Lévy processes, which include the possibility for jumps, see Figure 1.3. The Lévy measure $\nu(U)$, $U \subseteq \mathbb{B}_0(\mathbb{R})$ (i.e., $U$ is a Borel set whose closure does not contain 0), corresponding to a Lévy process is defined as the expected number of jumps of the process of size less than or equal $U$ that happen before or at time 1. It turns out (by the Itô-Lévy decomposition) that all Lévy processes $\eta(t)$ may be written as a sum of a deterministic term, a Brownian motion term and two jump terms; one "small-jump" term and one "big-jump" term. The big jump term is an integral with respect to a Poisson random measure $N(t,U)$, while the small jumps are integrals with respect to a compensated Poisson random measure $\tilde{N}(t,U)$ :

$$\eta(t) = \alpha t + \sigma B(t) + \int_{|z|<R} z\tilde{N}(t,dz) + \int_{|z|\geq R} zN(t,dz).$$

where $R \in \mathbb{R}$.

Also, one can show that under some finiteness conditions on the Lévy measure, the large jumps can be removed. This leads to jump process stochastic differential equations of the form:

$$\begin{aligned} dX(t) &= b(t,X(t))dt + \sigma(t,X(t))dB(t) + \int_{\mathbb{R}} \gamma(t,X(t),z)\tilde{N}(dt,dz), \ t \in [S,T] \\ X(S) &= x. \end{aligned}$$

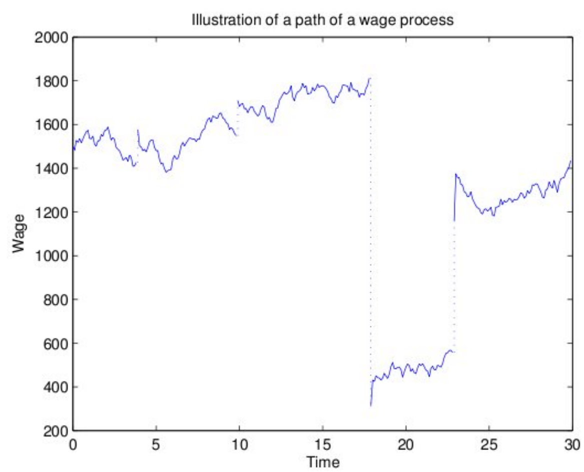


Figure 1.3: Path of a jump process (here: a wage process)

Brownian motion is such an important stochastic process because its properties make it attractive for mathematical analysis, but also because stochastic models based on this process turn out to fit real world data and phenomenons in a good way. The Brownian motion is actually the (random) movement of a particle in a fluid resulting from collisions with the molecules of the fluid. Other, examples of applications of stochastic models are biological population models, climate models, modeling the spreading of a virus or seeds from a plant, traffic models or modeling stock prices. The mathematical study of asset markets, i.e., mathematical finance, is one of the prime applications of stochastic analysis. The applications in this thesis are mainly from economics and mathematical finance. However, the same results and the same kind of analysis may be relevant to other fields of application as well.

Naturally, due the the randomness, the probabilities of the various scenarios are very important. In the initial coin toss example, the probability of all the outcomes is the same. That is, $P(\omega) = \frac{1}{8}$ for all $\omega \in \Omega$. In real world applications of stochastic analysis, this probability is sometimes determined by the model-maker, and sometimes one considers a set of potential probability measures and study what will happen in the worst possible case. However, in this thesis, we will mainly take the probability measure as given.

A crucial reason for wanting to model something, is to be able to make better decisions. This is the purpose of *stochastic optimization* (stochastic control). For instance, in mathematical finance a trader would like to maximize his expected utility over some period of time, given an initial wealth. In biology, one may aim to regulate hunting in order to keep an animal population at a sustainable level: too little hunting may lead to too many animals which can cause diseases, while too much hunting may lead to destruction of the population. Clearly, this illustrates the importance of stochastic optimization, which is a central topic in this thesis. The previous examples are all illustrations of *optimal control problems*, also called *stochastic control*, which is a kind of stochastic optimization problem. Another important kind of stochastic optimization is *optimal stopping*, where we aim to determine the optimal time to do a certain action. For example, one may study the optimal time to buy a house or to exercise an American option. However, in this thesis, we will mostly study optimal control.

In optimal control problems, we would like to find the best possible action for an agent over a time period, depending on the information available at each time. A feasible action is called an *admissible control*, and the best possible action is the one optimizing a performance function (e.g. minimizing a cost function or maximizing a reward function) over the set of all admissible controls. The best choice of control is called an *optimal control*. Information over time is represented in the models via *filtrations*. A filtration is a nested sequence of $\sigma$-algebras. The nesting of the $\sigma$-algebras correspond to the gradual revealing of information in the model as time evolves.

Hence, the standard optimal control problem (in the Brownian motion case) is as follows: Let $\mathscr{A}$ be the set of admissible controls, contained in the set of all adapted processes (wrt. the filtration generated by the Brownian motion). Let $X(t)$ be a stochastic process determining e.g. the market situation, defined as the solution of the stochastic differential equation (1.1). Let $f(u,x)$ be a running profit function, and let $g(x)$ be a function representing the terminal value. Define the *performance function* (for $X(S)=x$) by

$$J_u(x)=E^x\left[\int_S^T f(u(t),X(t))dt+g(X(T))\right].$$

Then, the basic optimal control problem is to find an admissible control $u^*$ which maximizes the performance function, i.e., $J_{u^*}(x)=\sup_{u\in\mathcal{A}} J_u(x)$.

There are two main approaches to solving such optimal control problems: *stochastic dynamic programming* (the Hamilton-Jacobi-Bellman equation) and stochastic maximum principles. The *Hamilton-Jacobi-Bellman* (HJB) approach is based on Bellman's optimality principle, which states that for an optimal control, the decisions from a certain time and until the end, must be optimal for the sub-problem starting at that that time and state. Hence, the problem is divided into sub-problems. When passing to the limit in time, this leads to a deterministic partial differential equation (PDE) where the unknown is the value function. Solving this PDE corresponds to solving the optimal control problem. However, this PDE can rarely be solved analytically, so numerical methods must be applied.

A weakness of the (classical) HJB approach is that it can only be applied to Markovian (memoryless) systems, see Øksendal et al. [73]. For instance, it is difficult to use the HJB method for problems with general partial information, see Haadem et al. [32]. However, the *stochastic maximum principle* can be used for non-Markovian systems as well. There are also examples of stochastic systems with time-inconsistencies such that the Bellman optimality principle does not hold, but where the maximum principle can be applied, see Buckdahn et al. [12]. The stochastic maximum principle is based on introducing a *Hamiltonian function* and *adjoint processes* corresponding to the control problem. Then, the idea is to prove that if some concavity conditions are satisfied, and a control maximizes the Hamiltonian function, then it is an optimal control. Hence, the stochastic control problem reduces to maximizing a real-valued function (and solving a so-called BSDE, which will be discussed shortly). However, as mentioned, some concavity conditions are required for this. These concavity conditions are not necessary for the HJB method, and this makes the HJB technique more suitable in some cases. Another advantage with the HJB method is that one has to solve a PDE instead of a BSDE. Usually, both of these equations must be solved numerically, and numerical approximation schemes for PDEs are very well known and developed. The basic idea behind the stochastic maximum principle is similar to that of the Pontryagin maximum principle, which is a deterministic maximum principle. The idea is to perturb an optimal control on a small time interval, do a Taylor expansion with respect to the time, and then let the time step tend towards zero. Then, one obtains a variational inequality, from which the maximum principle follows.

In Chapters 5, 6 and 7 maximum principles are derived. Furthermore, we consider problems with only one agent, but also stochastic control games, where two players interact so their choice of controls influence each other. Also in the case of two agents interacting, one can derive suitable stochastic maximum principles, see Chapter 7.

The HJB method leads to solving a partial differential equation. However, in order to determine the optimal control using the maximum principle approach, one has to solve a *backward stochastic differential equation* (BSDE) in the adjoint processes. In the case where we derive maximum principles for stochastic control games, we get *forward-backward stochastic differential equations* (FBSDEs) in the adjoint variables which need to be solved. Hence, we need an understanding of theory related to BSDEs and FBSDEs. Questions on existence and uniqueness of solutions of such equations are studied, as well as some solution methods. Only some of these equations may be solved analytically, so for practical applications, numerical methods are necessary.

There is a resemblance between the stochastic maximum principle and the *Lagrange duality method*. In both cases, one introduces a new function, the Hamiltonian and the Lagrangian respectively, which is a perturbed version of the objective function. Also, this transformation involves adjoint or dual variables. While Lagrange multipliers are introduced in order to handle deterministic constraints at a specific time, the adjoint stochastic processes are introduced in order to handle a stochastic differential equation constraint.

One may say that both the Lagrange multiplier method and the stochastic maximum principle are examples of *duality methods*. Depending on the type of optimization problem and the type of constraint for the problem, different variations of such duality solution methods can be applied. Table 1.1 gives a sketch of some common situations. For instance, if we are in a deterministic,

| | Discrete time | Continuous time |
|---|---|---|
| Deterministic | Lagrange multiplier | Pontryagin max. principle |
| Stochastic | Convex duality or Stochastic Lagrange | Stochastic max. principle |

Table 1.1: Duality methods for different kinds of constraints

discrete time setting (i.e., there are a finite number of constraints), the Lagrange multiplier method can be applied. In the stochastic, discrete time setting, *convex duality theory* (also called conjugate duality theory) may be used to solve the problem. Convex duality was first introduced by Rockafellar [85], and has over the past years been used in connection to stochastic analysis and mathematical finance by for instance Pennanen [78], Pennanen and Perkkiö [77] and Korf [47].

In this thesis, we will use several of the methods in Table 1.1 in order to solve stochastic control problems. In Chapter 2, we use Lagrange duality in order to price a claim under partial information, while in Chapter 4, a stochastic Lagrange multiplier method is introduced and combined with a stochastic maximum principle in order to study optimal consumption for an agent. Convex duality is used in Chapter 3, while different varieties of stochastic maximum principles are derived and applied in Chapters 5, 6 and 7.

The kind of *adjoint processes*, or dual variables, that are introduced in connection to these duality methods will vary according to the constraints at hand. If there is a finite number of deterministic constraints, there will be a finite number of deterministic dual variables. If the constraint is a differential equation, the adjoint variables will be given as a solution of a specific differential equation. If there is only one random constraint, the stochastic Lagrange dual variable will be a random variable. Finally, when the constraint is an SDE, as is the case when using the stochastic maximum principle, the adjoint variables will be the solution of a specific BSDE. Hence, the dual or adjoint variables are of the "same form" as the constraints in the original problem.

Also, note that the core of all duality methods is to transform one *primal* constrained optimization problem into another *dual optimization problem*. In the simplest kind of duality, linear programming duality (used in discrete state space, discrete time mathematical finance by for example Pliska [82]), these problems correspond to the standard linear programming primal and dual problems. In the case of the stochastic maximum principle, this corresponds to the original optimal control problem and the maximum principle with the adjoint variable as a constraint.

Another central topic of this thesis is *information*. As mentioned, information is included into the stochastic models via filtrations. However, in the real world, people may have different levels of information at the same time. There may be underlying processes that are hidden to some, but visible to others, there may be a delay in the information available or the memory may be noisy, i.e., influenced by randomness. Often, lack of complete information or difference in levels of information means that that standard optimal control results and maximum principles do not apply. Therefore, we develop new results adapted to these varying levels of information. In Chapters 2 and 3, we consider *partial information* in a financial market. In Chapter 2, we study delayed information, while in Chapter 3 we consider a completely general partial information. An example of partial information which is not delayed information is when there are hidden processes in the price dynamics, visible to some agents in the market, but not to all. In Chapters 6 and 7, we consider agent(s) who may have a *delay* in their information, as well as a *noisy memory* influencing their decisions. Also, in Chapter 7 there are two agents interacting, and they may have different delays and different length of (noisy) memory.

A lot of the stochastic differential equations arising in these problems cannot be solved analytically. Hence, there is a need for numerical methods. In Chapter 8, we derive an Euler type numerical method for approximating solutions of stochastic differential equations with noisy memory, as seen in Chapters 6 and 7.

## 1.1 Structure of the thesis and summary of the papers

This thesis consists of seven papers which, though all connected via the topic of stochastic control (optimization), can be divided into three categories: Duality theory, stochastic maximum principles and stochastic differential equations with noisy memory. However, the noisy memory SDE papers also use optimal control techniques, and as mentioned, the stochastic maximum principle can be viewed as a sort of duality method. In addition, the effect of different levels of information is also a core topic in several of the papers. Nevertheless, we divide the papers into these groups to make the basic ideas more prominent:

- **Part I:** Duality theory.
  - Pricing of claims in discrete time with partial information.
  - Convex duality methods for pricing contingent claims under partial information and short selling constraints.
  - Stochastic maximum principle with Lagrange multipliers and optimal consumption with Lévy wage.
- **Part II:** Stochastic maximum principles.
  - Singular recursive utility.
  - Optimal control of systems with noisy memory and BSDEs with Malliavin derivatives.
- **Part III:** Stochastic differential equations with noisy memory.
  - Forward backward stochastic differential equation games with delay and noisy memory.
  - A numerical method for the solution of stochastic differential equations with noisy memory.

The remaining part of this chapter consist of a brief summary of the papers.

### 1.1.1 Part I: Pricing of claims in discrete time with partial information

*This paper has been published in Applied Mathematics & Optimization (October 2013, Volume 68, Issue 2, pp 145-155).*

We consider the pricing problem of a seller of a contingent claim $B$ in a financial market with a finite scenario space $\Omega$ and a finite, discrete time setting. The seller is assumed to have information modeled by a filtration $(\mathcal{G}_t)_t$ which is generated by a delayed price process, so the seller has delayed price information. This delay of information is a realistic situation for many financial market traders. Actually, traders may pay to get updated prices. The seller's problem is to find the smallest price of $B$, such that there is no risk of her losing money. We solve this by deriving a dual problem via Lagrange duality, and use the linear programming duality theorem to show that there is no duality gap.

This paper considers the case of finite $\Omega$ and discrete time. Although this is not the most general situation, it is of practical use, since one often envisions only a few possible world scenarios, and has a finite set of times where one wants to trade. Also, for this and similar problems in mathematical finance, discretization is necessary to find efficient computational methods. There are many advantages to working with finite $\Omega$ and discrete time. The information structure of an agent can be illustrated in a scenario tree, making the information development easy to visualize. Conditions on adaptedness and predictability, are greatly simplified. Adaptedness of a process to a filtration means that the process takes one value in each vertex (node) of the scenario tree representing the filtration. Moreover, the general linear programming theory (see Vanderbei [97]) and Lagrange duality framework (see Bertsekas et al. [8]) apply. This allows for application of powerful theorems such as the linear programming duality theorem. Also, computational algorithms from linear programming, such as the simplex algorithm and interior point methods, can be used to solve the seller's problem in specific situations.

### 1.1.2 Part I: Convex duality methods for pricing contingent claims under partial information and short selling constraints

This paper analyzes an optimization problem from mathematical finance using conjugate duality. We consider the pricing problem of a seller of a contingent claim $B$ in a discrete time, arbitrary scenario space setting. The seller has a general level of partial information, and is subject to short selling constraints.

The seller's (stochastic) optimization problem is to find the minimum price of the claim such that she, by investing in a self-financing portfolio, has no risk of losing money at the terminal time $T$. The price processes are only assumed to be non-negative, stochastic processes, so the framework is model independent (in this sense).

The main contribution of the paper is a characterization of the seller's price of the claim $B$ as a $Q$-expectation of the claim, where $Q$ is a super-martingale measure with respect to the optional projection of the price process. The conjugate duality technique, which we use to prove this characterization, is different from what is common in the mathematical finance literature, and results in (fairly) brief proofs. Moreover, it does not rely on the reduction to a one-period model. This feature makes it possible to solve the optimization problem even though it contains partial information.

### 1.1.3 Part I: Stochastic maximum principle with Lagrange multipliers and optimal consumption with Lévy wage

*This paper is written in collaboration with PhD-student Espen Stokkereit, and has been accepted for publication in Afrika Matematika (DOI 10.10007/sIB370-015-0360-5).*

This paper derives a stochastic Lagrange multiplier method for solving constrained optimal control problems for jump diffusions. This can be used in combination with methods of optimal control, such as the stochastic maximum principle. Two different terminal constraints are considered, one that holds in expectation, and one that holds almost surely. Moreover, this method is used to analyze an interesting optimal consumption problem with wage jumps and stochastic inflation.

To analyze our version of the optimal consumption problem, we first impose a constraint on the expected terminal level of savings. This constraint transfers all the risk to the relevant financial institution (bank), and the consumers behave as if the market was complete. We assume that the agents have constant relative risk aversion (CRRA) utility functions and seek to maximize expected utility over a finite time horizon. Consequently, we are able to arrive at an explicit expression for an agent's optimal consumption process. Second, we impose an almost sure constraint on the terminal level of savings. This constraint is similar to the concept of admissibility widely used in the finance literature (see e.g. Karatzas and Shreve [45]), and makes the consumers bear all market risk. Thus, two extremes of risk sharing are considered.

### 1.1.4 Part II: Singular recursive utility

*This paper is written in collaboration with Bernt Øksendal, and has been submitted.*

Let $c(t) \geq 0$ be a consumption rate process. The classical way of measuring the total utility of $c$ from $t = 0$ to $t = T$ is by the expression

$$J(c) = E[\int_0^T U(t, c(t))dt]$$

where $U(t, \cdot)$ is a utility function for each $t$. This way of adding utility rates over time has been criticized from an economic and modeling point of view, see e.g. Mossin [62] and Hindy, Huang & Kreps [33]. Instead, Duffie and Epstein [25] proposed to use *recursive utility* $Y(t)$, defined as the solution of the backward stochastic differential equation (BSDE)

$$Y(t) = E[\int_t^T g(s, Y(s), c(s))ds | \mathcal{F}_t]; \ t \in [0, T]. \tag{1.2}$$

How should we model the recursive utility of a *singular* consumption process $\xi$? A natural proposal would be

$$Y(t) = E[\int_t^T g(s, Y(s), \xi(s))d\xi(s) | \mathcal{F}_t]. \tag{1.3}$$

We get, by the martingale representation theorem (see for instance Øksendal [64]), that $(Y, Z)$ solves the *singular BSDE*

$$\begin{aligned} dY(t) &= -g(t, Y(t), \xi(t))d\xi(t) + Z(t)dB(t) \\ Y(T) &= 0. \end{aligned} \tag{1.4}$$

To the best of our knowledge, such singular BSDEs have not been studied before. We show conditions for the existence and uniqueness of a solution for this kind of singular BSDE. Furthermore, we analyze the problem of maximizing the singular recursive utility. We derive sufficient and necessary maximum principles for this problem, and connect it to the Skorohod reflection problem. Finally, we apply our results to a specific cash flow. In this case, we find that the optimal consumption rate is given by the solution to the corresponding Skorohod reflection problem.

### 1.1.5 Part II: Optimal control of systems with noisy memory and BSDEs with Malliavin derivatives

*This paper is written in collaboration with S. E. A. Mohammed, B. Øksendal and E. E. Røse. It has been submitted.*

In this article, we develop two approaches for analyzing optimal control for a new class of stochastic systems with noisy memory. The main objective is to derive necessary and sufficient criteria for maximizing the performance functional on the underlying set of admissible controls. One should note the following unique features of the analysis:

- The state dynamics follows a controlled stochastic differential equation (SDE) driven by *noisy memory*: The evolution of the state $X$ at any time $t$ is dependent on its past history $\int_{t-\delta}^{t} X(s)\,dB(s)$ where $\delta$ is the memory span and $B$ is the driving white noise.
- The maximization problem is solved through a new backward stochastic differential equation (BSDE) that involves not only partial derivatives of the Hamiltonian but also their Malliavin derivatives.
- Two independent approaches are adopted for deriving necessary and sufficient maximum principles for the stochastic control problem: The first approach is via Malliavin Calculus and the second is a reduction of the dynamics to a two-dimensional controlled SDE with *discrete delay* and no noisy memory. In the second approach, the optimal control problem is then solved without resort to Malliavin calculus.

- A natural link between the above two approaches is established by using the solution of the two-dimensional BSDE in order to solve the noisy memory BSDE.

### 1.1.6 Part III: Forward backward stochastic differential equation games with delay and noisy memory

*This paper has been submitted.*

The aim of this paper is to study a stochastic game between two players. The game is based on a forward stochastic differential equation (SDE) in the process $X$, which determines the market situation. This SDE includes two kinds of memory of the past; regular memory and noisy memory. Regular memory (also called delay) means that the SDE can depend on previous values of the process $X$, while noisy memory means that the SDE may involve an Itô integral over previous values of the process.

Coupled to this market SDE are two backward stochastic differential equations (BSDEs). Each of these BSDEs corresponds to one of the players in the stochastic game; corresponding to player $i$ is a BSDE in the process $W_i$, $i = 1, 2$. Similar to the SDE, these BSDEs involve regular and noisy memory of the process $X$. However, the length of memory can be different for the two players. The players may also have different levels of information, which is illustrated by their different filtrations. For each of the players, the goal of the game is to find an optimal control $u_i$ which maximizes their personal performance function, $J_i$. This performance function depends on the player's profit rate, the market process $X$ and the process $W_i$ coming from the player's BSDE. Such FBSDE stochastic games have been studied by Øksendal and Sulem [68], however they do not include memory in their model. We study conditions for a pair of controls $(u_1, u_2)$ to be a Nash equilibrium for such a stochastic game. In order to do so, we derive sufficient and necessary maximum principles giving conditions for a control to be Nash optimal.

### 1.1.7 Part III: A numerical method for the solution of stochastic differential equations with noisy memory

*This paper has been submitted.*

Generalized noisy memory SDEs are stochastic differential equations where the system depends, in a noisy way, on the past values of the state process scaled by some time-dependent function. Applications of such equations can be modeling of animal populations where the population growth depends in some stochastic way on the previous population states, as well as the current number of animals. This effect may be influenced by time, for instance seasonal weather effects.

We show that such noisy memory SDEs are at least as difficult to solve as stochastic Volterrra equations. This means that noisy memory SDEs are often impossible to solve analytically. Therefore, we derive a numerical Euler scheme for such equations. Using, among other things, Grönwall's inequality and the Itô formula, we prove that the mean-square error of this scheme is of order $\sqrt{\Delta t}$. This is, perhaps somewhat surprisingly, the same order as the Euler scheme for regular SDEs, despite the added complexity from the noisy memory. To illustrate the numerical method, we apply it to a noisy memory SDE which can be solved analytically.

Chapter 2

# Pricing of claims in discrete time with partial information

By Kristina Rognlien Dahl.

Published in Applied Mathematics & Optimization October 2013, Volume 68, Issue 2, pp 145-155 (minor typos have been corrected).

**Abstract**

We consider the pricing problem of a seller with delayed price information. By using Lagrange duality, a dual problem is derived, and it is proved that there is no duality gap. This gives a characterization of the seller's price of a contingent claim. Finally, we analyze the dual problem, and compare the prices offered by two sellers with delayed and full information respectively.

## 2.1 Introduction

We consider the pricing problem of a seller of a contingent claim $B$ in a financial market with a finite scenario space $\Omega$ and a finite, discrete time setting. The seller is assumed to have information modeled by a filtration $(\mathcal{G}_t)_t$ which is generated by a delayed price process, so the seller has delayed price information. This delay of information is a realistic situation for many financial market traders. Actually, traders may pay to get updated prices.

The seller's problem is to find the smallest price of $B$, such that there is no risk of her losing money. We solve this by deriving a dual problem via Lagrange duality, and use the linear programming duality theorem to show that there is no duality gap. A related approach is that of King [46], where the fundamental theorem of mathematical finance is proved using linear programming duality. Vanderbei and Pilar [98] also use linear programming to price American warrants.

A central theorem of this paper is Theorem 2.1, which characterizes the seller's price of the contingent claim. This generalizes a pricing result by Delbaen and Schachermayer to a delayed information setting (see [18], Theorem 5.7). In Section 2.4, we compare the constrained and partially informed seller's price to that of an unconstrained seller. As one would expect, the seller with delayed information will offer $B$ at a higher price than a seller with full information.

Since the seller's pricing problem is parallel to the buyer's problem, of how much she is willing to pay for the claim, the results will carry through analogously for buyers. This implies that a buyer with delayed information is willing to pay less for the claim than a buyer with full information. Hence, it is less likely that a seller and buyer with delayed information will agree on a price than it is for fully informed agents.

This paper considers the case of finite $\Omega$ and discrete time. Although this is not the most general situation, it is of practical use, since one often envisions only a few possible world scenarios, and has a finite set of times where one wants to trade. Also, for this and similar problems in mathematical finance, discretization is necessary to find efficient computational methods.

There are many advantages to working with finite $\Omega$ and discrete time. The information structure of an agent can be illustrated in a scenario tree, making the information development easy to visualize. Conditions on adaptedness and predictability, are greatly simplified. Adaptedness of a process to a filtration means that the process takes one value in each vertex (node) of the scenario tree representing the filtration. Moreover, the general linear programming theory (see Vanderbei [97]) and Lagrange duality framework (see Bertsekas et al. [8]) apply. This allows application of powerful theorems such as the linear programming duality theorem. Also, computational algorithms from linear programming, such as the simplex algorithm and interior point methods, can be used to solve the seller's problem in specific situations. Note that the simplex algorithm is not theoretically efficient, but works very well in practice. Interior point methods, however, are both theoretically and practically efficient. Both algorithms will work well in practical situations where one considers a reasonable amount of possible world scenarios. Theoretically, they may nevertheless be inadequate for a very large number of possible scenarios.

Those familiar with linear programming may wonder why Lagrange duality is used to derive the dual problem instead of standard linear programming techniques. There are two important reasons for this. First of all, the Lagrange duality approach provides better economic understanding of the dual problem and allows for economic interpretations. Secondly, the Lagrange duality method can be explained briefly, and Lagrange methods are familiar to most mathematicians. Hence, using Lagrange duality makes this paper self-contained. The reader does not have to be familiar with linear programming or other kinds of optimization theory.

Other papers discussing the connection between mathematical finance and duality methods in optimization are Pennanen [78], King [46], King and Korf [47] and Pliska [82]. Pennanen [78] considers the connection between mathematical finance and the conjugate duality framework of Rockafellar [85]. King [46] proves the fundamental theorem of mathematical finance via linear programming duality, and King and Korf [47] derive a dual problem to the seller's pricing problem via the conjugate duality theory of Rockafellar. Pliska [82] also uses linear programming duality to prove that there exists a linear pricing measure if and only if there are no dominant trading strategies.

Examples of papers considering models with different levels of information in mathematical finance are Di Nunno et. al [24], Hu and Øksendal [37], Biagini and Øksendal [9], Lakner [53] and Platen and Rungaldier [81].

The remaining part of the paper is organized as follows. Section 2.2 explains the setting. The financial market is defined, the use of scenario trees to model filtrations is explained and the notation is introduced. Section 2.3 analyzes the seller's pricing problem with partial information via Lagrange duality. This leads to the central Theorem 2.1. In Section 2.4 we analyze the dual problem, and compare the result of Theorem 2.1 with the price offered by a seller will full information. This leads to Proposition 2.2. Section 2.5, concludes and poses questions for further research. For those interested, the appendix (Section 2.6) covers some background theory, namely Lagrange duality.

## 2.2 The model

The financial market is modeled as follows. We are given a probability space $(\Omega, \mathcal{F}, P)$ consisting of a finite scenario space, $\Omega = \{\omega_1, \omega_2, \dots, \omega_M\}$, a ($\sigma$-)algebra (here, there is no difference between $\sigma$-algebras and algebras since $\Omega$ is finite) $\mathcal{F}$ on $\Omega$ and a probability measure $P$ on the measurable space $(\Omega, \mathcal{F})$.

The financial market consists of $N+1$ assets: $N$ risky assets (stocks) and one non-risky asset (a bond). The assets each have a price process $S_n(t, \omega)$, $n = 0, 1, \dots, N$, for $\omega \in \Omega$ and $t \in \{0, 1, \dots, T\}$ where $T < \infty$, and $S_0$ denotes the price process of the bond. The price processes $S_n$, $n = 0, 1, \dots, N$, are stochastic processes. We denote by $S(t, \omega) := (S_0(t, \omega), S_1(t, \omega), \dots, S_N(t, \omega))$ the vector in $\mathbb{R}^{N+1}$ consisting of the price processes of all the assets. For notational convenience, we sometimes suppress the randomness, and write $S(t)$ instead of $S(t, \omega)$. Let $(\mathcal{F}_t)_{t=0}^T$ be the filtration generated by the price processes. We assume that $\mathcal{F}_0 = \{\emptyset, \Omega\}$ (so the prices at time 0, $S(0)$, are deterministic) and $\mathcal{F}_T$ is the algebra corresponding to the finest partition of $\Omega$, $\{\{\omega_1\}, \{\omega_2\}, \dots, \{\omega_M\}\}$.

We also assume that $S_0(t, \omega) = 1$ for all $t \in \{0, 1, \dots, T\}$, $\omega \in \Omega$. This corresponds to having divided through all the other prices by $S_0$, and hence turning the bank into the numeraire of the market. This altered market is a discounted market. To simplify notation, the price processes in the discounted market are denoted by $S$ as well. Note that the stochastic process $(S_n(t))_{t=0}^T$ is adapted to the filtration $(\mathcal{F}_t)_{t=0}^T$.

Consider a contingent claim $B$, i.e., a non-negative, $\mathcal{F}_T$-measurable random variable. $B$ is a financial asset which may be traded in the market. Therefore, consider a seller of the claim $B$. This seller has price information which is delayed by one time step. We let $(\mathcal{G}_t)_t$ be the filtration modeling the information structure of the seller. Hence, we let $\mathcal{G}_0 = \{\emptyset, \Omega\}$, $\mathcal{G}_t = \mathcal{F}_{t-1}$ for $t = 1, \dots, T-1$

and $\mathscr{G}_T = \mathscr{F}_T$. These assumptions imply that at time 0 the seller knows nothing, while at time $T$ the true world scenario is revealed. Note that since $\Omega$ is finite, there is a bijection between partitions and algebras (the algebra consists of every union of elements in the partition). The sets in the partition are called blocks.

One can construct a scenario-tree illustrating the situation, with the tree branching according to the information partitions. Each vertex of the tree corresponds to a block in one of the partitions. Each $\omega \in \Omega$ represents a specific development in time, ending up in the particular world scenario at the final time $T$. Denote the set of vertices at time $t$ by $\mathscr{N}_t$, and let the vertices themselves be indexed by $v = v_1, v_2, \ldots, v_V$.

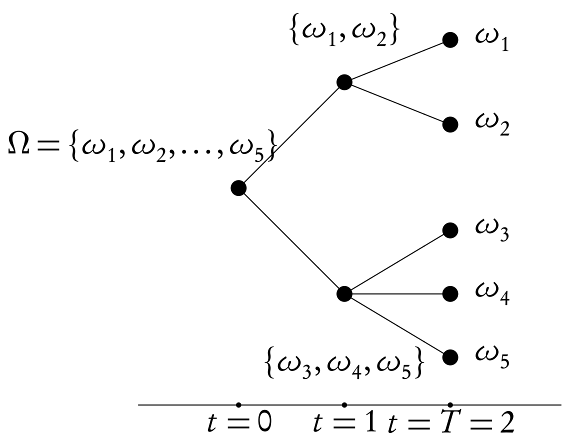


Figure 2.1: A scenario tree.

In the example illustrated in Figure 2.1, $V = 8$ and $M = 5$. The filtration $(\mathscr{G}_t)_{t=0,1,2}$ corresponds to the partitions $\mathscr{P}_1 = \{\Omega\}$, $\mathscr{P}_2 = \{\{\omega_1, \omega_2\}, \{\omega_3, \omega_4, \omega_5\}\}$, $\mathscr{P}_2 = \{\{\omega_1\}, \{\omega_2\}, \ldots, \{\omega_5\}\}$.

Some more notation is useful. The parent $a(v)$ of a vertex $v$ is the unique vertex $a(v)$ preceding $v$ in the scenario tree. Note that if $v \in \mathscr{N}_t$, then $a(v) \in \mathscr{N}_{t-1}$. Every vertex, except the first one, has a parent. Each vertex $v$, except the terminal vertices $\mathscr{N}_T$, have children vertices $\mathscr{C}(v)$. This is the set of vertices immediately succeeding the vertex $v$ in the scenario tree. For each non-terminal vertex $v$, the probability of ending up in vertex $v$ is called $p_v$, and $p_v = \sum_{u \in \mathscr{C}(v)} p_u$. Hence, from the original probability measure $P$, which gives probabilities to each of the terminal vertices, one can work backwards, computing probabilities for all the vertices in the scenario tree.

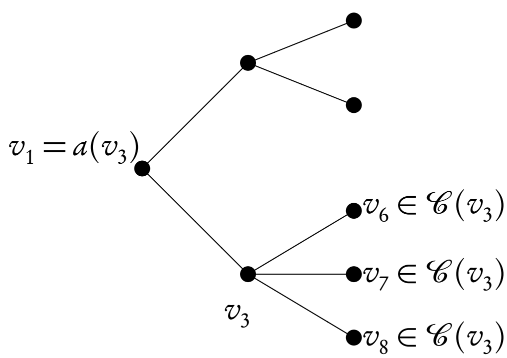


Figure 2.2: Parent and children vertices in a scenario tree.

The adaptedness of the price process $S$ to the filtration $(\mathscr{F}_t)_t$ means that, for each asset $n$, there is one value for the price $S_n$ in each vertex of the scenario tree. This value is denoted by $S_n^v$.

## 2.3 The pricing problem with partial information

Consider the model and the seller of Section 2.2, with $T \geq 4$. Following the same approach for a smaller $T$ is not a problem, but requires different notation and must therefore be considered separately. Hence, we consider a seller of a contingent claim $B$ who has price information that is delayed with one time step. Recall that the seller's filtration $(\mathscr{G}_t)_t$ is such that $\mathscr{G}_0 = \{\emptyset, \Omega\}$, $\mathscr{G}_t = \mathscr{F}_{t-1}$ for $t = 1, \ldots, T-1$, $\mathscr{G}_T = \mathscr{F}_T$.

The pricing problem of this seller is

$$
\begin{array}{llll}
\text{minimize} & x & & \\
\text{subject to} & & & \\
& S_0 \cdot H_0 \leq x, & & \\
& B_v \leq S_v \cdot H_{a_{\mathcal{G}}(v)} & & \text{for all } v \in \mathcal{N}_T^{\mathcal{G}}, \\
& S_{\mathcal{C}_{\mathcal{G}}(v)} \cdot H_v = S_{\mathcal{C}_{\mathcal{G}}(v)} \cdot H_{a_{\mathcal{G}}(v)} & & \text{for all } v \in \mathcal{N}_t^{\mathcal{G}} \\
& & & \text{and for all } \mathcal{C}_{\mathcal{G}}(v) \in \mathcal{N}_{t+1}^{\mathcal{G}}, \\
& & & t = 1, \ldots, T-2, \\
& S_{\mathcal{C}_{\mathcal{F}}(v)} \cdot H_v = S_{\mathcal{C}_{\mathcal{F}}(v)} \cdot H_{a_{\mathcal{G}}(v)} & & \text{for all } v \in \mathcal{N}_{T-1}^{\mathcal{G}} \\
& & & \text{and for all } \mathcal{C}_{\mathcal{F}}(v) \in \mathcal{N}_{T-1}^{\mathcal{F}}
\end{array}
\tag{2.1}
$$

where the minimization is done with respect to $x \in \mathbb{R}$ and $H_v \in \mathbb{R}$ for $v \in \mathcal{N}_t^{\mathcal{G}}$, for $t = 0, 1, \ldots, T-1$. Moreover, $\mathcal{N}_t^{\mathcal{G}}$ denotes the set of time $t$ vertices in the scenario tree representing the filtration $\mathcal{G}$, and similarly for the filtration $\mathcal{F}$. $B_v$ denotes the value of the claim $B$ in the vertex $v \in \mathcal{N}_T^{\mathcal{G}}$ (note that each $v \in \mathcal{N}_T^{\mathcal{G}}$ corresponds to an $\omega \in \Omega$). Also, $a_{\mathcal{G}}(v)$ denotes the parent of vertex $v$ w.r.t. the filtration $\mathcal{G}$ (see Section 2.2). Similarly, $\mathcal{C}_{\mathcal{G}}(v)$ and $\mathcal{C}_{\mathcal{F}}(v)$ denote the children vertices of vertex $v$ w.r.t. $\mathcal{G}$ and $\mathcal{F}$, respectively.

Hence, the seller's problem is: Minimize the price $x$ of the claim $B$ such that the seller is able to pay $B$ at time $T$ from investments in a self-financing, $\mathcal{G}$-adapted portfolio that costs less than or equal to $x$ at time 0. Note that the feasibility constraints in problem (2.1) imply that the seller acts in a self-financing matter w.r.t. the *actual* prices in the market. Let $\tilde{p}$ denote the seller's price of the claim $B$, so $\tilde{p}$ is the optimal value of problem (2.1). Problem (2.1) is a linear programming problem. Hence, there are efficient algorithms, such as the simplex algorithm or interior point methods, for solving problem (2.1), at least if the scenario tree is not too large.

We determine the dual problem of (2.1), using Lagrange duality techniques. In order to use the Lagrange duality method, rewrite the equality feasibility constraints as two inequality constraints. Let $y_0 \geq 0$, $z_v \geq 0$ for all $v \in \mathcal{N}_T^{\mathcal{G}}$, $y_v^1, y_v^2 \geq 0$ for all $v \in \mathcal{N}_t^{\mathcal{G}}$, for $t = 2, 3 \ldots, T-1$ and $w_v^1, w_v^2 \geq 0$ for all $v \in \mathcal{N}_{\mathcal{F}}^{T-1}$ be the Lagrange dual variables. Let $z$ denote the vector of all the $z_v$'s and similarly $y^i$, $w^i$ the vector of all the $y_v^i$'s and $w_v^i$'s for $i = 1,2$. Then, the Lagrange dual problem is

$$\begin{array}{ll}
\sup_{y_0,z,y^1,y^2,w^1,w^2\geq 0} & \inf_{x,H}\{x + y_0(S_0\cdot H_0 - x) + \sum_{v\in\mathcal{N}_T^{\mathcal{G}}} z_v(B_v - S_v\cdot H_{a_{\mathcal{G}}(v)}) \\
& +\sum_{t=1}^{T-2}\sum_{v\in\mathcal{N}_t^{\mathcal{G}}}\sum_{u\in\mathcal{C}_{\mathcal{G}}(v)}(y_u^1 - y_u^2)S_u\Delta H_v \\
& +\sum_{v\in\mathcal{N}_{T-2}^{\mathcal{F}}}\sum_{u\in\mathcal{C}_{\mathcal{F}}(v)}(w_u^1 - w_u^2)S_u\cdot\Delta H_v\} \\
= \sup_{y_0,z\geq 0,y,w} & \{\inf_x\{x(1-y_0)\} + \inf_{H_0}\{H_0\cdot(y_0S_0 - \sum_{u\in\mathcal{C}_{\mathcal{G}}(1)} y_uS_u)\} \\
& +\sum_{t=1}^{T-3}\sum_{v\in\mathcal{N}_t^{\mathcal{G}}}\inf_{H_v}\{H_v\cdot\sum_{u\in\mathcal{C}_{\mathcal{G}}(v)}(y_uS_u \\
& -\sum_{\mu\in\mathcal{C}_{\mathcal{G}}(u)} y_\mu S_\mu)\} + \sum_{v\in\mathcal{N}_{T-2}^{\mathcal{G}}}\inf_{H_v}\{H_v\cdot\sum_{u\in\mathcal{C}_{\mathcal{G}}(v)}(y_uS_u \\
& -\sum_{\mu\in\mathcal{C}_{\mathcal{F}}(u)} w_\mu S_\mu)\} + \sum_{v\in\mathcal{N}_{T-1}^{\mathcal{G}}}\inf_{H_v}\{H_v\cdot \\
& (\sum_{u\in\mathcal{C}_{\mathcal{F}}(v)} w_uS_u - \sum_{u\in\mathcal{C}_{\mathcal{G}}(v)} z_uS_u)\} + \sum_{v\in\mathcal{N}_T^{\mathcal{G}}} z_vB_v\}
\end{array}$$

where $y_v := y_v^1 - y_v^2$ and $w_v := w_v^1 - w_v^2$ are free variables, $\Delta H_v := H_v - H_{a_{\mathcal{G}}(v)}$ and we have exploited that the Lagrange function is separable.

Consider each of the minimization problems separately. In order to have a feasible dual solution, all of these minimization problems must have optimal value greater than $-\infty$:

- $\inf_x\{x(1-y_0)\} > -\infty$ if and only if $y_0 = 1$. In this case, the infimum is 0.
- $\inf_{H_0}\{H_0\cdot(y_0S_0 - \sum_{u\in\mathcal{C}_{\mathcal{G}}(1)} y_uS_u)\} > -\infty$ if and only if $y_0S_0 = \sum_{u\in\mathcal{C}_{\mathcal{G}}(1)} y_uS_u$. In this case, the infimum is 0.
- Note that

  $$\inf_{H_v}\{H_v\cdot\sum_{u\in\mathcal{C}_{\mathcal{G}}(v)}(y_uS_u - \sum_{\mu\in\mathcal{C}_{\mathcal{G}}(u)} y_\mu S_\mu)\} > -\infty$$

  if and only if $\sum_{u\in\mathcal{C}_{\mathcal{G}}(v)}(y_uS_u - \sum_{\mu\in\mathcal{C}_{\mathcal{G}}(u)} y_\mu S_\mu) = 0$. Therefore, in order to get a dual solution, this must hold for all $v\in\mathcal{N}_t^{\mathcal{G}}$ for $t = 1,2,\ldots,T-3$. In this case, the infima are 0.
- Furthermore, $\inf_{H_v}\{H_v\cdot\sum_{u\in\mathcal{C}_{\mathcal{G}}(v)}(y_uS_u - \sum_{\mu\in\mathcal{C}_{\mathcal{F}}(u)} w_\mu S_\mu)\} > -\infty$ if and only if $\sum_{u\in\mathcal{C}_{\mathcal{G}}(v)}(y_uS_u - \sum_{\mu\in\mathcal{C}_{\mathcal{F}}(u)} w_\mu S_\mu) = 0$. Again, in this case, the infimum is 0.

- Finally, $\inf_{H_v}\{H_v\cdot(\sum_{u\in\mathscr{C}_{\mathscr{F}}(v)} w_u S_u - \sum_{u\in\mathscr{C}_{\mathscr{G}}(v)} z_u S_u)\} > -\infty$ if and only if $\sum_{u\in\mathscr{C}_{\mathscr{F}}(v)} w_u S_u = \sum_{u\in\mathscr{C}_{\mathscr{G}}(v)} z_u S_u$. Hence, this must hold for all $v\in\mathscr{N}_{T-1}^{\mathscr{G}}$. In this case the infimum is 0.

Hence, the dual problem is

$$
\begin{array}{ll}
\sup_{y_0,z\geq 0,y,w} & \sum_{v\in\mathscr{N}_T^{\mathscr{G}}} z_v B_v \\
\text{subject to} &
\end{array}
$$

$$
\begin{array}{rcl}
y_0 & = & 1, \\
y_0 S_0 & = & \sum_{u\in\mathscr{C}_{\mathscr{G}}(1)} y_u S_u, \\
\sum_{u\in\mathscr{C}_{\mathscr{G}}(v)}(y_u S_u - \sum_{\mu\in\mathscr{C}_{\mathscr{G}}(u)} y_\mu S_\mu) & = & 0 \quad \text{for all } v\in\mathscr{N}_t^{\mathscr{G}}, t=1,2,\dots,T-3, \\
\sum_{u\in\mathscr{C}_{\mathscr{G}}(v)}(y_u S_u - \sum_{\mu\in\mathscr{C}_{\mathscr{F}}(u)} w_\mu S_\mu) & = & 0 \quad \text{for all } v\in\mathscr{N}_{T-2}^{\mathscr{G}}, \\
\sum_{u\in\mathscr{C}_{\mathscr{F}}(v)} w_u S_u & = & \sum_{u\in\mathscr{C}_{\mathscr{G}}(v)} z_u S_u \quad \text{for all } v\in\mathscr{N}_{T-1}^{\mathscr{G}}.
\end{array}
\tag{2.2}
$$

Note that the dual feasibility conditions are vector equations. From the linear programming duality theorem, see Vanderbei [97], there is no duality gap. Hence, the optimal value of problem (2.1) equals the optimal value of problem (2.2).

By analyzing the dual feasibility conditions, we can remove the variable $w$ and rewrite problem (2.2) so that it is expressed using the filtration $(\mathscr{F}_t)_t$:

$$
\begin{array}{ll}
\sup_{y_0,z\geq 0,y} & \sum_{v\in\mathscr{N}_T^{\mathscr{G}}} z_v B_v \\
\text{subject to} &
\end{array}
$$

$$
\begin{array}{rcl}
y_0 & = & 1, \\
y_0 S_0 & = & \sum_{u\in\mathscr{C}_{\mathscr{F}}(0)} y_u S_u, \\
\sum_{u\in\mathscr{C}_{\mathscr{F}}(v)}(y_u S_u - \sum_{\mu\in\mathscr{C}_{\mathscr{F}(u)}} y_\mu S_\mu) & = & 0 \quad \text{for all } v\in\mathscr{N}_t^{\mathscr{F}}, \\
 & & \quad t=0,1,\dots,T-4, \\
\sum_{u\in\mathscr{C}_{\mathscr{F}}(v)}(y_u S_u - \sum_{\mu\in\mathscr{C}_{\mathscr{F}}(u)} \sum_{\gamma\in\mathscr{C}_{\mathscr{F}}(\mu)} z_\gamma S_\gamma) & = & 0 \quad \text{for all } v\in\mathscr{N}_{T-3}^{\mathscr{F}}.
\end{array}
\tag{2.3}
$$

It is difficult to interpret problem (2.3) in its present form. It turns out that we can rewrite this problem slightly so that it is easier to understand. Note that

$$\sum_{v \in \mathcal{N}_T^{\mathcal{F}}} z_v S_v = \sum_{u \in \mathcal{N}_1^{\mathcal{F}}} y_u S_u = y_0 S_0 \tag{2.4}$$

where the first equality follows from using the dual feasibility conditions inductively, and summing over all vertices at each time. Equation (2.4) is a vector equation. Since the market is normalized, the first component of the price process vector is 1 at each time $t$. Hence, equation (2.4) implies that $\sum_{v \in \mathcal{N}_t^{\mathcal{F}}} z_v = y_0 = 1$ where the final equality uses the first dual feasibility condition. Recall that $z$ is non-negative from problem (2.3). Hence, $z$ can be identified with a probability measure on the terminal vertices of the scenario tree. Denote this probability measure by $Q$. Then, problem (2.3) can be rewritten

$$\begin{array}{lrl}
\sup_{Q,y} & E_Q[B] & \\
\text{subject to} & & \\
(i) & S_0 & = \sum_{u \in \mathcal{C}_{\mathcal{F}}(0)} y_u S_u, \\
(ii) & \sum_{u \in \mathcal{C}_{\mathcal{F}}(v)} (y_u S_u - \sum_{\mu \in \mathcal{C}_{\mathcal{F}}(u)} y_\mu S_\mu) & = 0 \ \text{ for } v \in \mathcal{N}_t^{\mathcal{F}}, t = 0, \ldots, T-4, \\
(iii) & \sum_{u \in \mathcal{C}_{\mathcal{F}}(v)} y_u S_u & = \sum_{u \in \mathcal{C}_{\mathcal{F}}(v)} \sum_{\mu \in \mathcal{C}_{\mathcal{F}}(u)} \sum_{\gamma \in \mathcal{C}_{\mathcal{F}}(\mu)} q_\gamma S_\gamma \\
 & & \quad \text{for } v \in \mathcal{N}_{T-3}^{\mathcal{F}}
\end{array} \tag{2.5}$$

where $Q$ is a probability measure and $q_\gamma$ denotes the $Q$-probability of ending up in vertex $\gamma$ at time $T$.

The dual problem is to maximize the expectation of the contingent claim $B$ over a set of probability measures, and some constraints regarding the price process and a free variable $y$. However, there is no martingale measure interpretation of the dual problem. Let $\tilde{d}$ denote the optimal value of the transformed dual problem (2.5).

The previous derivation gives us the following theorem.

**Theorem 2.1.** *Consider a seller of a contingent claim $B$ who has partial information in the sense that her price information is delayed by one time step. Then, $\tilde{p} = \tilde{d}$, i.e. the seller's price of $B$ is equal to the optimal value of problem (2.5).*

Note that for a specific problem, one can solve problem (2.5) using the simplex algorithm or interior point methods (for a reasonably sized scenario tree).

Also, the same kind of argument as above can be done from the buyer's point of view, yielding dual problem similar to problem (2.5), but with infimum instead of the supremum.

## 2.4 Some comments on the dual problem

In this section, we connect Theorem 2.1 to some previously known results.

### 2.4.1 Connection to full information

From Delbaen and Schachermayer [18] (or a derivation similar to that of Section 2.3), we know that the seller's price of $B$ with full information is

$$\alpha := \sup_{Q \in \mathcal{M}(S,\mathcal{F})} E_Q[B] \tag{2.6}$$

where $\mathcal{M}(S,\mathcal{F})$ denotes the set of equivalent martingale measures w.r.t. the filtration $(\mathcal{F}_t)_t$. In the following, assume there exists a $Q \in \mathcal{M}(S,\mathcal{F})$. From [18], this means that there is no arbitrage in the market.

**Theorem 2.2.** *The difference between the price of B offered by a seller with delayed information and a seller with full information is*

$$\tilde{d} - \alpha \geq 0. \tag{2.7}$$

*Proof.* From the definition of $\tilde{d}$ and $\alpha$, it suffices to prove that each $Q \in \mathcal{M}(S,\mathcal{F})$ corresponds to a solution $y, \tilde{Q}$ of problem (2.5). Hence, let $Q \in \mathcal{M}(S,\mathcal{F})$. Define $\tilde{Q} := Q$, and for each $v \in \mathcal{N}^{\mathcal{F}}_{T-1}$, define $y_v := \sum_{u \in C_{\mathcal{F}}(v)} \tilde{q}_u$ (where $\tilde{q}_u$ is the probability corresponding to node $u \in \mathcal{N}_T$). Similarly, for each $v \in \mathcal{N}^{\mathcal{F}}_{T-2}$, define $y_v := \sum_{u \in C_{\mathcal{F}}(v)} y_u$. Iteratively, we define $y_v := \sum_{u \in C_{\mathcal{F}}(v)} y_u$ for each $v \in \mathcal{N}^{\mathcal{F}}_t$, $t = 0, \ldots, T-3$. We would like to show that $\tilde{Q}, y$ are feasible for problem (2.5).

- $(i)$ : Since $Q \in \mathcal{M}(S,\mathcal{F})$, $S_0 = E_Q[S_1|\mathcal{F}_0]$, which from the definition of conditional expectation implies $(i)$.

$(ii)$ : $Q \in \mathcal{M}(S, \mathcal{F})$ implies that $E_Q[S_{t+1}|\mathcal{F}_t] = S_t$. Hence, from the definition of conditional expectation, $y_u S_u = \sum_{\mu \in \mathcal{C}_{\mathcal{F}}(u)} y_\mu S_\mu$ for all $u \in \mathcal{N}_t^{\mathcal{F}}$, so $(ii)$ holds.

$(iii)$ : Again, since $Q \in \mathcal{M}(S, \mathcal{F})$, $E_Q[S_T|\mathcal{F}_{T-2}] = S_{T-2}$. Hence, $y_u S_u = \sum_{\mu \in \mathcal{C}_{\mathcal{F}}(u)} \sum_{\gamma \in \mathcal{C}_{\mathcal{F}}(\mu)} q_\gamma S_\gamma$, so $(iii)$ holds.

Hence, the theorem follows. □

The difference in Theorem 2.2 can be computed for specific examples.

Theorem 2.2 implies that, as one would expect, the seller with only partial information will demand a higher price for $B$ than a fully informed seller. As in Section 2.3, the same kind of argument goes through for a buyer of the claim. Hence, the chance of a seller and buyer agreeing on a price for the claim is smaller in a market with delayed information, than in the fully informed case.

### 2.4.2 A closer bound

We can find an interpretable problem which has optimal value closer to that of problem (2.5) than the full information problem (2.6). Consider the following optimization problem

$$\begin{array}{ll} \sup_Q & E_Q[B] \\ \text{subject to} & \end{array}$$

$$\begin{array}{rcll} S_0 & = & E_Q[S_1], & \\ E_Q[S_{t+1}|\mathcal{F}_t] & = & E_Q[S_{t+2}|\mathcal{F}_t] & \text{for } t = 0, 1, \ldots, T-4, \\ E_Q[S_{T-2}|\mathcal{F}_{T-3}] & = & E_Q[S_T|\mathcal{F}_{T-3}]. & \end{array} \tag{2.8}$$

Let $\beta$ denote the the optimal value of problem (2.8).

**Theorem 2.3.** *The optimal value of problem (2.8) lies between the price of B offered by the seller with full information and the price offered by the seller with delayed information, i.e.,*

$$\alpha \leq \beta \leq \tilde{d}$$

*Proof.* Clearly, $\alpha \leq \beta$, from the definition of $\mathcal{M}(S, \mathcal{F})$.

It remains to prove that $\beta \leq \tilde{d}$. Hence, consider $Q$ feasible in problem (2.8). It suffices to prove that $Q$ corresponds to a feasible solution $\tilde{Q}, y$ for problem (2.5). Define $\tilde{Q}$ and $y$ as in the proof of Proposition 2.2. We check the feasibility constraints of problem (2.5).

$(i)$ : Since $Q$ is feasible in (2.8), $S_0 = E_Q[S_1]$. Hence, from the definition of conditional expectation, $S_0 = \sum_{u \in \mathcal{C}_{\mathcal{F}}(0)} y_u S_u$.

$(ii)$ : Again, since $Q$ is feasible in (2.8), $E_Q[S_{t+1}|\mathcal{F}_t] = E_Q[S_{t+2}|\mathcal{F}_t]$ for $t = 0, 1, \ldots, T-4$. From the definition of conditional expectation, this implies that $\sum_{u \in \mathcal{C}_{\mathcal{F}}(v)}(y_u S_u - \sum_{\mu \in \mathcal{C}_{\mathcal{F}}(u)} y_\mu S_\mu) = 0$ for all $v \in \mathcal{N}_t^{\mathcal{F}}$, $t = 0, \ldots, T-4$. Hence, $(ii)$ holds.

$(iii)$ : $(iii)$ follows similarly from the feasibility of $Q$ in (2.8) and the definition of conditional expectation.

Hence, $\beta \leq \tilde{d}$. □

## 2.5 Final remarks

A main idea of this paper has been to illustrate the close connection between pricing problems in mathematical finance and duality methods in optimization.

Some questions open for further research are:

- Can these results be generalized to a model with continuous time, possibly using a discrete time approximation?
- Is it possible to characterize the partially informed seller's dual problem more explicitly?

**Acknowledgements:** We would like to thank an anonymous referee for careful reading and many helpful suggestions. We are also very grateful to Professor Bernt Øksendal (University of Oslo) for several useful comments.

## 2.6 Appendix: Lagrange duality

This section reviews some basic ideas and results concerning Lagrange duality which will be useful in the following. For more on Lagrange duality and optimization theory, see Bertsekas et al. [8].

Let $X$ be a general inner product space with inner product $\langle \cdot, \cdot \rangle$. Consider a function $f : X \to \mathbb{R}$ and the very general optimization problem

$$\text{minimize } f(x) \text{ subject to } g(x) \leq 0, \ x \in S \tag{2.9}$$

where $g$ is a function such that $g : X \to \mathbb{R}^n$ and $S$ is a non-empty subset of $X$. Here, $g(x) \leq 0$ means component-wise inequality. This will be called the primal problem.

Define the Lagrange function, $L(x, \lambda)$, for $\lambda \in \mathbb{R}^n$, $\lambda \geq 0$ (component-wise), to be

$$L(x, \lambda) \ = \ f(x) + \lambda \cdot g(x).$$

where $(\cdot)$ denotes Euclidean inner product.

Then, for all $x \in X$ such that $g(x) \leq 0$ (component-wise) and all $\lambda \in \mathbb{R}^n$, $\lambda \geq 0$, we have $L(x, \lambda) \leq f(x)$. This motivates the definition $L(\lambda) := \inf_{x \in S} L(x, \lambda)$ for all $\lambda \geq 0$ (note that $L(\lambda) = -\infty$ is possible), and the Lagrange dual problem

$$\sup_{\lambda \geq 0} L(\lambda).$$

This gives the following result called weak Lagrange duality.

**Proposition 2.4.**

$$\sup\{L(\lambda) : \lambda \geq 0\} \ \leq \ \inf\{f(x) : g(x) \leq 0, \ x \in S\}.$$

Hence, the Lagrange dual problem gives the greatest lower bound on the optimal value of problem (2.9), based on $L$. Often the Lagrange dual problem is separable, and therefore fairly easy to solve. For some problems, one can proceed to show duality theorems, proving that $\sup_{\lambda \geq 0} L(\lambda) = \inf\{f(x) : g(x) \leq$

$0,\ x \in S\}$. In this case, one says that there is no duality gap. This typically occurs in convex optimization problems under certain assumptions. For instance, the linear programming duality theorem (see Vanderbei [97]) may be derived using Lagrange duality.

Chapter 4

# Stochastic maximum principle with Lagrange multipliers and optimal consumption with Lévy wage

By Kristina Rognlien Dahl[1] and Espen Stokkereit[2]
To be published in Afrika Matematika (minor typos have been corrected).
DOI: 10.1007/s13370-015-0360-5.

**Abstract**

We show how a stochastic version of the Lagrange multiplier method can be combined with the stochastic maximum principle for jump diffusions to solve certain constrained stochastic optimal control problems. Two different terminal constraints are considered; one constraint holds in expectation and the other almost surely.

As an application of this method, we study the effects of inflation- and wage risk on optimal consumption. To do this, we consider the

[1]The research leading to these results has received funding from the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement no. 228087.

[2]E. Stokkereit, Department of Economics, University of Oslo, Pb. 1095 Blindern, 0316 Oslo, Norway. espen.stokkereit@econ.uio.no

optimal consumption problem for a budget constrained agent with a Lévy income process and stochastic inflation. The agent must choose a consumption path such that his wealth process satisfies the terminal constraint. We find expressions for the optimal consumption of the agent in the case of CRRA utility, and give an economic interpretation of the adjoint processes.

## 4.1 Introduction

This paper derives a stochastic Lagrange multiplier method to solve constrained optimal control problems for jump diffusions. This can be used in combination with methods of optimal control, such as the stochastic maximum principle. Two different terminal constraints are considered, one that holds in expectation (soft constraint), and one that holds almost surely (hard constraint). Moreover, this method is used to analyze an interesting optimal consumption problem with wage jumps and stochastic inflation.

The problem of determining optimal lifetime consumption under uncertainty dates back to the seminal papers of Merton [57], [58]. These papers treat the problem of portfolio choice in continuous time in a complete market. In particular, there is no risk in the wage-level and the price of the consumption good is constant. This is explored further by Karatzas and Shreve [44], by using a so-called martingale method to handle market incompleteness. Karatzas [43] also considers a dynamic, stochastic economy with several heterogeneous agents.

To analyze our version of the optimal consumption problem, we first impose a constraint on the expected terminal level of savings. This constraint transfers all the risk to the relevant financial institution (bank), and the consumers behave as if the market was complete. We assume that the agents have constant relative risk aversion (CRRA) utility functions and seek to maximize expected utility over a finite time horizon. Consequently, we are able to arrive at an explicit expression for an agent's optimal consumption process. Second, we impose an almost sure constraint on the terminal level of savings. This constraint is similar to the concept of admissibility widely used in the finance literature (see e.g. Karatzas and Shreve [45]), and makes the consumers bear all market risk. Thus, two extremes of risk sharing are considered.

A motivation for studying a consumption optimization problem with stochastic income and a CRRA utility function is found in Zeldes [102]. The paper argues that such uncertainties dramatically affects the consumption function, and links this to three classical empirical consumption puzzles. The model in Zeldes [102] uses discrete time and numerical approximations for the solution. In contrast, we consider continuous time and focus on analytical solutions.

A paper which, similarly to our paper, has a more analytical approach to such a consumption optimization problem is El Karoui and Jeanblanc-Picqué [26]. The authors solve the consumption-portfolio problem for an agent with a stochastic, insurable income under a liquidity constraint. Opposed to our situation, El Karoui and Jeanblanc-Picqué [26] assume that the wage is not a source of new uncertainty, and they have a liquidity constraint which prohibits all borrowing against future income.

Koo [49] also considers a consumption-portfolio problem for a liquidity constrained agent, but who has an uninsurable income risk. In this case, the wage process is modeled using only a Brownian motion. This is contrary to our situation, where the wage process is a jump process. Note that none of the articles mentioned so far specifically consider the inflation risk, as we do in this paper.

It is interesting to include inflation risk because having a stochastic future consumption price adds more realism to the model, yet it does not make the problem significantly more complicated. A paper which does consider the inflation risk, as well as the income risk, is Battocchio and Menoncin [7]. However, their wage process does not include jumps (only a Brownian motion).

The structure of this paper is as follows. In Section 4.2 we introduce a general stochastic optimal control problem with constraints, and prove a stochastic Lagrange multiplier method for solving this type of problem. This is the theoretical and methodical foundation for the rest of the paper. Then, in Section 4.3, we treat an interesting application from economy in detail: We introduce a specific stochastic control problem involving inflation- and wage risk. In Section 4.4 the problem is solved using the stochastic multiplier method of Section 4.2. We consider the case with a soft end constraint (i.e., the constraint holds in expectation), where an explicit solution is derived using the Lagrange multiplier method and the stochastic maximum principle.

## 4.2 A stochastic Lagrange multiplier method

Let $(\Omega, \mathscr{F}, P)$ be a probability space, where $\Omega$ is the scenario space, $\mathscr{F}$ a $\sigma$-algebra and $P$ the probability measure. We consider continuous time, $t \in [0,T]$. Let $\{B(t)\}_{t\in[0,T]}$ be a Brownian motion, and $\int_{\mathbb{R}} z\tilde{N}(dt,dz)$ a pure jump process independent of $B(t)$. Let $\{\mathscr{F}_t\}_{t\in[0,T]}$ be the filtration generated by the Brownian motion and the pure jump process. In the following, by an adapted processes, we mean adapted with respect to this filtration.

Also, let $f : \mathbb{R}_+ \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$ be given, continuous functions. Then, we consider the stochastic optimal control problem which comes in two versions $(i)$ and $(ii)$:

$$\begin{aligned}
&\sup_{u\in\mathscr{A}} \quad E^x[\textstyle\int_0^T f(t,X(t),u(t))dt + g(X(T))] \\
&\text{subject to} \\
&\qquad dX(t) = b(t,X(t),u(t))dt + \sigma(t,X(t),u(t))dB(t) \\
&\qquad\qquad + \textstyle\int_{\mathbb{R}} \gamma(t,X(t^-),u(t^-),z)\tilde{N}(dt,dz) \\
&\qquad (i)\ E^x[M(X(T))] = 0 \text{ or } (ii)\ M(X(T)) = 0 \text{ a.s.,}
\end{aligned} \tag{4.1}$$

where $M : \mathbb{R} \to \mathbb{R}$ is some given continuous function, $\mathscr{U} \subset \mathbb{R}$ is a given set, $b : \mathbb{R}_+ \times \mathbb{R} \times \mathscr{U} \to \mathbb{R}$, $\sigma : \mathbb{R}_+ \times \mathbb{R} \times \mathscr{U} \to \mathbb{R}$ and $\gamma : \mathbb{R}_+ \times \mathbb{R} \times \mathscr{U} \times \mathbb{R} \to \mathbb{R}$ . Here, $E^x[\cdot]$ denotes the expectation given that the state process $(X(t))_{t\in[0,T]}$ starts in $x$, i.e. $X(0) = x$.

In problem (4.1), the stochastic process $u(t) = u(t,\omega) : \mathbb{R}_+ \times \Omega \to \mathscr{U}$ is our control process. We say that this control process $u(t)$ is admissible, and write $u \in \mathscr{A}$ if the SDE in problem (4.1) has a unique, strong solution for all $x \in \mathbb{R}$, and

$$E^x[\int_0^T f(t,X(t),u(t))dt + g(X(T))] < \infty.$$

There is a sufficient stochastic maximum principle for jump diffusions by Framstad et al. [29] (see also Tang and Li [95]), which can be used to find the optimal control of problem (4.1) without the additional constraints $(i)$ or $(ii)$. This maximum principle converts the simplified problem into the problem of maximizing a Hamiltonian function, and solving the so-called adjoint backward stochastic differential equation (BSDE). For reference purposes, this stochastic maximum principle, Theorem 4.9, is written out in Appendix 4.6.

However, if we add a constraint such as $(i)$ or $(ii)$ to the problem, such as in (4.1), the stochastic maximum principle cannot be used directly. We consider these two different kinds of constraints, and show how the constrained stochastic optimal control problems can be solved using a combination of a generalized Lagrange duality method and the stochastic maximum principle.

**Remark 4.1.** *For notational simplicity, problem (4.1) is assumed to be in one dimension. However, the results of this section also apply to multi-dimensional stochastic optimal control problems.*

### 4.2.1 Constraint of type $(i)$

Consider problem (4.1) with a type $(i)$ constraint, i.e.:

$$\phi(x) := \sup_{u\in\mathcal{A}} E^x\left[\int_0^T f(t,X(t),u(t))dt + g(X(T))\right]$$

subject to

$$\begin{aligned} dX(t) &= b(t,X(t),u(t))dt + \sigma(t,X(t),u(t))dB(t) \\ &\quad + \int_{\mathbb{R}} \gamma(t,X(t^-),u(t^-),z)\tilde{N}(dt,dz) \\ E^x[M(X(T))] &= 0. \end{aligned} \tag{4.2}$$

This problem can be solved using the standard Lagrange multiplier method, and then applying some method of stochastic control, for instance the stochastic maximum principle. Hence, let $\lambda \in \mathbb{R}$ be a Lagrange multiplier. Then, we introduce the unconstrained stochastic control problem

$$\begin{aligned} &\phi_\lambda(x) := \sup_{u\in\mathcal{A}} E^x\left[\int_0^T f(t,X(t),u(t))dt + g(X(T)) + \lambda M(X(T))\right] \\ &\text{subject to} \\ &\qquad dX(t) = b(t,X(t),u(t))dt + \sigma(t,X(t),u(t))dB(t) \\ &\qquad\qquad + \int_{\mathbb{R}} \gamma(t,X(t^-),u(t^-),z)\tilde{N}(dt,dz). \end{aligned} \tag{4.3}$$

This solution strategy is explored in Section 11.3 in Øksendal [64] for the no-jump case. However, the proof of this theorem generalizes in a straight-forward manner to the Lévy case (i.e., where we may have jumps in the dynamics of the state process $X(t)$). Therefore, we have the following theorem.

**Theorem 4.2.** *(Theorem 11.3.1 in Øksendal [64], adapted to the jump case) Suppose that we for all $\lambda \in \mathbb{R}$ can find $\phi_\lambda(y)$ and $u^*_\lambda$ solving the unconstrained stochastic control problem (4.3). Moreover, suppose there exists $\lambda_0 \in \mathbb{R}$ such that*

$$E^x[M(X_{u^*_{\lambda_0}}(T))] = 0.$$

*Then, $\phi(x) := \phi_{\lambda_0}(x)$ and $u^* := u^*_{\lambda_0}$ solves the constrained stochastic control problem (4.2).*

*Proof.* See Øksendal [64], Theorem 11.3.1 and also the proof of the following Theorem 4.3. □

From Theorem 4.2, it is sufficient to solve problem (4.3) in order to solve problem (4.2). Problem (4.3) can be solved using the stochastic maximum principle (see Theorem 4.9 in the Appendix).

### 4.2.2 Constraint of type $(ii)$

Now, consider problem (4.1) with a type $(ii)$ constraint. That is, consider the stochastic optimal control problem

$$\phi(x) := \sup_{u \in \mathcal{A}} E^x[\int_0^T f(t, X(t), u(t))dt + g(X(T))]$$

subject to

$$\begin{aligned} dX(t) &= b(t, X(t), u(t))dt + \sigma(t, X(t), u(t))dB(t) \\ &\quad + \int_{\mathbb{R}} \gamma(t, X(t^-), u(t^-), z)\tilde{N}(dt, dz) \\ M(X(T)) &= 0 \text{ a.s.} \end{aligned} \tag{4.4}$$

where, as before, $M : \mathbb{R} \to \mathbb{R}$ is a given, continuous function. For notational simplicity, define

$$J^u(x) := E^x[\int_0^T f(t, X(t), u(t))dt + g(X(T))].$$

We would like to use the Lagrange multiplier concept to solve problem (4.4) by solving an unconstrained stochastic control problem. However, since we have an almost sure constraint, it is not sufficient to introduce a single scalar Lagrange multiplier $\lambda \in \mathbb{R}$. The Lagrange multiplier must be stochastic in order to handle the stochastic constraint $M(X(T)) = 0$ a.s. Hence, we introduce an $\mathcal{F}_T$-measurable stochastic Lagrange multiplier $\mu : \Omega \to \mathbb{R}$ (which we will also call a *stochastic multiplier*). Note that $\mu$ must be $\mathcal{F}_T$-measurable, since $M(X(T))$ is $\mathcal{F}_T$-measurable.

Assume that the stochastic multiplier $\mu$ satisfies $E[\mu] < \infty$. Moreover, assume that $E^x[M(X_u(T)] < \infty$ for all $u \in \mathcal{A}$. We introduce a new, but related stochastic control problem

$$\phi_\mu(x) := \sup_{u\in\mathcal{A}} E^x[\int_0^T f(t,X(t),u(t))dt + g(X(T)) + \mu M(X(T))]$$

subject to

$$\begin{aligned} dX(t) &= b(t,X(t),u(t))dt + \sigma(t,X(t),u(t))dB(t) \\ &\quad + \int_{\mathbb{R}} \gamma(t,X(t^-),u(t^-),z)\tilde{N}(dt,dz), \end{aligned} \tag{4.5}$$

and define

$$J^u_\mu(x) := E^x[\int_0^T f(t,X(t),u(t))dt + g(X(T)) + \mu M(X(T))].$$

We also define the set of stochastic multipliers by

$$\Lambda := \{\mu : \Omega \to \mathbb{R} \mid \mu \text{ is } \mathcal{F}_T\text{-measurable and } E[\mu] < \infty\}.$$

Now, we will prove Theorem 4.3, which says that if we can find a solution to the unconstrained problem (4.5) with a stochastic multiplier which ensures that the constraint $M(X(T)) = 0$ a.s. is satisfied, then we have a corresponding solution to our original problem (4.4).

**Theorem 4.3.** *Suppose that we for all $\mu \in \Lambda$ can find $\phi_\mu(x)$ and $u^*_\mu$ solving the unconstrained stochastic control problem (4.5). Moreover, suppose there exists $\mu_0 \in \Lambda$ such that*

$$M(X_{u^*_{\mu_0}}(T)) = 0 \text{ a.s.}$$

*Then,* $\phi(x) := \phi_{\mu_0}(x)$ *and* $u^* := u^*_{\mu_0}$ *solves the constrained stochastic control problem (4.4).*

*Proof.* Let $\mu$ be $\mathscr{F}_T$-measurable. Then,

$$\begin{aligned} &E^x[\int_0^T f(t, u^*_\mu, X_{u^*_\mu})dt + g(X_{u^*_\mu}(T)) + \mu M(X_{u^*_\mu}(T))] = J_\mu^{u^*_\mu}(x) \\ &\geq J_\mu^u(x) = E^x[\int_0^T f(t, u, X_u)dt + g(X_u(T)) + \mu M(X_u(T))] \end{aligned}$$

where the first equality uses the definition of $J_\mu^u$, the inequality uses the definition of $u^*_\mu$ and the final equality uses the definition of $J_\mu^u$.

In particular, if $\mu = \mu_0$ a.s. and $u$ is feasible in the constrained control problem (4.4), then

$$M(X_{u^*_{\mu_0}}(T)) = 0 = M(X_u(T)) \text{ a.s.} \tag{4.6}$$

from the definition of $\mu_0$ and the assumption that $u$ is feasible in problem (4.4).

Hence,

$$\begin{aligned} &J_{\mu_0}^{u^*_{\mu_0}}(x) = E^x[\int_0^T f(t, u^*_{\mu_0}, X_{u^*_{\mu_0}})dt + g(X_{u^*_{\mu_0}}(T)) + \mu_0 M(X_{u^*_{\mu_0}}(T))] \\ &\geq J_{\mu_0}^u(x) = E^x[\int_0^T f(t, u, X_u)dt + g(X_u(T)) + \mu_0 M(X_u(T))] \end{aligned}$$

But $M(X_{u^*_{\mu_0}}(T)) = 0 = M(X_u(T))$ a.s. from equation (4.6), so

$$J^{u^*_{\mu_0}}(x) = J_{\mu_0}^{u^*_{\mu_0}}(x) \geq J_{\mu_0}^u(x) = J^u(x)$$

for all stochastic controls $u$ feasible in the constrained problem (4.4). Note that $u^*_{\mu_0}$ is feasible in problem (4.4), therefore it is an optimal control for this problem. □

Note that problem (4.5) is a stochastic optimal control problem of the form in Øksendal and Sulem [66], with $f_\mu = f$ and $g_\mu(x) = g(x) + \mu M(x)$. Therefore, we may use some known methods of stochastic control, for example the stochastic maximum principle, to solve the problem. Clearly, $f_\mu$ and $g_\mu$ are continuous functions. However, in order to use the stochastic maximum principle, we also need $g_\mu(x)$ to be a concave function. If this is the case, we can solve this problem using the maximum principle for jump processes, Theorem 4.9.

It is irrelevant for this solution strategy whether the unconstrained stochastic control problem coming from the stochastic Lagrange multiplier method is solved using the maximum principle (Theorem 4.9), or some other method of stochastic control. If it is more suitable for the problem, the dynamic programming (Hamilton-Jacobi-Bellman) approach to stochastic control of jump diffusions can also be used, see Øksendal and Sulem [66], Theorem 3.1. For the dynamic programming approach, the concavity of the function $g_\mu = g + \mu M$ is not necessary. However, the problem must have a Markovian structure.

**Remark 4.4.** *Note that the only thing that distinguishes problem (4.5) from an unconstrained version of problem (4.4) (i.e. without the constraint $M(X(T)) = 0$ a.s.) is the $g$-function in the objective function. Hence, when we apply the stochastic maximum principle, Theorem 4.9, the Hamiltonian is equal for these two problems. Therefore, also the control maximizing the Hamiltonian will be equal. The only difference is the terminal condition of the BSDE for the adjoint processes $p, q, r$, equation (4.19). However, this altered terminal condition clearly affects the solution of the adjoint BSDE, and hence can also affect the optimal control process.*

**Remark 4.5.** *It is important that the constraint of the stochastic optimal control problem depends on $X(T)$, i.e. the state process at the* terminal time. *If we had a constraint of the type $M(X(\tilde{t})) = 0$ a.s., where $\tilde{t} < T$, then the new stochastic control problem coming from the Lagrange multiplier method would not fit the setting of Øksendal and Sulem [66]. This complicates matters significantly.*

## 4.3 The economic model: Optimal consumption with Lévy wage

In this section, we derive an economic model which we will use to analyze wage- and inflation risk in an optimal consumption problem with a terminal constraint. We also give interpretations of the two different types of terminal constraints $(i)$ and $(ii)$ (see Section 4.2). In the following Section 4.4, we will apply the stochastic multiplier method of Section 4.2 in order to solve this optimal consumption problem.

As before, let $(\Omega, \mathcal{F}, P)$ be a probability space, where $\Omega$ is the scenario space, $\mathcal{F}$ a respective $\sigma$-algebra and $P$ the probability measure. Consider an agent who is planning for times $t \in [a, T+a]$, $a \geq 0$. Again, let $\{B(t,\omega)\}_{t\in[a,T+a]}$ be a Brownian motion, and $\int_{\mathbb{R}} \gamma_i(t,z,\omega)N_i(dt,dz), i = 1,2$, two pure jump processes independent of $B(t)$ and each other (we will often abbreviate by omitting $\omega$ in the notation). Let $\{\mathcal{F}_t\}_{t\in[a,T+a]}$ be the filtration generated by the Brownian motion and the pure jump processes, including all null sets. In the following, by adapted processes, we mean adapted with respect to this filtration.

An agent in this market receives an exogenous nominal wage, chooses a consumption function and has the possibility to save or borrow (i.e., go long or short) in an asset with risk free nominal payoff. For a specific agent, the starting time of planning is $a \geq 0$, $W_n(t,\omega)$ is the nominal wage rate at time $t$ and $X_n(t,\omega)$ is the nominal level of savings at time $t$. In this economy, $\xi(t,\omega)$ is the time $t$ price of the consumption good, and $S(t)$ is the time $t$ price level of the risk free asset. Our goal is to study the effect of inflation- and wage risk. Hence, to avoid unnecessary complication, we do not include any risky assets. For more about the economic terms, see Romer [90].

The inflation, denoted $\pi(t,\omega)$, can be decomposed into a drift term $\hat{\pi}$, and the deviation from that level. The deviation from $\hat{\pi}$ is stochastic, and modeled by $\Delta(t,\omega)dt = \tilde{\pi}dB(t)$ ($\tilde{\pi}$ is constant), i.e., $\pi(t,\omega) = \hat{\pi} + \Delta(t,\omega)$. Since inflation is defined by the identity

$$d\xi(t,\omega) := \xi(t,\omega)\pi(t,\omega)dt,$$

the price $\xi$ consequently is a geometric Brownian motion. The market is normalized by setting $\xi(0) := 1$.

The changes in the nominal wage rate is a pure jump process

$$dW_n(t) = \xi(t)\int_{\mathbb{R}} zN_1(dt,dz) - (1-\epsilon)W_n(t)\int_{\mathbb{R}} N_2(dt,dz).$$

Here, $N_i(dt,dz)$, $i = 1,2$, represent two independent Poisson random measures, $z > 0$ is the size of each jump, which will always be positive, and $\epsilon \in (0,1)$

is a constant. Thus, we model the nominal wage process such that an agent receives positive wage gains, and bears the risk of losing a portion of wage (for instance by becoming unemployed, but receiving a welfare benefit). The wage gain pressure is proportional to the consumption price, since an agent cares about the real wage, rather than the nominal wage. Using the Lévy-Kintchine representation, these terms are decomposed into martingale and non-martingale expressions. In order to do this decomposition, we assume that $\int_{|z|\geq 1} |z|\nu_i(dz) < \infty$, $i = 1,2$ (alternatively, the stronger assumption of finite variance of the two Lévy processes). Thus,

$$\begin{aligned}\xi(t)\int_{\mathbb{R}} zN_1(dt,dz) &= \xi(t)\int_{\mathbb{R}} z(E[N_1(dt,dz)]+\tilde{N}_1(dt,dz)) \\ &= \xi(t)\Big(\int_{\mathbb{R}} z\nu_1(dz)dt+\int_{\mathbb{R}} z\tilde{N}_1(dt,dz)\Big) \\ &= \xi(t)\Big(\alpha dt+\int_{\mathbb{R}} z\tilde{N}_1(dt,dz)\Big),\end{aligned}$$

where $\alpha := \int_{\mathbb{R}} z\nu_1(dz)$. Similarly

$$(1-\epsilon)W_n(t)\int_{\mathbb{R}} N_2(dt,dz) := \beta(1-\epsilon)W_n(t)dt+(1-\epsilon)W_n(t)\int_{\mathbb{R}}\tilde{N}_2(dt,dz).$$

Here, $\beta := \int_{\mathbb{R}}\nu_2(dz)$. The terms

$$\tilde{N}_i(dt,dz) := N_i(dt,dz)-\nu_i(dz)dt,\ i=1,2,$$

are the compensated Poisson random measures, and $\nu_i(dz)$, $i = 1,2$, are the Lévy measures. For more on these measures and Lévy processes in general, see Øksendal and Sulem [66]. We let the consumption good (with price $\xi(t)$) be the numéraire for all $t\in[a,T+a]$, and define the real wage process by

$$W(t,\omega) := \frac{W_n(t,\omega)}{\xi(t,\omega)}.$$

Then, by Itô's formula (the product rule),

$$dW(t,\omega) = d\left(\frac{W_n(t,\omega)}{\xi(t,\omega)}\right) = \frac{dW_n(t,\omega)}{\xi(t,\omega)} - \frac{W_n(t,\omega)d\xi(t,\omega)}{[\xi(t,\omega)]^2}\ P\text{-a.e.}$$

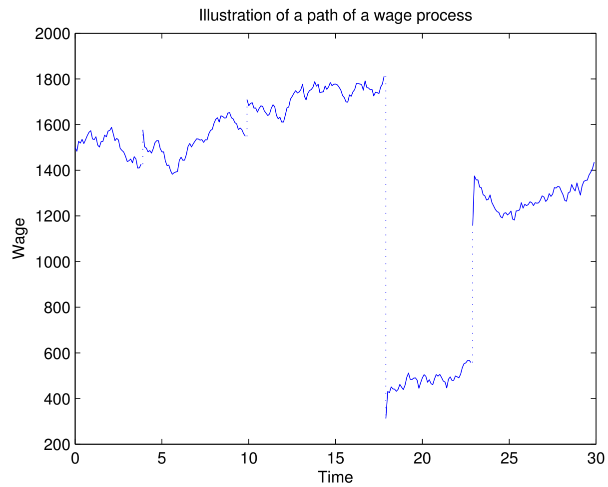


Figure 4.1: A path of a Lévy wage process.

Inserting the relevant terms, we get

$$\begin{aligned} dW(t) &= (\alpha - [\hat{\pi} + \beta(1-\epsilon)]W(t))dt - \tilde{\pi}W(t)dB(t) + \int_{\mathbb{R}} z\tilde{N}_1(dt,dz) \\ &\quad -(1-\epsilon)W(t)\int_{\mathbb{R}} \tilde{N}_2(dt,dz). \end{aligned}$$

Thus, the real wage is a process with both jumps and continuous movements due to the inflation. Figure 4.1 is an illustration of such a process.

The price of the asset is given by

$$dS(t) = r_n S(t)dt,$$

where the interest rate $r_n$ is a constant. Since this is the only financial object of this market, nominal savings is given by $X_n(t) = \eta S(t)$, where $\eta$ is the number of assets held by the agent. Nominal savings are assumed to be $\Gamma$-financing, meaning that $dX_n(t) = (W_n(t) - \xi(t)c(t))dt + r_n X_n(t)dt$ (this corresponds to the budget constraint), where $\{c(t)\}_{t\in[a,T+a]}$ is an adapted stochastic process representing the real consumption rate of the agent at time $t$.

Define the real value of savings

$$X(t,\omega) := \frac{X_n(t,\omega)}{\xi(t,\omega)}.$$

Written in differential form (using Itô's formula), we get

$$\begin{aligned} dX(t,\omega) = d\left(\tfrac{X_n(t,\omega)}{\xi(t,\omega)}\right) = \tfrac{dX_n(t,\omega)}{\xi(t,\omega)} - X(t,\omega)\pi(t,\omega)dt \\ = \tfrac{r_n X_n(t,\omega)dt + W_n(t,\omega)dt - \xi(t,\omega)c(t)dt}{\xi(t,\omega)} - X(t,\omega)\pi(t,\omega)dt. \end{aligned}$$

Thus, the real value of savings is

$$dX(t,\omega) = (r_n - \hat{\pi})X(t)dt - \tilde{\pi}X(t)dB(t) + W(t)dt - c(t)dt.$$

Given the market situation just described, an agent planning consumption and saving at a given time $a \geq 0$ would like to solve the following stochastic optimal control problem:

$$\begin{array}{ll} \sup_{\{c(t)\geq 0 \ \forall \ t \ a.s.\}} & E[\int_a^{T+a} e^{-\delta(t-a)} u(c(t))dt] \\ \text{subject to} & \end{array} \tag{4.7}$$

$$\begin{array}{rcl} dX(t) & = & (r_n - \hat{\pi})X(t)dt - \tilde{\pi}X(t)dB(t) + W(t)dt - c(t)dt, \\ dW(t) & = & (\alpha - [\hat{\pi} + \beta(1-\epsilon)]W(t))dt - \tilde{\pi}W(t)dB(t) \\ & & + \int_{\mathbb{R}} z\tilde{N}_1(dt,dz) - (1-\epsilon)W(t)\int_{\mathbb{R}} \tilde{N}_2(dt,dz) \\ X(a) & = & x_a, \ W(a) = w_a \\ \text{(i) } E[X(T+a)] & \geq & K, \quad or \quad \text{(ii) } X(T+a) \geq K \text{ a.s.,} \end{array}$$

where $K \leq 0$ is a given constant, and either condition (i) or condition (ii) holds. We call this problem (OCP) (optimal consumption problem).

In problem 4.7, the function $u(\cdot): \mathbb{R}_+ \mapsto \mathbb{R}_+$ is the utility function of the agent, which we later will assume is of CRRA form, i.e. $u(c) = \frac{c^\gamma}{\gamma}$, where $\gamma < 1$. Furthermore, $\delta > 0$ is the agent's time preference discount factor and, as mentioned, $\{X(t)\}_{t\in[a,T+a]}$ is the stochastic saving process for the time $t$ amount

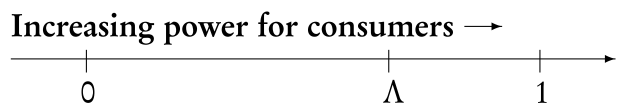


Figure 4.2: Illustration of the two constraints

of real wealth placed in the bank. Moreover, $\{W(t)\}_{t\in[a,T+a]}$ is the stochastic wage process for the time $t$ real wage level. We assume that $X(a)=x_a>0$ and $W(a)=w_a\geq 0$. These two initial levels are both exogenous. Note also that the wage process $W$ is given exogenously, while the agent can control the process $X$.

**Remark 4.6.** *Note that the constraints in problem (4.7) correspond to choosing $M(x)=x-K$ in the notation of Section 4.2, where $K$ is a constant.*

Regarding the two versions of the terminal condition, condition (i) is very mild for the agent. In this case the bank takes on all the market risk. On the other hand, condition (ii) is a strict constraint from the agent's point of view, since it says that the agent must end up with greater than or equal $K$ at the final time a.s. When considering this constraint, the agent bears all market risk. Reality is most likely somewhere in-between the two extremes (i) and (ii). In general, one could introduce a risk sharing parameter $\Lambda\in[0,1]$, and introduce a constraint of the form $\Lambda E[X(T+a)]+(1-\Lambda)X(T+a)\geq K$ a.s. Such a parameter $\Lambda$ characterizes the power of the banks in the market. Note that $\Lambda=1$ is constraint (i), while $\Lambda=0$ gives constraint (ii). Hence, the smaller $\Lambda$ is, the greater is the power of the banks in the market. See Figure 4.2 for an illustration.

The objective is now to solve the stochastic optimal control problem (4.7), i.e. find an expression for the optimal consumption process given the dynamics and the terminal condition on savings.

## 4.4 Stochastic multiplier approach and (OCP)

In this section, we return to our original problem (4.7) with constraint $(i)$. This problem will be solved using the techniques of Section 4.2.1.

Problem (4.7) with constraint $(i)$ is equivalent to the problem where the constraint is binding, i.e. $E[X(T+a)]=K$. To prove this, assume we have an optimal control $\tilde{c}$ for problem (4.7) where $E[X_{\tilde{c}}(T+a)]>K$. We will show that there exists an $\epsilon>0$ such that $c:=\tilde{c}+\epsilon$ satisfies $E[X_c(T+a)]=K$ (i.e. it is a feasible control in problem (4.7)). By Lemma 4.10 in the Appendix and the definition of $c$,

$$X_c(t)=X_{\tilde{c}}(t)-\epsilon F,$$

where $F$ is a known, positive random variable composed of exponentials and depending only on the parameters of the model, see equation (4.21). Hence,

$$E[X_c(T+a)]=E[X_{\tilde{c}}(T+a)]-\epsilon E[F].$$

Let $A:=E[X_{\tilde{c}}(T+a)]-K>0$ (by assumption). Then, $E[X_c(T+a)]-K=A-\epsilon E[F]$. We would like to choose $\epsilon$ such that $E[X_c(T+a)]-K=0$, that is $A-\epsilon E[F]=0$, i.e.

$$\epsilon=\frac{A}{E[F]}>0 \text{ (positive since } E[F]>0).$$

Thus, the consumption process $c=\tilde{c}+\epsilon$ is feasible in problem (4.7). By the definition of $c$, $J(c)>J(\tilde{c})$, so since $c$ is feasible, $\tilde{c}$ cannot have been an optimal control.

Thus, the problem becomes

$$\sup_{\{c(t)\geq 0\ \forall\ t\ a.s.\}} \quad E\Big[\int_a^{T+a} e^{-\delta(t-a)}u(c(t))dt\Big]$$

subject to

$$\begin{aligned}
dX(t) &= (r_n - \hat{\pi})X(t)dt - \tilde{\pi}X(t)dB(t) + W(t)dt - c(t)dt \\
dW(t) &= (\alpha - [\hat{\pi} + \beta(1-\epsilon)]W(t))dt - \tilde{\pi}W(t)dB(t) \\
&\quad + \int_{\mathbb{R}} z\tilde{N}_1(dt,dz) - (1-\epsilon)W(t)\int_{\mathbb{R}} \tilde{N}_2(dt,dz) \\
E[X(T+a)] &= K,\ X(a) = x_a,\ W(a) = w_a.
\end{aligned}$$

We rewrite this problem as an unconstrained two-dimensional stochastic control problem using the stochastic multiplier method of Section 4.2.1,

$$\sup_{\{c(t)\geq 0\ \forall\ t\ a.s.\}} \quad E\left[\int_a^{T+a} e^{-\delta(t-a)}u(c(t))dt + \lambda(X(T+a) - K)\right]$$

subject to

$$\begin{aligned}
dY(t) &= \begin{bmatrix} (r_n - \hat{\pi})X(t) + W(t) - c(t) \\ \alpha - [\hat{\pi} + \beta(1-\epsilon)]W(t) \end{bmatrix} dt + \begin{bmatrix} -\tilde{\pi}X(t) \\ -\tilde{\pi}W(t) \end{bmatrix} dB(t) \\
&\quad + \int_{\mathbb{R}} \begin{bmatrix} 0 & 0 \\ z & -(1-\epsilon)W(t) \end{bmatrix} \begin{bmatrix} \tilde{N}_1(dt,dz) \\ \tilde{N}_2(dt,dz) \end{bmatrix} \\
Y(a) &= \begin{bmatrix} x_a \\ w_a \end{bmatrix}
\end{aligned} \tag{4.8}$$

where $Y(t) := (X(t), W(t))^T$, and $\lambda \in \mathbb{R}$ is a Lagrange multiplier introduced to handle the constraint $E[X(T+a)] = K$.

We solve problem (4.8) using the stochastic maximum principle for jump diffusions from Øksendal and Sulem [66], Theorem 4.9. Note that this theorem is easily generalized to our setting, where we start at time $t = a$ instead of $t = 0$. We apply a subjective current value version of the Hamiltonian function, i.e. $\tilde{H} = He^{\delta(t-a)}$. All the corresponding adjoint functions will be marked with a tilde to emphasize this change.

In this case, the subjective current value Hamiltonian function is

$$\begin{aligned}
\tilde{H}(t,y,c,\tilde{p},\tilde{q},\tilde{r}) &= u(c) + \tilde{p}_1(\{r_n - \hat{\pi}\}x + w - c) + \tilde{p}_2(\alpha - [\hat{\pi} + \beta(1-\epsilon)]w) \\
&\quad -\tilde{q}_1\tilde{\pi}x - \tilde{q}_2\tilde{\pi}w + \int_{\mathbb{R}} z\tilde{r}_{21}(t,z)\nu_1(dz_1) - \int_{\mathbb{R}}(1-\epsilon)w\tilde{r}_{22}(t,z)\nu_2(dz_2)
\end{aligned}$$

where the adjoint processes $\tilde{p}(t) := (\tilde{p}_1(t), \tilde{p}_2(t))^T$, $\tilde{q}(t) := (\tilde{q}_1(t), \tilde{q}_2(t))^T$ and $\tilde{r} \in \mathbb{R}^{2\times 2}$ is the matrix with components $\tilde{r}_{ij}(t,z)$, $i,j \in \{1,2\}$ for $t \in [a, T+a]$,

$z \in \mathbb{R}$.

The set of adjoint backward stochastic differential equations (BSDEs) corresponding to this Hamiltonian is

$$\begin{aligned} d\tilde{p}_1(t) &= -\tilde{p}_1(t)(r_n - \hat{\pi} - \delta)dt + \tilde{q}_1(t)\tilde{\pi}dt + \tilde{q}_1(t)dB(t) \\ &\quad + \int_{\mathbb{R}} \tilde{r}_{11}(t,z_1)\tilde{N}_1(dt,dz_1) + \int_{\mathbb{R}} \tilde{r}_{12}(t,z_2)\tilde{N}_2(dt,dz_2) \\ \tilde{p}_1(T+a) &= \tilde{\lambda} \end{aligned} \tag{4.9}$$

$$\begin{aligned} d\tilde{p}_2(t) &= \{-\tilde{p}_1(t) + [\hat{\pi} + \beta(1-\epsilon) + \delta]\tilde{p}_2(t) \\ &\quad + \tilde{q}_2(t)\tilde{\pi} + \int_{\mathbb{R}}(1-\epsilon)\tilde{r}_{22}\nu_2(dz_2)\}dt + \tilde{q}_2(t)dB(t) \\ &\quad + \int_{\mathbb{R}} \tilde{r}_{21}(t,z_1)\tilde{N}_1(dt,dz_1) + \int_{\mathbb{R}} \tilde{r}_{22}(t,z_2)\tilde{N}_2(dt,dz_2) \\ \tilde{p}_2(T+a) &= 0. \end{aligned}$$

The first order condition for the maximization of the Hamiltonian is:

$$\tilde{p}_1(t) = u'(\hat{c}(t)) = \hat{c}^{\gamma-1}(t),$$

where the final equality holds for CRRA utility.

To determine the optimal consumption, it suffices to find the adjoint process $\tilde{p}_1(t)$. This process is possibly stochastic due to the randomness in wage and inflation. Now, since $\tilde{\lambda}$ is constant, $\tilde{p}_1(t)$ has a deterministic terminal value. Note that $\tilde{p}_1(t) = \tilde{\lambda}\exp(\{r_n - \hat{\pi} - \delta\}(T+a-t))$, $\tilde{q}_1(t) = 0$, $\tilde{r}_{11}(t,z) = \tilde{r}_{12}(t,z) = 0$ for $t \in [a, T+a]$, $z \in \mathbb{R}$ is the solution of equation (4.9). So, $\tilde{p}_1(t)$ is deterministic. Thus, equivalently to equation (4.4), we have

$$\tilde{\lambda}e^{(r_n - \hat{\pi} - \delta)(T+a-t)} = \tilde{p}_1(t) = u'(\hat{c}(t)) = \hat{c}^{(\gamma-1)}(t). \tag{4.10}$$

To determine the Lagrange multiplier $\tilde{\lambda}^*$ such that $E[X_{\hat{c}}(T)] = K$, we solve the stochastic differential equation (SDE) for the savings process $X(t)$:

$$\begin{aligned} dX(t) &= \big((W(t) - \hat{c}(t)) + (r_n - \hat{\pi})X(t)\big)dt - \tilde{\pi}X(t)dB(t) \\ X(a) &= x_a. \end{aligned} \tag{4.11}$$

From Lemma 4.10, the solution of equation (4.11) is

$$X(t) = x_a e^{R(t)-R(a)} + \int_a^t e^{R(t)-R(s)}(W(s) - \hat{c}(s))ds, \tag{4.12}$$

where we for notational convenience define the stochastic processes

$$\begin{aligned} \Pi(t) &:= \hat{\pi}t + \tfrac{1}{2}\tilde{\pi}^2 t + \tilde{\pi}B(t), \\ R(t) &:= r_n t - \Pi(t). \end{aligned} \tag{4.13}$$

We also need to solve the stochastic differential equation for the wage process. This solution is given by the following lemma.

**Lemma 4.7.** *The solution of the stochastic differential equation*

$$\begin{aligned} dW(t) &= (\alpha - [\hat{\pi} + \beta(1-\epsilon)]W(t))dt - \tilde{\pi}W(t)dB(t) + \int_{\mathbb{R}} z\tilde{N}_1(dt,dz) \\ &\quad -(1-\epsilon)W(t)\int_{\mathbb{R}} \tilde{N}_2(dt,dz) \\ W(a) &= w_a \end{aligned} \tag{4.14}$$

*is*

$$W(t) = w_a e^{-(\zeta(t)-\zeta(a))} + \int_a^t \alpha e^{-(\zeta(t)-\zeta(s))} ds + \int_a^t \int_{\mathbb{R}} z e^{-(\zeta(t)-\zeta(s))}\tilde{N}_1(ds,dz), \tag{4.15}$$

*where*

$$\zeta(t) := \Pi(t) + \beta(1-\epsilon)t - \ln(\epsilon)\int_a^t \int_{\mathbb{R}} \tilde{N}_2(du,dz).$$

For the proof of Lemma 4.7, see the Appendix.

The $\Pi(t)$ occurring in $\zeta(t)$ in Lemma 4.7 adjusts for changes in the numéraire consumption good, while the term $\beta(1-\epsilon)t - \ln(\epsilon)\int_a^t \int_{\mathbb{R}} \tilde{N}_2(du,dz)$ represents the geometric effect of unemployment risk.

By inserting the wage expression (4.15) into $X_{\hat{c}}(T+a)$, we find

$$
\begin{aligned}
X_{\hat{c}}(T+a) &= x_a e^{R(T+a)-R(a)} + \int_0^T w_a e^{R(T+a)-R(t+a)} e^{-(\zeta(t+a)-\zeta(a))} dt \\
&+ \int_0^T \Big( \int_0^t \alpha e^{R(T+a)-R(t+a)} e^{-(\zeta(t+a)-\zeta(s+a))} ds \Big) dt \\
&+ \int_0^T \Big( \int_0^t \int_{\mathbb{R}} z e^{R(T+a)-R(t+a)} e^{-(\zeta(t+a)-\zeta(s+a))} \tilde{N}_1(ds,dz) \Big) dt \\
&- \int_0^T \tilde{\lambda}^{\frac{1}{\gamma-1}} e^{R(T+a)-R(t+a)} e^{\frac{(r_n-\hat{\pi}-\delta)(T-t)}{\gamma-1}} dt.
\end{aligned}
$$

Now, we would like to choose $\tilde{\lambda} \in \mathbb{R}$ such that $E[X_{\hat{c}}(T+a)] = K$. Since $\tilde{\lambda}$ is a constant, it is separable from the integral containing it. The equation can thus be solved, and we find

$$
\tilde{\lambda}^* = (\hat{\mathscr{B}}^{-1}\, \hat{\mathscr{C}})^{(\gamma-1)},
$$

where

$$
\begin{aligned}
\hat{r} &:= r_n - \hat{\pi}, \\
\hat{\mathscr{C}} &:= x_a e^{\hat{r}T} + \int_0^T w_a E[e^{R(T+a)-R(t+a)-(\zeta(t+a)-\zeta(a))}] dt \\
&\quad + \int_0^T \Big( \int_a^t \alpha E[e^{R(T+a)-R(t+a)-(\zeta(t+a)-\zeta(s+a))}] ds \Big) dt - K, \\
\hat{\mathscr{B}} &:= \Big( \int_0^T e^{(\hat{r}-\hat{\Gamma})(T-t)} dt \Big) = \frac{e^{(\hat{r}-\hat{\Gamma})T}-1}{\hat{r}-\hat{\Gamma}},
\end{aligned}
$$

and $\hat{\Gamma} := \frac{\hat{r}-\delta}{1-\gamma}$.

Intuitively, $\hat{\mathscr{C}}$ is the expectation of the time $T+a$ future value of consumption, and $\hat{\mathscr{B}}^{-1}$ is a time $T+a$ future value risk aversion weight of the expected subjective real interest rate.

From this, the candidate for the optimal consumption process is

$$
\hat{c}(t) = \hat{\mathscr{B}}^{-1} \cdot \hat{\mathscr{C}} \cdot e^{-\hat{\Gamma}(T+a-t)},\ t \in [a, T+a]. \tag{4.16}
$$

Now, we would like to apply the stochastic maximum principle, see the Appendix Theorem 4.9, to conclude that $\hat{c}(t)$ actually is the optimal consumption process. In order to do this, the boundedness conditions of Theorem 4.9 must hold. However, this is the case for our problem. The functions $\sigma$ and $\gamma$ (for our specific problem) are continuous, and the Brownian motion $B(t)$ has a continuous version. Since the conditions of Theorem 4.9 involve the expectation of the integral over a finite time interval, the integral of these processes will be finite. The adjoint processes $\tilde{p}$ and $\tilde{q} = \tilde{r} = 0$ do not cause any finiteness-problems, hence, we may use the stochastic maximum principle.

Therefore, by the stochastic maximum principle, Theorem 4.9, the consumption process $\hat{c}$ is optimal if it is feasible, i.e. if $\hat{c}(t) \geq 0$ for all $t \in [a, T+a]$. Clearly, since $K \leq 0$, this holds. Therefore, $c^*(t) := \hat{c}(t)$, $t \in [a, T+a]$ is the optimal consumption process.

Note that when we are using a soft end constraint (type $(i)$), there are no stochastic elements in the resulting optimal control. The result is that $\lambda$ is a constant, thus the shadow price of the constraint $K$ is constant. Furthermore the adjoint function $\tilde{p}_1(t)$ also becomes deterministic, making the optimal consumption process deterministic. With the soft end constraint, the consumer makes a consumption plan at time $a$ only expecting to reach the terminal level $K$. Then, independently of which outcomes of wage and inflation are realized, the consumer adjusts savings, keeping the same consumption path. The result is that the bank bears all market risk, and the value function of the consumer is independent of which states are realized.

To conclude, the soft end constraint implies that the consumer is unaffected by the uncertainty in inflation and wage level, in the sense that only expected total income and expected real interest rate are relevant for the consumption plan, and thus the value function.

**Remark 4.8.** *The version of problem (4.7) with a type $(ii)$ constraint (i.e. an almost sure constraint on terminal savings) is more complicated. The technique is the same as before, except for determining the Lagrange multiplier $\tilde{\mu}^*$. Now, we would like to find $\tilde{\mu}^*$ s.t. $X_{\hat{c}}(T+a) = K$ a.s. To do this, we insert the expressions for $W$ and $\hat{c}$ into $X(T+a)$. Due to the stochastic nature of $\tilde{\mu}$, the Lagrange multiplier is no longer separable from the integral expression containing it, and hence it is more difficult to determine. However, the adjoint variables have the same interpretations as before. By this we deduce that the stochastic Lagrange multiplier has the ordinary shadow price interpretation, but in this case (with a hard end constraint) the shadow price of $K$ is a random variable, as the marginal utility of changing $K$ will depend on which states are realized.*

*The value function depends on the changes in the inflation and wage level, and thus it is stochastic. The consumer with the a.s. constraint bears all the market risk. This implies that consumption is strongly influenced by the uncertainty in the inflation and wage level, and in this case the lender bears no risk at all. The possibility of a loss of wage or an increased inflation has a negative effect for a power utility consumer, even though the consumer may be better off than in the soft end constraint case.*

## 4.5 Concluding remarks

In this paper we have derived a stochastic Lagrange multiplier method, and showed how this can be combined with the stochastic maximum principle for jump diffusions to solve constrained stochastic optimal control problems. As an application, we have studied an optimal consumption problem with inflation- and wage risk.

We have observed that the Lagrange multiplier behaves very differently, depending on whether the end constraint holds in expectation (type $(i)$ constraint) or almost surely (type $(ii)$). This affects the value function of the problem. In the first case, the value function is deterministic and almost identical to a version of the problem without uncertainty (see Sydsæther et al. [94] for such a problem). In the second case, the value function depends on the changes in the inflation and wage level, and thus it is stochastic. By economic intuition, this is equivalent to a consumer bearing no risk in the first case, and all risk in the second case. Hence, the behavior of the consumer is unaffected by inflation and wage risk when there is a terminal constraint which needs to hold in expectation, while the behavior is strongly affected by the risk when there is an a.s. end constraint.

The fact that different terminal constraints influence the outcome of the stochastic optimal consumption problem should be taken into consideration when stochastic optimal control is applied for analysis of practical problems, especially if the stochastic components have large variations, or when worst case scenarios are of interest.

**Acknowledgments:** We would like to thank an anonymous reviewer for helpful suggestions. We are also very grateful for the help of our supervisors, Professor Bernt Øksendal and Professor Kjell Arne Brekke, both at the University of Oslo.

## 4.6 Appendix: Some results from stochastic analysis

Let $f : \mathbb{R}_+ \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$ be given, continuous functions. Consider the stochastic optimal control problem

$$\begin{aligned} &\sup_{u \in \mathcal{A}} \quad E\left[\int_0^T f(t, X(t), u(t))dt + g(X(T))\right] \\ &\text{subject to} \\ &\qquad dX(t) = b(t, X(t), u(t))dt + \sigma(t, X(t), u(t))dB(t) \\ &\qquad\qquad + \int_{\mathbb{R}} \gamma(t, X(t^-), u(t^-), z)\tilde{N}(dt, dz) \\ &\qquad X(0) = x \end{aligned} \tag{4.17}$$

where $\mathcal{U} \subset \mathbb{R}$ is a given set, $b : \mathbb{R}_+ \times \mathbb{R} \times \mathcal{U} \to \mathbb{R}$, $\sigma : \mathbb{R}_+ \times \mathbb{R} \times \mathcal{U} \to \mathbb{R}$ and $\gamma : \mathbb{R}_+ \times \mathbb{R} \times \mathcal{U} \times \mathbb{R} \to \mathbb{R}$. The control $u(t) = u(t, \omega) : \mathbb{R}_+ \times \Omega \to \mathcal{U}$ is admissible, denoted $u \in \mathcal{A}$, if the dynamics of $X$ has a unique, strong solution for all $x \in \mathbb{R}$ and $E^x[\int_0^T f(t, X(t), u(t))dt + g(X(T))] < \infty$.

In the following theorem, the function $H : [0, T] \times \mathbb{R} \times \mathcal{U} \times \mathbb{R} \times \mathbb{R} \times \mathcal{R} \to \mathbb{R}$ is the Hamiltonian function, defined by

$$H(t, x, u, p, q, r) = f(t, x, u) + b(t, x, u)p + \sigma(t, x, u)q + \int_{\mathbb{R}} \gamma(t, x, u, z)\nu(dz) \tag{4.18}$$

where $\mathcal{R}$ is the set of functions such that the integral above converges.

**Theorem 4.9.** *(A sufficient maximum principle for stochastic optimal control with jumps, Framstad et al. [29])*

*Let $\hat{u}$ be an admissible control, i.e. $\hat{u} \in \mathcal{A}$, with corresponding state process $\hat{Y} = Y^{\hat{u}}$ and suppose there exists a solution $(\hat{p}(t), \hat{q}(t), \hat{r}(t, z))$ of the corresponding adjoint equation*

$$\begin{aligned} dp(t) = & -\nabla_y H(t, Y(t), u(t), p(t), q(t), r(t,\cdot))dt \\ & +q(t)dB(t) + \int_{\mathbb{R}} r(t^-, z)\tilde{N}(dt, dz), \qquad t < T, \\ p(T) = & \nabla g(Y(T)) \end{aligned} \tag{4.19}$$

*satisfying*

$$E\Big[\int_0^T (\hat{Y}(t) - Y^u(t))^2 \Big\{\hat{q}^2(t) + \int_{\mathbb{R}} r^2(t,z)\nu(dz)\Big\}dt\Big] < \infty$$

*and*

$$\begin{aligned} & E\Big[\hat{p}^2(t)\Big\{\sigma^2(t, Y^u(t), u(t)) + \int_{\mathbb{R}} \gamma^2(t, Y^u(t), u(t), z)\nu(dz)\Big\}dt\Big] \\ & \quad < \infty \text{ for all } u \in \mathscr{A}. \end{aligned}$$

*Moreover, suppose that*

$$H(t, \hat{Y}(t), \hat{u}(t), \hat{p}(t), \hat{q}(t), \hat{r}(t,\cdot)) = \sup_{v \in \mathscr{U}} H(t, \hat{Y}(t), v, \hat{p}(t), \hat{q}(t), \hat{r}(t,\cdot))$$

*for all $t$, that $g$ is a concave function of $y$ and that*

$$\hat{H}(y) := \max_{v \in \mathscr{U}} H(t, y, v, \hat{p}(t), \hat{q}(t), \hat{r}(t,\cdot))$$

*exists and is a concave function of $y$, for all $t \in [0, T]$. Then, $\hat{u}$ is an optimal control.*

*Proof.* See Framstad et al. [29]. □

**Lemma 4.10.** *Consider the stochastic differential equation,*

$$\begin{aligned} dX(t) &= (\mu(t) + \alpha X(t))dt + \beta X(t)dB(t) \\ X(a) &= x_a \end{aligned} \tag{4.20}$$

*where $\mu(t)$ is an adapted stochastic process and $\alpha, \beta \in \mathbb{R}$.*

*Then,*

$$\begin{aligned} X(t) &= \exp(\alpha t - \tfrac{1}{2}\beta^2 t + \beta B(t))\big(x_a \exp\{-\big(\alpha a - \tfrac{1}{2}\beta^2 a + \beta B(a)\big)\} \\ &\quad + \int_a^t \exp(-\alpha s + \tfrac{1}{2}\beta^2 s - \beta B(s))\mu(s)ds\big). \end{aligned} \tag{4.21}$$

*Proof.* The idea is to get rid of the terms of the SDE involving $X(t)$ by multiplying with the integrating factor

$$J(t) := \exp\{-\big(\alpha t + \frac{1}{2}\beta^2 t - \beta B(t)\big)\}.$$

By Itô's product rule (see Exercise 4.3 i Øksendal [64]),

$$d(X(t)J(t)) = X(t)dJ(t) + J(t)dX(t) + dX(t)dJ(t).$$

Itô's formula implies that

$$dJ(t) = \big((-\alpha + \beta^2)dt - \beta dB(t)\big)\exp(\alpha t + \frac{1}{2}\beta^2 t - \beta B(t)).$$

Hence,

$$d(X(t)J(t)) = \exp(\alpha t + \frac{1}{2}\beta^2 t - \beta B(t))\mu(t)dt.$$

So, by integrating from $a$ to $t$ on both sides and multiplying by $\frac{1}{J(t)}$, we find the solution

$$\begin{aligned} X(t) &= \exp(\alpha t - \tfrac{1}{2}\beta^2 t + \beta B(t))\big(x_a \exp\{-\big(\alpha a - \tfrac{1}{2}\beta^2 a + \beta B(a)\big)\} \\ &\quad + \int_a^t \exp(-\alpha s + \tfrac{1}{2}\beta^2 s - \beta B(s))\mu(s)ds\big). \end{aligned}$$

This concludes the proof. □

**Lemma 4.11.** *(Solution of linear BSDE, Proposition 1.3, El Karoui et al. [27]) Consider a linear BSDE of the form*

$$\begin{aligned} -dY(t) &= (\phi(t) + Y(t)\beta(t) + Z(t)\mu(t))dt - Z(t)dB(t) \\ Y(\bar{T}) &= \xi(\bar{T}) \end{aligned}$$

*where $\xi(\bar{T})$ is an $\mathscr{F}_{\bar{T}}$-measurable random variable. This BSDE has a unique solution $(Y, Z)$, where $Y$ is explicitly given by*

$$Y(t) = E[\xi(\bar{T})\Gamma_{t,\bar{T}} + \int_t^{\bar{T}} \Gamma_{t,s}\phi(s)ds|\mathscr{F}_t]$$

*where*

$$\begin{aligned} d\Gamma_{t,s} &= \Gamma_{t,s}(\beta(s)ds + \mu(s)dB(s)) \\ \Gamma_{t,t} &= 1,\ \Gamma_{t,s}\Gamma_{s,u} = \Gamma_{t,u}\ \forall\ t \leq s \leq u. \end{aligned}$$

*Proof.* See El Karoui et al. [27]. □

Finally, we give the proof of Lemma 4.7.

*Proof.* (Proof of Lemma 4.7) We solve the SDE by multiplying with the integrating factor

$$J(t) = \exp\left([\hat{\pi} + \beta(1-\epsilon)]t + \tilde{\pi}B(t) + \frac{1}{2}\tilde{\pi}^2 t - \ln(\epsilon)\int_a^t \int_{\mathbb{R}} \tilde{N}_2(du, dz)\right),$$

(chosen to get rid of the terms involving $W$ in equation (4.14)). By the Itô product rule for jump processes (see Øksendal and Sulem [66], Exercise 1.2),

$$\begin{aligned} d(W(t)J(t)) &= J(t)\big(\alpha - [\hat{\pi} + \beta(1-\epsilon)]W(t)\big)dt - J(t)\tilde{\pi}W(t)dB(t) \\ &\quad + W(t)J(t)\big(\hat{\pi} + \beta(1-\epsilon) + \tilde{\pi}^2\big)dt + W(t)J(t)\tilde{\pi}dB(t) \\ &\quad - \tilde{\pi}^2 W(t)J(t)dt + \int_{\mathbb{R}} J(t-)z\tilde{N}_1(ds, dz) \\ &\quad + \int_{\mathbb{R}} \big(-(1-\epsilon)W(t)J(t)(\tfrac{1}{\epsilon} - 1) + W(t)J(t)(\tfrac{1}{\epsilon} - 1) \\ &\quad - W(t)J(t)(1-\epsilon)\big)\tilde{N}_2(dt, dz) \\ &= \alpha J(t)dt + \int_{\mathbb{R}} zJ(t-)\tilde{N}_1(dt, dz). \end{aligned}$$

So,

$$W(t)J(t) = w_a J(a) + \int_a^t \alpha J(s)ds + \int_a^t \int_{\mathbb{R}} zJ(s-)\tilde{N}_1(ds, dz),$$

which gives the solution in equation (4.15). □

# Bibliography


[1] Aase, K., Øksendal, B., Privault, N. and Ubøe, J., White noise generalizations of the Clark-Haussmann-Ocone theorem with application to mathematical finance, Finance and Stochastics, 4, 465-496, 2000.

[2] Agram, N. and Øksendal, B., Malliavin calculus and optimal control of stochastic Volterra equations. Journal of Optimization Theory and Applications, 167, 1070-1094, 2015.

[3] Allen, E., *Modelling with Itô Stochastic Differential Equations*, Springer, The Netherlands, 2007.

[4] Anderson, E. J. and Nash, P., *Linear Programming in Infinite Dimensional Spaces: Theory and Applications*, Wiley-Interscience Series in Discrete Mathematics and Optimization, 1987.

[5] Applebaum, D., *Lévy Processes and Stochastic Calculus*, second ed., Cambridge University Press, Cambridge, 2009.

[6] Banos, D., Cordoni, F., Di Nunno, G., Di Persio, L. and Røse, E., A note on the computation of the delta in a market with memory driven by Lévy noise, in preparation, 2016.

[7] Battocchio, P., Menoncin, F., Optimal pension fund management in a stochastic framework, Insurance: Mathematics and Economics, 34, 79-95, 2004.

[8] Bertsekas, D. P., Nedic, A. and Ozdaglar, A.E., *Convex Analysis and Optimization*, first ed., Athena Scientific, Belmont, 2003.

[9] Biagini, F. and Øksendal, B., A general stochastic calculus approach to insider trading, Applied Mathematics and Optimization, 52, 167-181, 2005.

[10] Bouchard, B., No-arbitrage in discrete-time markets with proportional transaction costs and general information structure, Finance and Stochastics, 10, 276-297, 2006.

[11] Boyd, S., Vandenberghe, L., *Convex Optimization*, first ed., Cambridge University Press, Cambridge, 2004.

[12] Buckdahn, R., Djehiche, B., Li, J., A general stochastic maximum principle for SDEs of mean-field type Applied Mathematics and Optimization, 64, 197-216, 2011.

[13] Buckwar, E., Introduction to the numerical analysis of stochastic delay differential equations, Journal of Computational and Applied Mathematics, 125, 297-307, 2000.

[14] Carletti, M., Numerical solution of stochastic differential problems in the biosciences, Journal of Computational and Applied Mathematics, 185, 422-440, 2006.

[15] Cvitanić, J., Karatzas, I., Hedging contingent claims with constrained portfolios, Annuals of Applied Probability, 3, 652-681, 1993.

[16] Dahl, K., Mohammed, S., Øksendal, B. and Røse, E., Optimal control of systems with noisy memory and BSDEs with Malliavin derivatives, submitted, arXiv:1403.4034, 2015.

[17] Dahl, K. and Øksendal, B., Singular recursive utility, submitted, arXiv:1504.08170, 2015.

[18] Delbaen, F., Schachermayer, W., A general version of the fundamental theorem of asset pricing, Mathematische Annalen, 300, 463-520, 1994.

[19] De Vallière, D., Kabanov, Y., Stricker, C., No-arbitrage criteria for financial markets with transaction costs and incomplete information, Finance and Stochastics, 11, 237-251, 2007.

[20] Devolder, O., Glineur, F. and Nesterov, Y., Solving infinite-dimensional optimization problems by polynomial approximation, Center for Research and Econometrics Discussion Paper, 2010.

[21] Diehl, J. and Friz, P., Backward stachastic differential equations with rough drivers, The Annals of Probability, 4, 1715-1758, 2012.

[22] Ding, X. and Wu, R., A new proof for comparison theorems for stochastic differential inequalities with respect to semimartingales, Stochastic Processes and their Applications, 78, 155-171, 1998.

[23] Di Nunno, G., Øksendal, B. and Proske, F., *Malliavin Calculus for Lévy Processes with Applications to Finance*, Springer, Berlin Heidelberg, 2009.

[24] Di Nunno, G., Kohatsu-Higa, A., Meyer-Brandis, T., Øksendal, B., Proske, F.N. and Sulem, A., Anticipative stochastic control for Lévy processes with application to insider trading, Mathematical Modelling and Numerical Methods in Finance, 15, Amsterdam: North Holland, 573-595, 2008.

[25] Duffie, D. and Epstein, L.G., Stochastic differential utility, Econometrica, 60, 353-394, 1992.

[26] El Karoui, N. and Jeanblanc-Picqué, M., Optimization of consumption with labor income, Finance and Stochastics, 2, 409-440, 1998.

[27] El Karoui, N., Peng, S. and Quenez, M., Backward stochastic differential equations in finance. Mathematical Finance, 7, 1-71, 1997.

[28] Föllmer, H., Kramkov, D., Optional Decompositions under Constraints, Probability Theory and Related Fields, 109, 1-25, 1997.

[29] Framstad, N. C., Øksendal, B. and Sulem, A., Sufficient stochastic maximum principle for optimal control of jump diffusions and applications

to finance, Journal of Optimization Theory and Application, 121, 77-98, 2007.

[30] Freidlin, M., *Functional Integration and Partial Differential Equations*, Princeton University Press, Princeton NJ, USA, 1985.

[31] Frittelli, M., Rozassa Gianin, E., Putting order in risk measures. Journal of Banking and Finance, 26, 1473-1486, 2002.

[32] Haadem, S., Øksendal, B., Proske, F., Maximum principles for jump diffusion processes with infinite horizon, Automatica, 49, 2267–2275, 2013.

[33] Hindy, A., Huang, C. and Kreps, D., On intertemporal preferences in continuous time: The case of certainty, Journal of Mathematical Economics, 21, 401-440, 1992.

[34] Higham, D. J., Mao, X. , Stuart, A. M., Strong convergence of Euler methods for non-linear stochastic differential equations, SIAM Journal of Numerical Analysis, 40, 1041-1063, 2002.

[35] Holden, H., Øksendal, B., Ubøe, J., Zhang, T., *Stochastic Partial Differential Equations: A Modeling, White Noise Functional Approach*, Springer, Berlin Heidelberg, 2010.

[36] Hu, Y., Øksendal, B. and Sulem, A., Singular mean-field control games with applications to optimal harvesting and investment problems, arXiv:1406.1863, 2014.

[37] Hu, Y. and Øksendal, B., Partial information linear quadratic control for jump diffusions, SIAM Journal on Control and Optimization, 47, 1744-1761, 2008.

[38] Iacus, S. M., *Simulation and Inference for Stochastic Differential Equations*, Springer, Berlin Heidelberg, 2008

[39] Jouini, E., Kallal, H., Arbitrage in securities markets with short-sales constraints, Mathematical Finance, 3, 197-232, 1995.

[40] Kabanov, Y., Stricker, C., The Dalang-Morton-Willinger theorem under delayed and restricted information, In Memoriam Paul-André Meyer, Springer, 2006.

[41] Kabanov, Y., Stricker, C., Yu, M., A teacher's note on no arbitrage criteria, Séminaire de Probabilités XXXV, Lecture Notes in Matheamtics, 1755, 149-152, 2001.

[42] Karatzas, I., Kou, S. G., Hedging American contingent claims with constrained portfolios, Finance and Stochastics, 2, 215-258, 1998.

[43] Karatzas, I., Equilibrium in a simplified dynamic, stochastic economy with heterogeneous agents, Department of Mathematical Sciences, Carnegie Mellon University, paper 366, 1989.

[44] Karatzas, I. and Shreve, S., *Brownian Motion and Stochastic Calculus*, second ed., Springer, New York, 1998.

[45] Karatzas, I. and Shreve, S., *Methods of Mathematical Finance*, Springer, New York, 1998.

[46] King, A. J., Duality and martingales: A stochastic programming perspective on contingent claims, Mathematical Programming, 91, Ser. B, 543-562, 2002.

[47] King, A. J., and L. Korf, Martingale pricing measures in incomplete markets via stochastic programming duality in the dual of $\mathscr{L}^\infty$, preprint, University of Washington, 2001.

[48] Klein, I., Rogers, L. C. G., Duality in optimal investment and consumption problems with market frictions, Mathematical Finance, 17, 225-247, 2007.

[49] Koo, H. K., Consumption and portfolio selection with labor income: A continuous time approach, Mathematical Finance, 1, 49-65, 1998.

[50] Kramkov, D., Schachermayer, W., The asymptotic elasticity of utility functions and optimal investment in incomplete markets, Annals of Applied Probability, 9, 904-950, 1999.

[51] Kramkov, D., Schachermayer, W., Necessary and sufficient conditions in the problem of optimal investment in incomplete markets, Annals of Applied Probability, 13, 1504-1516, 2003.

[52] Kunita, H., Stochastic differential equations based on Lévy processes and stochastic flows of diffeomorphisms, Real and stochastic analysis, Trends Math., 305-373. Birkhäuser, Boston MA, 2004.

[53] Lakner, P., Utility maximization with partial information, Stochastic Processes and their Applications, 56, 247-273, 1995.

[54] Ma, J., Yin, H. and Zhang, J., On non-Markovian forward-backward SDEs and backward stochastic PDEs, Stochastic Processes and their Applications, 122, 3980-4004, 2012.

[55] Mao, X., *Stochastic Differential Equations and Applications*, second ed., Horwood Publishing, Chichester UK, 2007.

[56] Xuerong, M and S. Sabanis, Numerical solutions of stochastic differential delay equations under local Lipschitz condition, Journal of Computational and Applied Mathematics, 151, 215 - 227, 2003.

[57] Merton, R. C., Lifetime portfolio selection under uncertainty: The continuous-time case, The Review of Economics and Statistics, 3, 247-257, 1969.

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

[59] Mohammed, S., Stochastic functional differential equations, Volume 99 of Research Notes in Mathematics, Pitman (Advanced Publishing Program), Boston, MA, 1984.

[60] Mohammed, S. Stochastic Differential Systems with Memory: Theory, Examples and Applications, Stochastic Analysis and Related Topics VI, vol. 42 of the series Progress in Probability, 1-77, 1998.

[61] Mohammed, S. and Zhang, T., Anticipating stochastic differential systems with memory, Stochastic Processes and their Applications, 119, 2773-2802, 2009.

[62] Mossin, J., A note on uncertainty and preferences in a temporal context, The American Economic Review, 59, 172-174, 1969.

[63] Nualart, D., *The Malliavin Calculus and Related Topics*. second ed., Springer, Berlin, 2006.

[64] Øksendal, B., *Stochastic Differential Equations*, Springer, Berlin Heidelberg, 2007.

[65] Øksendal, B. and Røse, E. E., Application of white noise theory and Hida-Malliavin calculus in finance, in preparation, 2016.

[66] Øksendal, B. and Sulem, A., *Applied Stochastic Control of Jump Diffusions*, Springer, Berlin Heidelberg, 2007.

[67] Øksendal, B. and Sulem, A., Singular stochastic control and optimal stopping with partial information of Itô-Lévy processes, SIAM Journal of Control and Optimization, 50, 2254-2287, 2012.

[68] Øksendal, B. and Sulem, A., Forward-backward stochastic differential games and stochastic control under model uncertainty, Journal of Optimization Theory and Applications, 161, 22-55, 2014.

[69] Øksendal, B. and Sulem, A., Risk minimization in financial markets modeled by Itô-Lévy processes, Afrika Matematika , DOI 10.1007/s13370-014-0248-9, 2014.

[70] Øksendal, B. and Zhang, T., Backward stochastic differential equations with respect to general filtrations and applications to insider finance, Communications on Stochastic Analysis, 6, 703-722, 2012.

[71] Øksendal, B. and Zhang, T., The general linear stochastic Volterra equation with anticipating coefficients, Proceedings of the Fifth Gregynog Symposium, edited by I. M. Davies, A. Truman and K. D. Elworthy, Stochastic analysis and applications, World scientific, 1996.

[72] Øksendal, B. and Zhang, T., The stochastic Volterra equation, Barcelona seminar on stochastic analysis, edited by D. Nualart and M. S. Solé, Birkhäuser, 168-202, 1993.

[73] Øksendal, B., Sulem, A. and Zhang, T., Optimal control of stochastic delay equations and time-advanced backward stochastic differential equations, Advances in Applied Probability, 43, 572–596, 2011.

[74] Øksendal, B. and Sulem, A. Optimal control of predictive mean-field equations and applications to finance, arXiv:1505.04921, 2015.

[75] Pennanen, T., Dual representation of superhedging costs in illiquid markets, Mathematics and Financial Economics, 5, 233-248, 2012.

[76] Pennanen, T., Introduction to convex optimization in financial markets, Mathematical Programming, 134, 157-186, 2012.

[77] Pennanen, T. and Perkkiö, A. P., Stochastic programs without duality gaps, Mathematical Programming, 136, 91-110, 2012.

[78] Pennanen, T., Convex duality in stochastic optimization and mathematical finance, Mathematics of Opererations Research, 36, 340-362, 2011.

[79] Pham, H., Continuous-time stochastic control and optimization with financial applications, Springer, Berlin Heidelberg, 2009.

[80] Pulido, S., The fundamental theorem of asset pricing, the hedging problem and maximal claims in financial markets with short selling prohibitions, The Annals of Applied Probability, 2014, Vol. 24, No. 1, 54–75.

[81] Platen, E., Rungaldier, W. J., A benchmark approach to portfolio optimization under partial information, Research Paper Series from Quantitative Finance Research Centre, No 191, University of Technology, Sydney, 2007.

[82] Pliska, S., *Introduction to Mathematical Finance*, first ed., Blackwell Publishing, Malden, 1997.

[83] Prévôt, C. I. and Röckner, M., A concise course on stochastic partial differential equations, Lecture Notes in Mathematics 1905, Springer-Verlag, Berlin Heidelberg, 2007.

[84] Protter, P., *Stochastic Integration and Differential Equations* Springer, Berlin Heidelberg, 2003.

[85] Rockafellar, R. T., *Conjugate Duality and Optimization*, first ed., Society for Industrial and Applied Mathematics, Philadelphia, 1974.

[86] Rockafellar, R. T., *Convex Analysis*, Princeton University Press, Princeton, 1972.

[87] Rockafellar, R. T. and Wets, R. J.-B., *Variational Analysis*, Springer, Berlin Heidelberg, 2004.

[88] Rogers, L. C. G., Duality in constrained optimal investment and consumption problems: A synthesis, Lecture Notes in Mathematics, Springer, Berlin, 2003.

[89] Rogers, L. C. G., Dual valuation and hedging of Bermudan options, SIAM Journal of Financial Mathematics, 1, 604-608, 2010.

[90] Romer, D., *Advanced Macroeconomics*, third ed., McGraw-Hill/Irwing, New York, 2006.

[91] Sanz-Solè, M., *Malliavin Calculus*, EPFL Press, Lausanne, 2005.

[92] Shapiro, A., Dentcheva, D., Ruszczyński, A., Lectures on stochastic programming: Modeling and theory, SIAM and MPS, Philadelphia, 2009.

[93] Shilling, R. L., *Measures, Integrals and Martingales*, Cambridge University Press, Cambridge, 2007.

[94] Sydsæter, K., Seierstad, A. and Strøm, A., *Matematisk Analyse: Bind 2*, fourth ed., Gyldendal Akademisk, Oslo, 2012.

[95] Tang, S. and Li, X., Necessary conditions for optimal control of stochastic systems with random jumps, SIAM Journal of Control and Optimization, 32, 1447-14575, 1994.

[96] Totoki, H., A method of construction of measures on function spaces and its applications to stochastic processes, Mem. Fac. Sci. Kyushu Univ., Ser. A, 15, 1961.

[97] Vanderbei, R. J., *Linear Programming: Foundations and Extensions*, third ed., Springer, Berlin, 2008.

[98] Vanderbei, R. J., Ç- Pinar, M., Pricing American perpetual warrants by linear programming, SIAM Rev., 4, 767–782, 2009.

[99] Veraar, M., The stochastic Fubini theorem revisited, Stochastics An International Journal of Probability and Stochastic Processes, 84, 543-551, 2012.

[100] Wang, B., Singular control of stochastic linear systems with recursive utility, Systems and Control Letters, 51, 105-122, 2004.

[101] Wu, Z., Maximum principle for optimal control problem of fully coupled forward-backward stochastic systems, Systems science and Mathematical Sciences, 11, 249-259, 1998.

[102] Zeldes, S. P., Optimal consumption with stochastic income: Deviations from certainty equivalence, The Quarterly Journal of Economics, 275-298, 1989.