INAUGURAL - DISSERTATION

zur Erlangung der Doktorwürde
an der
Naturwissenschaftlichen-Mathematischen Gesamtfakultät
der
Rubrecht-Karls Universität Heidelberg

Vorgelegt von
Stephan Ernst Ludwig
Diplom Mathematiker
Geb. am 29.06.1981 in Heidelberg

Tag der mündlichen Prüfung:

Optimal portfolio allocation of commodity related assets using a controlled forward-backward stochastic algorithm

Advisors:
Prof. Dr. Dr. h.c. mult. Willi Jäger
Prof. George C. Papanicolaou

Abstract

In the first part of this thesis we develop an investment-consumption model with convex transaction costs and optional stochastic returns for a finite time horizon. The model is a simplified approach for the investment in a portfolio of commodity related assets like real options or production facilities. In contrast to common models like [Awerbuch, Burger 2003] our model is a multi time step approach that optimizes the investment strategy rather then calculating a static imaginary optimal portfolio. On one hand, our numerical results are consistent with the well-known investment-consumption theory in the literature. On the other hand, this is the first in-depth numerical study of a case with convex transaction costs and optional returns. Our focus in the analyses is the form of the investment strategy and its behavior with respect to model parameters.

In the second part, an algorithm for solving continuous-time stochastic optimal control problems is presented. The numerical scheme is based on the Stochastic Maximum Principle (SMP) as an alternative to the widely studied dynamic programming principle (DPP). By using the SMP, [Peng 1990] obtained a system of coupled forward-backward stochastic differential equations (FBSDE) with an external optimality condition. We extend the numerical scheme of [Delarue, Menozzi 2005] by a Newton-Raphson method to solve the FBSDE system and the optimality condition simultaneously. This is the first fully implemented algorithm for the solution of stochastic optimal control problems through the solution of the corresponding extended FBSDE system. We show that the key to its success and numerical advantage is the fact that it tracks the gradient of the value function and an adjusted Hessian backwards in time. The additional information is then exploited for the optimization.

Zusammenfassung

Im ersten Teil dieser Arbeit wird ein Investition-und-Konsum Modell mit konvexen Transaktionskosten und optionalem stochastischen Ertrag entwickelt. Der Betrachtungshorizont ist endlich. Dies ist ein vereinfachter Ansatz für die Modellierung von Investitionen in ein Portfolio von rohstoffverarbeitenden Industrieanlagen oder Realen Optionen. Im Unterschied zu bisherigen Ansätzen wie in [Awerbuch, Burger 2003] benutzt das Modell mehrere Zeitschritte. Dadurch berechnet es kein statisches und imaginäres, optimales Portfolio, sondern es optimiert die Investitionsstrategie über den gesamten Zeithorizont. Die Ergebnisse sind einerseits konsistent mit der Theorie über bekannte Investition-und-Konsum Modelle in der Literatur. Andererseits ist unsere Studie die erste grundsätzliche Untersuchung der Fälle von konvexen Transaktionskosten und optionalem Ertrag. Unser Hauptaugenmerk ist das Verhalten der Investitionsstrategie mit Blick auf die Wahl von Modellparametern.

Im zweiten Teil dieser Arbeit wird ein Algorithmus zur Lösung von stochastischen Kontrollproblemen in kontinuierlicher Zeit vorgestellt. Das numerische Schema basiert auf dem stochastischen Maximums Prinzip (SMP) als eine Alternative zu der weit erforschten Dynamischen Programmierung (DPP). Bei der Studie des SMP fand [Peng 1990] ein System von stochastischen, gekoppelten vorwärts-rückwärts Differentialgleichungen (FBSDE) mit externer Optimalitätsbedingung. In der vorliegenden Arbeit wird das numerische Schema von [Delarue, Menozzi 2005] um eine Newton-Raphson Methode erweitert, um die FBSDE und die Optimierung simultan zu lösen. Dies ist der erste, vollständig implementierbare Algorithmus für die Lösung von stochastischen Kontrollproblemen durch die Lösung des entsprechenden gekoppelten FBSDE Systems. Der Schlüssel für den Erfolg und die Vorteile des Algorithmus liegen darin begründet, dass der Algorithmus den Gradienten und eine adjustierte Hessian bei der Rückwärtsinduktion berechnet. Diese zusätzliche Information wird dann zur Lösung des Optimierungsproblems ausgenutzt.

Acknowledgment

I use this opportunity to express my profound thankfulness to many people that supported me during my PhD program at the University of Heidelberg.

First of all I want to thank Prof. Willi Jäger for his upfront trust in my work and my person; and also for his financial support. The meetings with him where highly efficient and he was always able to find the major issues of my current problems in a moment.

Equivalently, I want to thank Prof. George Papanicolaou for his lasting continuously support and motivation; and for his generous financial support during my two exchanges at Stanford University. His famous idiom "keep going" kept me going even during uncertain times.

My mentor Prof. Sebastian Sager helped me in the beginning phase of the project and inspired me to participate in an international conference. Prof. Hans-Georg Bock redeemed formal difficulties whenever they occured. I am thankful for their support and interest in my work.

I had a wonderful time with my office mates including political, social, economic and (even) mathematical discussions. Thanks to Lilian Kramer for checking my thesis for mistakes and to Sabrina Kellner for keep reminding me about my missed appointments.

I am also grateful that the Heidelberg Graduate School MathComp supported me financially in connection with my three journeys. Dr. Michael Winkler and Sarah Steinbach are only two names of the HGS team that helped me with petitions while keeping paper-work simple.

Last but not least, I would like to thank my whole family for their personal support. In particular, I feel very thankful to my Dad for his additional financial support and to my Mom for her everlasting care.

Contents

0 Introduction

From a financial point of view, every machine, production facility and commercial service can be seen as a production asset that is owned by a company. The asset refines raw commodities to enhanced products or it is a service that turns labor and materials into commercial products. In general, the production asset transforms inputs into outputs and by doing so, it produces a return. In this thesis:

1. We apply financial portfolio theory to these commodity related assets that are held by a single company. The goal is to maximize the companies's risk averse utility through an optimal investment strategy over a given future time horizon.

2. We also investigate a numerical method that is based on the Stochastic Maximum Principle in order to approximate solutions of the resulting stochastic optimal control problem.

0.1 History and literature

Driven by the deregulation of commodity and energy markets all over the world in recent years, financial theory has become an important tool for commodity processing companies to optimize their business. Complex commodity derivatives replicate production scheduling to some extend in order to give the commodity processing companies more opportunities to hedge their risks. More recently, companies start to assess the value of their physical facilities (production assets) from a financial market point of view. Due to the assets's flexibility (optional production) the valuation method is called real option valuation. The next obvious step is to consider the interaction of several production assets in a company's portfolio in order to optimize future investment.

Portfolio optimization for production assets has been studied in literature recently. In most cases, a standard portfolio model from the stock market is applied to production assets on a one-to-one basis. For example, [Awerbuch, Burger 2003], [Awerbuch, Stirling, Jansen, Beurskens 2005] or [Kleindorfer, Li 2004] discuss the optimal power plant portfolio in industrialized countries using modern portfolio theory. On one hand, each paper considers properties of production assets and their physical nature in a different way. On the other hand, all approaches are single time step (deterministic) optimizations and the result is then an imaginary, static, optimal portfolio.

This means that a company knows the optimal portfolio but there is no indication about the speed of adjustment or the investment process.

In contrast to the single time step approach, the consideration of the full investment strategy over a certain time horizon leads to a multi-dimensional stochastic optimal control problem. Common numerical approaches - like the Dynamic Programming Principle (DPP), see [Kushner, Dupuis 2001] - are expensive in higher dimension. Recently, the Stochastic Maximum Principle (SMP) got attention in the stochastic optimal control community. It was first formulated by [Peng 1990] and it leads to a system of forward-backward stochastic differential equations (FBSDE) coupled through an optimality condition. Also recently, numerical methods became available that find approximative solutions of FBSDE systems. The first fully implementable algorithm is the stochastic forward-backward algorithm from [Delarue, Menozzi 2005]. It uses an explicit backward iteration. A different approach is proposed in [Bender, Zhang 2008] which uses a global (Picard) iteration.

0.2 Overview and main results

This thesis consists of four parts. In **Part I** we present the preliminaries which lay the foundation of our portfolio approach of commodity related assets. The preliminaries are a synthesis of fundamental concepts in mathematical finance and carefully synthesized from diverse literature. They include modeling approaches for general commodities in Section 1, energy commodities in Section 2 and portfolio theory in Section 3.

In **Part II** we derive and analyze our model for optimal portfolio allocation of commodity related assets. First, we develop a discrete-time portfolio model that considers all common aspects of real options in Section 4. This includes, in particular, operation flexibility and constraints. In contrast to the approaches in the literature, our model considers the today's actual portfolio and optimizes the future allocation process over time. This yields better practical results than a static portfolio. Additionally, we account for market illiquidity through a convex (transaction) cost term. The interactions along the production chain, competition and supply & demand are not modeled explicitly. These aspects are rather encapsulated in the commodity price dynamics. Our general discrete-time model, serves as a template that can be utilized by companies from various industry sectors to estimate their portfolio value of production assets and optimize their future investment strategy. It is formulated in Problem 1 (4.20 - 4.29)

Second, we reduce the complex discrete-time model to a continuous-time investment-consumption model in order to analyze the model's dynamics in general in Section 5. The resulting model is formulated in Problem 2 (5.13) - (5.16). It is similar to Merton's investment-consumption model [Merton 1971] with the following three extensions:

- Each asset yields a stochastic return R_t which is a simplified version of the asset's production spread.
- The returns are optimized through $\max(R_t, -M)$ (M denotes the fixed costs) in order to account of the real option nature of the production assets.
- A convex transaction cost must be paid for the allocation of assets due to market illiquidity.

In Section 6 we consider analytical solutions of the extended investment-consumption model for one asset. We perform a comparative analysis of special cases for which there exist results in the literature and derive a new analytical solution for the case of stochastic returns. In the case of stochastic returns, we found that the optimal proportion of wealth invested in assets π^* depends linearly on the return rate R_t:

$$\pi^*(R_t) = \frac{R_t + \tilde{\mu} - r}{\sigma^2}, \tag{0.1}$$

where $\tilde{\mu}$ is the adjusted asset price's drift and σ is the asset price's volatility. So, in contrast to Merton's original model without returns, the optimal proportion π^* is not constant but rather stochastic. The case of quadratic transaction costs is much more challenging because the value function loses its homogeneity (6.5) and therefore the model cannot be simplified by a reduction of variables. The problem must be analyzed by numerical simulations.

In Section 7, we present numerical results of our reduced continuous-time model (Problem 2) in the case of one asset. In particular, the case of convex transaction costs has not been studied in the literature. We found that:

- The optimal consumption χ_t^* (defined in Section 3.1.5) is reduced in comparison to Merton's original case. We obtain that the reduction is almost proportional to the quadratic transaction costs.
- The optimal allocation rate α_t^* is continuous and directs towards the optimal portfolio. In particular, it is not a bang-bang control as in the case of proportional transaction costs.

- The optimal portfolio consists of fewer assets as in Merton's original model. The reason may be the asymmetry in the model: all consumption takes place from the bank account. When an assets is sold, the portfolio owner must pay transaction costs to consume the income.
- A higher return volatility σ_R increases the value function v and the amount of assets held in the optimal portfolio. The reason is the intrinsic option $\max(R_t, -M)$. The positive effect of higher volatility is reduced for assets with higher fixed costs M.
- The model is consistent with the analytical solutions of our comparative analysis.

In Section 8 we consider a portfolio of two production assets in order to show the applicability of our continuous-time model to multiple asset. Our controlled forward-backward (CFB) algorithm from Part III provides us the ability (in terms of performance) to study this multi-dimensional convex stochastic optimal control problem numerically. A detailed study of a certain business case is left to further research.

In **Part III** we investigate a new numerical method to solve stochastic optimal control problems that have the following two characteristic properties. These type of problems typically appear in mathematical finance and economics. A general formulation is given in Section 10:

- The stochastic state variable X_t is controlled, i.e., the SDE that describes the state's dynamic depends on the control π_t.
- The gain function $f(t,x)$ is concave w.r.t the state variable x.

The general stochastic algorithm is based on the Stochastic Maximum Principle (SMP) which we introduce in Section 11 and leads to a system of forward-backward stochastic differential equations (FBSDE) coupled through an external optimality condition. We present the connections between the dynamic programming and the maximum principle for the deterministic and stochastic case in an comparative manner in Table 9. Additionally, we provide a link between the SMP and HJB theory by showing that the adjoint variables Y_t and Z_t of the backward equation are related to the value function $v(t,x)$ that satisfies the HJB equation:

$$\big(\, Y_t(X_t),\; Z_t(X_t) \,\big) \;=\; \Big(\, \partial_x v(t,X_t),\; \sigma(t,X_t,\pi_t)\, \partial^2_{xx} v(t,X_t) \,\Big) \tag{0.2}$$

In Section 12 we synthesize the SMP together with a numerical scheme for FBSDE systems in order to obtain an implementable algorithm (12.36) -

(12.41). In particular, we extend [Delarue, Menozzi 2005]'s forward-backward stochastic algorithm by a Newton-Raphson method in order to handle the coupled optimality condition. As far as we are aware, this is the first stochastic approach for the solution of stochastic optimal control problems through the solution of the corresponding extended FBSDE system. Therefore, we call the scheme controlled forward-backward (CFB) algorithm. We transfer the convergence results from the FBSDE case to the CFB algorithm in Section 13.

The significance of our algorithm is that we track approximations of Y_t (the gradient of the value function v) and Z_t (the adjusted Hessian of v, see (0.2)) in the backward iteration instead of tracking v_t only. Therefore we can exploit the (additional) information provided in Y_t and Z_t for the optimization. In particular, due to the special choice of predictors, we can pre-calculate the optimal control π^* in a separate step, before the backward calculations of Y and Z. In Section 14 we show that this method leads to an additive effort (computational cost) for the optimization instead of a multiplicative effort when the optimization and the backward iteration are done simultaneously. The numerical complexity is:

$$L^d + K^r < L^d K^r \tag{0.3}$$

where L^d is the number of operations needed to calculate an expectation and K^r is the number of operations needed to solve the optimization problem. The parameter d denotes the number of sources of randomness and r denotes the dimensions of the control space. This advantage is key to the success of the CFB algorithm as an alternative to dynamic programming. We confirm these theoretical studies with results from an application in Section 15. In particular, we compare the performance between our CFB scheme and the DP scheme for the valuation of swing options with penalties and market feedback. We see clearly the structural advantage of (0.3). We published these results in [Ludwig et al. 2012].

In **Part IV** we present the implementation details for our controlled forward-backward algorithm of Part III applied to our model for optimal portfolio allocation of generation assets in Part II. We show the applicability of the CFB-algorithm to the corresponding stochastic optimal control problem in Section 16 and formulate the complete CFB algorithm for the specific problem in Section 17. We discuss implementation issues in Section 18.

0.3 The outline

In Part I, we introduce all concepts of financial theory that are necessary to understand the portfolio model for commodity related assets. In particular, we define commodities from a financial point of view in distinction to standard stocks in Section 1. We also present common commodity derivatives and the valuation for physical assets there. Section 2 is dedicated to the energy market as an important but complicated commodity market. Section 3 presents basic approaches in portfolio theory.

In Part II, we introduce our model on optimal portfolio allocation of commodity related assets. After discussing briefly the classification of the model we develop its general formulation in detail. This includes the choice of the utility function, the description of the state dynamics, the definition of decision variables and the determination of constraints in Section 4. The mathematical formulation is a high dimensional, non-linear stochastic optimal control problem, which is too complex for a meaningful analysis. Therefore, we perform a model reduction in Section 5. In Section 6 we present analytical solutions of the reduced model for simple cases and in Section 7 numerical solutions for the one-asset case. Additionally, we present numerical results of a multi-asset case from the energy sector in Section 8.

In Part III, we develop our controlled forward-backward algorithm for stochastic optimal control problems that is based on the SMP. In Section 10 we introduce the general problem statement. Then we show the connection of the SMP and the HJB theory in Section 11. We develop our algorithm in Section 12 and show its convergence in Section 13. Section 14 is dedicated to the discussion of the advantages of our new method compared to the Dynamic Programming method and Section 15 presents a performance comparison of an application in order to show that the advantages are obtained in practice.

In Part IV, we describe implementation issues that arose when we applied our controlled forward-backward algorithm to our portfolio model. We show the applicability of the CFB-algorithm to the stochastic optimal control problem in Section 16 and formulate the complete CFB algorithm for the specific problem in Section 17. The implementation issues are presented in Section 18.

In the Appendix A.1 we state a few basic definitions about probability spaces and stochastic processes.

Since this interdisciplinary thesis compose several scientific areas (e.g. mathematical finance, operation research, stochastic control and numerics) we use different notation for different contexts.

Part I
Preliminaries to commodity markets, mechanisms in energy production and portfolio theory

The following is a basic introduction into commodity markets, with a special emphasis on energy markets, and portfolio theory. We present fundamental concepts from diverse parts of the literature, which we synthesize later in the approach to an model in order to show how portfolio theory can be applied to commodity related assets.

1 Commodity markets

In the following section we present an overview of commodity markets and appropriate valuation techniques in contrast to standard stock or bond markets. An extensive introduction into commodity markets and price models can be found in [Geman 2005].

1.1 Definition of commodities

A commodity is a generic, relatively unprocessed, homogeneous good that can be processed and resold. It is interchangeable with other commodities of the same type. Most commodities are used as inputs for the production of goods and services, but often the output of processes can still be classified as commodities. Examples are precious metals, base metals, energy sources, agricultural products, timber, livestock, minerals, plastics, chemicals and refined products like gasoline. In contrast, electronic equipment is not a commodity because the properties differ from product to product depending on the producer. However, standardized electronic semiconductors could be regarded as a commodity. Similar definitions can be found at [the free dictionary].

In contrast to stock or bond markets, commodity prices depend on the specific location, the specific time and the specific quality of the physical good. To make this clearer, two commodities of the same type are different assets, if one of these three key properties - location, time, quality - are not

equal. In order to trade commodities in a liquid market environment, standardized future contracts are available at commodity exchanges. Futures specify the three key properties in the contract agreements, through the place of delivery, the time of delivery and the quality of the delivered good. Examples of contract specifications can be found at the web page of the [CME Group, Webpage 2011].

According to [Geman 2005], commodity market prices are always determined by the intersection of supply & demand. Therefore, let us take a closer look on how various supply & demand patters - at different locations, times and for different quality products - influence market prices.

- **Location**: The price of a commodity depends on its location. More precisely, on the local demand-supply relation. At one location there might be an over-supply while there is a supply shortage at another location. For example crude oil of equal quality may have different prices in Europe, North America, Middle East and Asia at the same time.

  The effect of location is smoothed through an efficient transportation network and open markets. In particular, a price difference between two locations can be regarded as a lower bound of transportation costs for the commodity. The reason is that lower transportation costs would immediately activate a physical shipping into the appropriate direction.

- **Time**: The price of a commodity depends on the time of exchange. This effect can be observed through the differences between spot prices and prices for Futures. [Geman 2005] states that "demand for commodities is generally inelastic to prices, given the indispensable nature of the good." In particular, essential goods like food and energy are indispensable on a daily basis. This explains why commodity prices show periodic patterns when supply & demand oscillate over seasons. Examples for seasonal supply are agricultural commodities like soybeans or corn according to harvest seasons. Also irregular supply fluctuations may arise through good or bad weather conditions. An example for seasonal demand is natural gas. The consumption of natural gas for heating purposes increases in winter and decreases in summer while supply does not depend on seasons.

  The effect of time is smoothed through appropriate storage capacities, whenever the commodity can be stored. Inventories allow the owner

to meet unexpected demand and balance disruptions in supply. One the one hand, the holder must pay the cost of storage. One the other hand, he has a benefit of holding the physical commodity available. [Kaldor 1939] and [Working 1948] define the notion of *convenience yield* as a benefit that "accrues to the owner of the physical commodity but not to the holder of a forward contract." Both effects must be considered when we compare spot and forward prices.

- **Quality**: The price of a commodity depends also on its quality. According to our definition above, a commodity is a homogeneous good or at least essentially uniform across producers. Nevertheless, the quality of a commodity may differ slightly in reality and some commodities are more heterogeneous than others. Exchanges try to standardize commodities by defining grades. To receive a certain grade, the quality attributes of a commodity must lay within defined intervals. Wheat for example has many potential quality attributes, including protein content, hardness, foreign matter, toxins. The grade differentials for wheat Futures that are traded at the Chicago Board of Trade (CBOT) can be found at [CME Group, Wheat Future 2011]. In addition to the quality of supply, in some cases there is a need for different grades on the demand side too. For example cars, ships and airplanes need gasoline, diesel and jet-fuel, respectively.

  In some cases, the quality of a commodity can be changed through a refinement and the price differences can be regarded as a lower bound of refinement costs. This is often the case for energy sources. On the other side, exchanges sometimes mix commodities with different grades and issue a fixed premium for the delivery of low quality products or a discount for high quality products. This is done in order to keep contracts standardized and liquid, especially in agricultural markets.

  In some cases, a commodity can be replaced by an alternative commodity that fulfills the same purpose. To some extend, the alternative can be regarded as a similar commodity with different quality. Examples are a daily switch from oil to gas in a dual-fired power plant or a long-term change from oil to gas driven power plants.

However, a high price correlation can be observed in commodity markets between one physical good at different locations, different delivery times and different quality grades as well as its alternatives, see [Geman 2005].

1.1.1 Standardized financial commodity markets

The specific properties may be one reason why the largest amount of commodities is traded over the counter (OTC). Nevertheless, exchanges installed liquid financial commodity markets in the past where also third parties can get involved. To make liquid trading possible, exchanges must offer standardized contracts, i.e., decrease the number of different instruments and simultaneously increase the traded volume of each instrument. In order to do so, exchanges merge locations into areas, delivery times into delivery periods and quality into quality groups. Moreover, the settlement of these contracts is in cash and not physically. These arrangements establish liquid trading of standard instruments like Futures and options. For example a Light Sweet Crude Oil (WTI) Future at the New York Mercantile Exchange (NYMEX) has a delivery period of one month and the delivery may take place at several locations in the U.S., see [CME Group, WTI Future 2011].

Beside producers and consumers, also haulers, storage offerers and refining companies trade standardized spread options to hedge their risk exposure due to location, time and quality, respectively. The underlying of these options are price spreads like location spreads, calendar spreads and quality spreads. Examples are the crack spread (difference between crude and refined petroleum products) and the spark spread (difference between power and needed fuel). Therefore, spread options have become one of the most traded derivatives at standardized commodity exchanges. We introduce spread options in more detail in section 1.3.2.

1.2 Modeling commodity prices

Stochastic models for commodity prices differ from stock and bond price models. In this subsection, we first explain the fundamental price drivers of commodities that can be observed at standardized markets. Then we discuss different approaches for modeling commodity prices and show examples of spot price models and forward curve models, which are most often used in literature.

1.2.1 Fundamental price drivers

The following four effects are commonly observed at standardized commodity markets over time. We explain how each effect is related to the three key properties - location, time and quality - of Section 1.1 through supply & demand.

- **Price shocks and mean-reversion of spot prices**: On one hand, short term demand for commodities is inelastic to prices. If supply or demand outmatch the limited storage capacity, spot prices decrease or increase to extremes within a short time period. This is called price shock or price spike. On the other hand, average consumption is relatively constant or changing only slightly in the long run. Moreover, supply adjusts to demand patterns and vice versa. This gives reasoning to the observable effect of mean-reversion, where commodity prices fluctuate around some equilibrium price, which is stable or slightly changing over long-term. The equilibrium price can be regarded as the marginal cost of production. Such behavior has appreared a number of times in the literature, see for instance, [Pindyck 2001] for the case of energy and [Geman, Nguyen 2002] for agricultural commodities.

- **Backwardation and contango of forward prices**: Next to short term effects, the outlook of future supply & demand plays an important role in commodity markets too. The market's expectations can be observed through the forward prices (Future prices). In case that forward prices decrease with respect to time-to-maturity, the effect is called *backwardation*. In case that they increase with respect to time-to-maturity, the effect is called *contango*. It is reasonable that backwardation indicate a fear of oversupply while contango indicate a fear of supply shortage. But note that in both cases, the cost of storage and the introduced *convenience yield* must be subtracted from the forward prices first, before analyzing supply & demand expectations.

- **Seasonality**: As mentioned above, soft commodities are subject to seasonal supply fluctuations due to harvest and seed times; in particular because agricultural goods perish over time. Other commodities are subject to seasonal demand fluctuations since they are hardly storable (e.g. natural gas) or not storable at all (e.g. power). Both effects lead to seasonal price patterns and are analyzed for different commodities in more detail in [Geman 2005] chapter 7-11 and [Xu 2004] for example.

- **Inter-sector correlations**: Many commodities are similar to or even replaceable by alternatives. More precisely, commodities are related to each other in case they satisfy the same demand category as food, energy or working material. Therefore, high correlation of prices between alternatives or related commodities are observable at markets. For example [Sieczka, Holyst 2009] provides clear evidence for the existence of strong correlations of commodities within a given sector and even for "inter-sector correlations, depending on the level of similarity".

There is a vast amount of literature about modeling the price dynamics of commodities and how to design models that capture the described price drivers. In the following paragraphs, we present the most common stochastic price models. An extended introduction into modeling commodity price behavior can be found for example in [Burger, Graeber, Schindlmayr 2007] chapter 3 or [Geman 2005] chapter 3.

1.2.2 Different model approaches

Basically, there are two different categories of stochastic models for commodity prices. The first category is called *fundamental models*. They consider fundamental price drivers and market mechanisms in order to derive the price from the intersection of supply & demand, where the uncertain supply & demand curves are stochastic variables. These models usually assume some general equilibrium setting, such as Pareto efficiency. The disadvantage of such an approach is that extensive and complete data from production and consumption must be available. Otherwise market mechanisms and their parameters are hard to identify. Moreover, complex business connections and side effects from production or consumption need to be taken into account like resource mining, transportation, storage, capacities, physical constraints, alternatives and others. This may be one reason why fundamental models are more common in power markets than in other markets, since there is transparency of data. We give an example of a fundamental model for power prices in Section 2.3.

The second category is called *reduced-form models*. These models describe prices or price changes over time as a stochastic process directly. The form of the process attempts to capture the distribution of returns observed in current and historic data. These models are preferred when it comes to pricing derivatives. Once the model parameters are calibrated to market data of plain vanilla instruments, one can obtain prices (often as closed from solutions) for derivatives. This method is called *mark-to-market* pricing. In the following paragraphs we present examples of reduced-form models that are commonly used in the literature, or that I observed in practice at different financial institutions in Europe during my work as risk-management consultant at d-fine GmbH[1]. In particular, we show how these reduced-form models capture the fundamental price drivers described in Section 1.2.1.

A thrid less common category is the one of game theoretical models. Exam-

[1] www.d-fine.com

ples can be found in [Kannan, Zavala 2010].

1.2.3 Spot price models

Let S_t denote the spot price of a certain commodity at time t. In general, S_t is assumed to be a stochastic process on a well-defined probability space (Ω, F, P). By well-defined we mean that the probability space can be augmented with a filtration $\mathbb{F}_t = \{\mathrm{F}_t, t > 0\}$ that is naturally generated by all uncertain factors of the specific price model. Loosely speaking, F_t contains all information available until time t. A brief introduction into probability spaces, stochastic processes and related definitions that are used in the thesis can be found in the Appendix A.1.

In the following, W_t always denotes a Wiener process in an appropriate dimension that is compatible to the filtration $\mathbb{F}_t$.

Vasicek model

[Vasicek 1977] introduced the first continuous time model that represents the random evolution of short-term interest rates using a so-called Ornstein-Uhlenbeck process, see [Ornstein, Uhlenbeck 1930]:

$$dS_t = \kappa\left(\mu - S_t\right)dt + \sigma dW_t, \tag{1.1}$$

where κ, μ, σ are positive constants. If $S_t > \mu$, the expected change is negative. If $S_t < \mu$, the expected change is positive. In general, the process is always reverting towards the mean level μ where the parameter κ controls the relative speed of reversion. Therefore, the model is very appropriate to describe the mean-reverting effect of commodity spot prices and it is frequently used in literature.

Schwartz one factor model

To preclude negative values [Schwartz 1997] applied the Ornstein-Uhlenbeck process to the log price, $\ln S_t$, and obtained the following SDE:

$$dS_t = \kappa\left(\mu - \ln(S_t)\right)S_t dt + \sigma S_t dW_t. \tag{1.2}$$

A special feature of the Schwartz model is that at time t, $\ln S_T$ is normally distributed for all $T > t$. Moreover, a closed form solution of forward prices $F_t(T)$ with time-to-maturity T is available:

$$F_t(T) = E\left[S_T \middle| \mathbb{F}_t\right] = exp\left[e^{-\kappa T} lnS_t + \left(1 - e^{-\kappa T}\right)\alpha + \frac{\sigma^2}{4\kappa}\left(1 - e^{-2\kappa T}\right)\right], \tag{1.3}$$

where $E[\cdot|\,\mathbb{F}_t]$ denotes the conditional expectation and α is a derived parameter. One draw-back of Schwartz's model is that the volatility of $F_t(T)$ is decreasing with time-to-maturity:

$$\sigma^{F_t(T)} = \sigma e^{-\kappa(T-t)}. \tag{1.4}$$

A decay of volatility is not observable in all commodity markets and the user may struggle to capture the volatility of contracts with long maturities. Another draw-back is that the mean-reverting level is fixed while in reality the level may change in the long run.

Schwartz-Smith two factor model

[Schwartz, Smith 2000] proposed a two factor model combining mean-reverting behavior on a short time scale (variations) with a non-reverting behavior on a long time scale (dynamics):

$$\begin{aligned} \ln S_t &= \chi_t + \xi_t, \\ d\chi_t &= -\kappa\chi_t dt + \sigma_\chi dW_t^\chi, \\ d\xi_t &= \mu\xi dt + \sigma_\xi dW_t^\xi, \\ dW_t^\chi dW_t^\xi &= \rho dt, \end{aligned} \tag{1.5}$$

where κ, σ_χ, μ, σ_ξ are positive constants and $\rho \in [0,1]$ defines the correlation between dW_t^χ and dW_t^ξ. [Schwartz, Smith 2000] showed that the model is equal to the Schwartz one factor model with a "stochastic mean-reversion level". The spot price S_t itself is log-normally distributed and closed-form solutions for forward and option prices can be derived. The model gives more flexibility to the possible shapes of forward curves and time-dependent volatilities. Nevertheless, the model is not sufficient to match an arbitrary shape of volatility curves, which are observable in some commodity markets.

Stochastic volatility models

[Trolle, Schwartz 2009] observed that "volatility is largely unspanned by the Futures contracts" and therefore they introduced a stochastic volatility model to be able to fully capture market prices of options. [Eydeland, Geman 1998] extend this model to commodities such that it also includes the effect of

mean-reversion:

$$\begin{aligned} dS_t &= \kappa^1 \left(\mu^1 - \ln(S_t)\right) S_t dt + \sigma_t S_t dW_t^1, \\ d\Sigma_t &= \kappa^2 \left(\mu^2 - \Sigma_t\right) + e\sqrt{\Sigma_t} dW_t^2, \\ dW_t^1 dW_t^2 &= \rho dt, \end{aligned} \tag{1.6}$$

where $\Sigma_t = \sigma_t^2$ and κ^1, κ^2, μ^1, μ^2, e are positive constants. The correlation coefficient ρ is usually negative since, in contrast to stock prices, "the volatility of commodity prices tend to decrease with prices - the inverse leverage effect" as discussed in [Geman 2005] chapter 5.

Adding jumps

[Merton 1976] introduced a jump-diffusion model to describe the observations of violent movements of stock markets as a consequence of the arrival of good or bad news. Jumps are observed in commodity markets also for other reasons, see [Geman 2005]. The dynamics of the most simple jump-diffusion model are:

$$dS_t = \mu S_t + \sigma S_t dW_t + U_t S_t dN_t, \tag{1.7}$$

where N_t denotes a Poisson process with intensity λ, U_t is a normally distributed random variable and μ, σ are positive constants. All previously introduced models can be extended by jumps which leads for example to a mean-reverting jump-diffusion model. One draw-back of adding jumps is that the spot price S_t is not normally distributed anymore. On the contrary, jumps are often incorporated in models to account for fat tails in the price distributions.

Adding seasonality

To account for seasonality, many authors simply add an extra term to their model. For example [Sorensen 2002] introduced the following general model for agricultural commodities:

$$\begin{aligned} \ln S_t &= s(t) + \chi_t + \xi_t, \\ s(t) &= \sum_{k=1}^{K} \gamma_k \cos(2\pi kt) + \gamma_k^* \sin(2\pi kt), \end{aligned} \tag{1.8}$$

where χ_t, ξ_t are arbitrary stochastic processes and γ_k, γ_k^*, for $k = 1, ..., K$, are constant parameters to be estimated.

Regime switching models

Many economic time series occasionally exhibit dramatic breaks in their behavior that are associated with events such as financial crises, see for example [Jeanne, Masson 2000]. A mixed regime model is able to capture different behavior of economic variables during these periods. It appears that these types of models have been analyzed first by [Lindgren 1978].

Except for the occurance of dramatic economic changes, high price periods of short duration (spikes) are observable in commodity markets and regime switching models were introduced by many researchers to fit the shape of these spikes.[2] The simplest form of a regime switching model has two possible states, $R_t \in \{1, 2\}$, as presented in [Hamilton 1989], where the log-price is given by:

$$\ln S_t = \begin{cases} X_t^1 & \text{if } R_t = 1, \\ X_t^2 & \text{if } R_t = 2, \end{cases} \tag{1.9}$$

and a transition matrix is given by:

$$P = \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix} = \begin{pmatrix} p_{11} & 1 - p_{11} \\ 1 - p_{22} & p_{22} \end{pmatrix}, \tag{1.10}$$

where X_t^1, X_t^2 are stochastic state variables. The p_{ij}'s denote the probability of change from state i to state j.

There is extensive literature on other multi-factor models combining mean-reversion, stochastic convenience yields, stochastic volatility, jumps, regime switches, seasonality and seasonal volatilities.

In applications, one has to keep in mind that the more features a model captures, the more parameters are needed. The problem is that the parameters need to be estimated from often sparse market data. I gained experience with parameter estimation for spot price models from the analysis of more than 45 different commodities at a large German bank. I saw that on the one hand, almost every commodity needs its own adapted pricing model to adequately price traded contracts at commodity exchanges. On the other

[2] Spikes are hard to capture with jump diffusion models, since the duration of high price regimes is short compared to the duration of "normal" price regimes. For example [Weron, Bierbrauer, Trueck 2004] compared a jump-diffusion and a regime switching model for the spot electricity prices from the Nordic power exchange and found that the regime switching model leads to better estimates and better modeling performance.

hand, models are neither robust nor applicable in practice whenever they have more parameters than there are liquid (daily observable) contracts in the market.

1.2.4 Forward curve models

Usually a commodity is not traded at spot markets at all but only through forward contracts (e.g. standardized Futures). In particular, when a physical good needs time to be exchanged, a spot price makes no sense. In most cases, a commodity market consists of several tradeable forward contracts with different maturities. Therefore, a number of authors proposed to model the dynamics of the whole forward curve $\{F_t(T)|\ T \geq t\}$ directly, instead of modeling the spot price S_t.

[Clewlow, Strickland 1999] argued that the volatility and covariance structure of commodity markets are similar to interest rate markets and adapted the moel of [Heath, Jarrow, Morton 1992] (HJM model) to commodities. The HJM model describes the dynamics of the forward prices $F_t(T)$ by:

$$dF_t(T) = \mu_t(T)F_t(T)dt + \sum_{i=1}^{n} \sigma_t^i(T)F_t(T)dW_t^i, \tag{1.11}$$

where n is the number of risk factors and the Wiener processes W_t^i, for $i = 1, ..., n$, are possibly correlated. The key of the HJM model is that, no drift estimation is needed. Indeed, assuming an arbitrage free market, there exists a risk-neutral probability measure Q that makes the forward price $F_t(T) = E^Q[S_T|\ \mathrm{F}_t]$ a Q-martingale. Thus, under a Q-measure, $\mu_t(T)$ vanishes for all t and T.

HJM-type models capture the full dynamics of the entire forward rate curve, while short-rate models only capture the dynamics of a point on the curve. In particular, in equation 1.11, the drift and the volatility depend on two time variables t and T, in contrast to spot price models where there is only one time dependency. Nevertheless, a theoretical spot price can be obtained by $S_t = F_t(t)$ and its dynamics are given by:

$$dS_t = \left.\frac{\partial F_t(T)}{\partial T}\right|_{T=t} dt + dF_t(T)|_{T=t}. \tag{1.12}$$

S_t is log-normally distributed but non-Markovian since the drift part depends on the realized path of the Brownian motion.

One feature of the direct modeling of forward curves is that all option price formulas, derived from the [Black 1976] model, can be used to price options on forward contracts. Since commodities are always traded in forward contracts and therefore all options are written on these underlying forward contracts (e.g. standardized options are written on standardized Futures), this feature is often exploited for pricing commodity options.

When choosing the HJM framework to model a commodity forward curve, one may ask how many parameters such a model needs in reality. [Geman 2005] chapter 3.7 states that $n = 3$ factors "explain 97% of yield curve moves historically observed in various counties". The factors can be identified through a Principal Component Analysis (PCA) of the forward curve. Therefore, three to four parameters should be enough in most cases.

1.3 Commodity derivatives

In stock and bond markets, exotic[3] derivatives often attract speculators by providing individual leverage effects for them. In contrast, exotic derivatives on commodities often mirror the physical processing of goods and therefore provide producers and trading companies an opportunity to hedge their risk exposure due to their commodity holdings.

In the following paragraphs, we introduce several commodity derivatives. We start with basic contracts and then increase complexity. This prepares us for the valuation of physical assets (e.g. a complete production facility) in the next Subsection 1.4. Before we start, we briefly describe the risk neutral valuation of exotic derivatives.

1.3.1 Risk neutral valuation of derivatives

Let us assume a market in the Black-Scholes world (BS market) which 1) is arbitrage free, 2) is complete 3) has a constant interest rate r and 4) is perfect in terms of continuously liquid trading without transaction costs. The first fundamental theorem of asset pricing states, that there exists an equivalent martingale measure $Q \sim P$, often called risk neutral measure, such that the

[3]The term exotic means that a financial instrument is non-standard or difficult to evaluate. There is no general definition of the term exotic since even standard instruments become more complex over time. In spite of that, an option is most often called exotic if it has one or more of the following properties: 1) the pay-off at maturity T is path-dependent, 2) there exists more than one exercise possibility (multiple exercise option), 3) the option has more than one underlying (multi-asset option).

net present value of a derivative V_t with payoff $h_T(S_T)$ at time T is equal to the expected discounted cash-flow of the payoff under Q:

$$V_t = E^Q[e^{-rT} h_T(S_T)|\mathrm{F}_t]. \tag{1.13}$$

The second fundamental theorem of asset pricing states, that if and only if the market is complete, the risk measure Q is unique. Assuming a log-normally distributed spot price in these settings, [Black, Scholes 1973] derived the famous Black-Scholes formula for option pricing.

In view of the Black-Scholes formula, price models are often designed in a way that the spot prices are log-normally distributed. This is even the case for commodity spot price models in order to provide analytical solutions for option and forward contracts. The reason is that the model's parameter need to be calibrated to market data of liquid (standard) contracts and analytical solutions make the corresponding parameter estimation feasible in practice. Note, since the underlyings of commodity options are Futures, one has to use [Black 1976]'s formula instead of the original Black-Scholes formula. If there is no analytical solution available, numerical simulations have to be used to approximate the expected value in (1.13).

1.3.2 Exotic derivatives

Due to the physical nature of commodity markets, European type options are not feasible instruments for market participants and therefore are rarely traded. In the following we present the most commonly traded types of exotic options and explain how production companies use them for hedging.

Asian options

The underlying of an Asian type option is the arithmetic average of an index - representing the spot price in most cases - over a historic time horizon. Asian options became famous in currency markets where companies receive daily cash flows of similar amount in foreign currency. [Geman 2005] states that "in the case of oil, the quantity of time elapsed between the day a tanker leaves the production site and reaches its destination explains why oil indexes are arithmetic averages; accordingly, most options on oil are Asian." Other commodity indexes developed similarly.

In detail, an Asian contract defines the time points of measurement $t_i \in [0, T]$, $i = 1, ..., N$ over a certain time horizon. Practitioners call them fixing points.

The pay-off h_T of an Asian type call option at maturity T is then given by:

$$h_T(S_{t_1}, .., S_{t_N}) = \max\left[\left(\frac{1}{N}\sum_{i=1}^{N} S_{t_i}\right) - K, 0\right] \tag{1.14}$$

where S_t is the index price at time t and K is the strike. The pay-off depends on the path of the price process between t_1 and t_N. The valuation of Asian options has been a challenge for a long time since the arithmetic average is not (log-)normally distributed for most price models. [Geman, Yor 1993] together with [Eydeland, Geman 1995] provided an analytical solution method for Asian option prices by 1) finding the Laplace transformation with respect to maturity of the Asian option price and 2) showing how to invert this Laplace transform.

Other path-dependent options commonly traded in commodity markets are barrier or digital options which are used e.g. to hedge against a catastrophic natural event. I know from my experience at a commodity sales desk, that most corporate clients buy barrier and digital options only on an Asian type underlying these days.

Swaps and swaptions

Swaps on underlying indexes are very common in commodity markets. The pay-off of a swap is equal to a strip of forward contracts on the underlying with maturities $T_i \in [0, T]$ $i = 1, ..., N$ and fixed strike K:

$$h_{T_i}(S_{T_i}) = S_{T_i} - K, \quad \forall i = 1..N \tag{1.15}$$

At every T_i the buyer of the swap pays the fixed amount K of cash and receives the underlying spot price S_{T_i}, making the instrument path-dependent. A typical buyer of a swap needs the underlying commodity as an input for his production. He purchases the commodity physically for the price S_{T_i} but gets $S_{T_i} - K$ back if the price is higher. Thus, he pays the fixed price K and is able to establish a constant margin from production. With the rise of swaps, swaptions - options on swaps - became a commonly traded instrument too.

Exchange options

The holder of an exchange option has the right to exchange a given amount of one commodity for another. [Geman 2005] states that "The spread between two quantities is probably the most traded instrument in the world of commodities. Spread options come into play in the valuation of power

plants, oil refineries, storage facilities and transmission lines." Exchange options are mostly used to trade the differences of the three key properties - time, location and quality - introduced in Section 1.1. A typical pay-off of an exchange option has one of the following structures:

$$\begin{aligned}
h_T(S_T^1, S_T^2) &= \max\left(S_T^1 - S_T^2, 0\right), && \text{location spread,}\\
h_T(S_T^1, S_T^2) &= \max\left(S_T^1 - kS_T^2, 0\right), && \text{quality spread,}\\
h_T(S_T^1, S_T^2) &= \max\left((S_T^1 - kS_T^2), K\right), && \text{minimum spread,}\\
h_T(F_T(T^1), F_T(T^2)) &= \max\left(F_T(T^1) - F_T(T^2), 0\right), && \text{calendar spread.}
\end{aligned} \tag{1.16}$$

Here, S_t^1, S_t^2 are spot prices or prices of Asian underlyings at time t. $F_T(T^1)$, $F_T(T^2)$ denote prices of Futures at T with maturity $T^2 > T^1 > T$ and k, K are positive constants. The first pay-off in (1.16) occurs in spread options which are used to hedge the spread between a commodity and another similar commodity at a different location or of different quality. The second pay-off incorporates a weighting (production) factor k and reflects the margin between output S^1 and input S^2. The third pay-off ensures a fixed minimum margin K for the holder. The last pay-off is called calendar spread and is popular for hedging seasonality.

Note that an exchange option is a multi-asset option and the correlation of prices is a key factor for its valuation. Analytical solutions and optimal exercise regions for a wide range of American type exchange options can be found in [Broadie, Detemple 1994].

Swing options

In the energy sector, volume flexibility in contracts is an essential feature to adjust the production to fast changing demand. Swing options provide this volume flexibility and its holder is able to hedge volume risks. For later purposes, we describe them in more detail now.

Definition 1.1 (Swing option)
A Swing option is a multi-exercise option on a given time interval $[0,T]$ *with exercise times* $t_i \in [0,T]$, $i = 1,...,N$ *and fixed strikes* K_{t_i}. *The option agreement leaves the exchanged quantities (volume) open to the option holder but defines volume bounds or penalties. There are lower and upper bounds,* m *and* M, *for each single (local) exercise and also lower and upper bounds,* A *and* B, *for the total (global) volume. On each exercise time* t_i, *the option*

holder has the right to request a quantity q_{t_i} *such that:*

$$\begin{aligned} m \le q_{t_i} \le M, \\ A \le \textstyle\sum_{i=1}^{N} q_{t_i} \le B. \end{aligned} \tag{1.17}$$

The value of a swing option at time t *therefore is:*

$$V_t(S_t, Q_t) = \max_q E_t^q \left[\sum_{i=1}^{N} e^{-r(t_i - t)} (S_{t_i} - K_{t_i})\, q_{t_i} \,\middle|\, S_t,\ Q_t \right], \tag{1.18}$$

where $Q_t := \sum_{t_i \le t} q_{t_i}$ *is the total volume already called and:*

$$q \in \left\{ \{q_{t_i}\}_{i=1}^{N} \,\middle|\, m \le q_{t_i} \le M,\ A \le Q_t \le B \right\}. \tag{1.19}$$

In practice m is the minimal (daily) must run volume and M is the capacity limit. To fix the total amount recalled, issuers can just set $A = B$. Instead of sticking to the fixed bounds m, M, A and B, some issuers introduce (convex) penalties if (1.17) is violated.

Note that at any time $t \in [0, T]$, the contract has two states - the underlying price S_t and the total quantity already taken $Q_t = \sum_{t_i < t} q_{t_i}$ - plus one control variable q_t. Therefore, the pricing of a Swing option turns out to be a stochastic optimal control problem. We price Swing options in Section 15 as a test application for our later introduced controlled forward-backward stochastic algorithm. For more details about the valuation of Swing options see [Bodea 2012].

1.4 Valuation of production assets

The structure of the above derivatives replicate mechanisms of industrial processes where commodities are produced, transported, refined, stored and sold. Going one step further, a production facility or a processing service itself can be analyzed from a financial point of view and a net present value (NPV) of the physical asset, that processes commodities, can be estimated. In the following paragraphs, we briefly introduce how to valuate production facilities as financial assets in commodity markets. We call them production assets.

1.4.1 Discounted cash flow approach

The most common method to valuate production assets is the discounted cash flow (DCF) approach. In DCF, all expected future cash flows are summarized

and discounted to the present. In detail, let c_{t_i}, for $i = 1, ..., N$, denote the (uncertain) cash flows of an asset during time $[t_{i-1}, t_i)$. Here, t_N is the lifetime of the production asset. Then the NPV at t_0 is approximated by:

$$V_{t_0}^{DCF} = E\left[\sum_{i=1}^{N} e^{-r(t_i - t_0)}\, c_{t_i} \,|\, \mathrm{F}_{t_0}\right]. \tag{1.20}$$

The merit of DCF is its simplicity and applicability to a large class of valuation problems. The drawbacks are 1) the difficult estimation of all expected future cash flows given todays information, 2) the assessment of the influential (risk-adjusted) discount factor r and 3) the fact that DCF does not account for the optionality to adjust operation to optimal levels as soon as more information is provided (intrinsic options). In other words, the flexibility to react on the developments of uncertain variables has no value in DCF. Thus, the DCF value $V_{t_0}^{DCF}$ can be seen as a lower bound of the NPV of production assets. More about the DCF method can be found in [Geman 2005].

1.4.2 Strip of options approach

In the previous paragraph we learned that operational adjustments are a unique feature of production assets. In order to account for these intrinsic optionalities, a production asset can be modeled as a strip of spread options (SSO). Let us assume that a facility refines one input commodity to one output commodity. Let S_t^1 denote the price of the output and S_t^2 the price of the input. Let OM_t denote the cost of operation and maintenance, k the production factor and a the facility's capacity. Then, the NPV is approximated through a SSO by:

$$V_{t_0}^{SSO} = E\left[\sum_{i=0}^{N} e^{-r(t_i - t_0)}\, \max(S_{t_i}^1 - kS_{t_i}^2 - OM_{t_i}, 0)\, a \,\middle|\, \mathrm{F}_{t_0}\right]. \tag{1.21}$$

Since the operator of the facility has the ability to stop production whenever it is not profitable, negative net cash flows vanish. The advantage of the SSO approach is that, assuming a complete market, we can use risk-neutral valuation and the risk free interest rate to calculate $V_{t_0}^{SSO}$. A drawback of this approach is that it neglects any physical constraints of the production process and assumes perfect flexibility. Thus, the SSO value $V_{t_0}^{SSO}$ can be seen as an upper bound of the NPV of production assets.

1.4.3 Real option approach

In reality, changes in operations are restricted by physical constraints. For example, a production line can not be restarted immediately after a shut

down decision since each change in operation takes at least some time. In general, the possibility to change depends on the current state x_t of the system. Therefore, an advanced model must include variables that represent this current state. Usually, the current state x_{t_i} depends on the previous state $x_{t_{i-1}}$ and the operator's decision u_{t_i}:

$$x_{t_i} = x_{t_i}(x_{t_{i-1}}, u_{t_i}). \tag{1.22}$$

Let us assume the example facility from Section 1.4.2. Let u_{t_i} denote a ramping decision at t_i and let $x_{t_i} \in [0, a]$ denote the operation level during $[t_i, t_{i+1})$. Then, the NPV of the production asset in state x_0 at time t_0 is given by:

$$V_{t_0}^{RO}(x_0) = \max_{\{u_{t_i}\}_{i=1}^N} E\left[\sum_{i=0}^{N} e^{-r(t_i - t_0)} \left(S_{t_i}^1 - kS_{t_i}^2 - OM_{t_i} \right) x_{t_i} \,\middle|\, \mathrm{F}_{t_0} \right], \tag{1.23}$$

where every decision u_{t_i} have to be *feasible* according to all physical constraints. The following list shows examples of physical constraints for production assets. We present these constraints in more detail for the specific case of fuel fired power plants in Section 2.4.

- Time that is necessary to change the operation level (ramping time).
- Waiting time after a shut down of production before new decisions can be made (minimum down time).
- Waiting time after a start up of production before new decisions can be made (minimum up time).
- Minimal operational level (must-run level).
- Maximal operational level (capacity).

We call the valuation approach (1.23) a *real option* approach. The main difference between the real option approach and a strip of spread options (SSO) approach in Section 1.4.2 is, that *real* decisions are not reversible and affect future states and decisions. Therefore, the maximum operator in (1.23) is outside the expectation operator because we can not anticipate future; compare the SSO approach (1.21).

Note that a Swing option (1.18) can be seen as a replication of a real option (1.23) - when replacing $S_{t_i} - K_{t_i}$ with $S_{t_i}^1 - kS_{t_i}^2 - OM_{t_i}$ - but it only accounts for volume/operation level constraints. Since there are a lot of different

interpretations of the term *real option* in the literature and in practice, we clarify its meaning in our context through the following definition, which is based on [Ronn 2002].

Definition 1.2 (Real option)
A real option is an exotic financial multi-exercise option, where a single exercise decision affects the future state of the option and possible future decisions. The real option approach takes the supremum of the expectation value over all feasible control policies.

A big advantage of the real option approach is that its solution $V_{t_0}^{RO}(x_0)$ also provides an optimal commitment policy $\{u_{t_i}^*\}_{i=1}^N$. Operators can use this optimal decision path as long as no new information arrives. A drawback of the real option approach is that it leads to a (complex) stochastic optimal control problem.

Note, the name *option* does not imply that a real option is a commodity derivative liquidly tradable at exchanges. In contrast, a production assets is a highly illiquid commodity related asset. [Geman 2005] propose to use rather the real probability measure for pricing real options instead of the risk-neutral measure.

In practice, the real option approach is used to value power generation assets, naphtha crackers, software development projects and restructuring projects. A state of the art model for valuing flexible, gas-fired power plants under emission trading can be found in [Ludwig 2010]. Examples for applying the real option approach to general power plants (generation assets) can be found in [Deng, Johnson, Sogomonian 2001] or [Clewlow et al. 2009]. An extended introduction into real options can be found for example in [Ronn 2002].

2 Energy markets

In this section we introduce some unique properties of energy and power markets and show their implications to reduced-form price models. We also present a fundamental price model for power that is based on a deterministic supply curve and inelastic stochastic demand. Last but not least, we discuss power plants as real options and state specific physical constraints.

2.1 Fossil energy sources

Fossil energy sources are scarce and exist as persistent natural reserves only in a few regions of the world. Nevertheless, fossil resources are the primary source of energy in almost every nation and huge amounts of them are transported and stored around the world. Let us take a closer look at the three key properties, location, time and quality, for fossil energy sources:

- **Location**: Transportation of fossil energy sources is expensive because of their sizes and their toxic nature. For example, coal can be transported only by ship or train, oil tankers need special shells to prevent environmental damage, transportation of gas needs pipelines or costly liquefiers, biomass needs trucks because of the widespread planting area and uranium needs secured transport. The transportation from the actual source or mine to the power plants in different nations may take weeks, up to months. To establish power security, there is a need for well organized transportation networks and supply chain management to match the variable energy demand on time.

- **Time**: Fossils are storable but storage capacities are expensive or limited: oil needs silos, natural gas needs big underground caverns, uranium needs secured boxes, and biomass matures. Seasonality comes into play when energy sources are essential for heating or cooling. Especially the natural gas market shows different price levels in summer and winter times due to seasonal demand.

- **Quality**: Even if energy sources from different locations have different levels of quality (energy content), they can be regarded as homogeneous goods. Otherwise, they can be mixed up to an average quality level. Refined products differ in quality. For example, crude oil is split into driving fuel, jet fuel or heating fuel. This is done in so-called crackers, and that is where the name crack-spread comes from. Since the type of output can be adjusted in those crackers, there is a high price correlation among the refined products and the original energy source.

Most authors apply reduced-form models to price energy commodities. The models usually take the above properties into account and are similar to the models of Section 1.2. Particular examples can be found in [Deng 2000] and [Trolle, Schwartz 2009] for general energy commodities and [Xu 2004] for natural gas.

2.2 Power and renewables

Power has become a indispensable good in industrial nations. In the past, power markets were regulated by authorities to ensure supply stability to inelastic consumer demand. Even after the deregulation of power markets in Europe, North America and some Asian countries[4], a large amount of power is still pre-sold to customers at a constant price. This fixed price policy strongly supports the inelasticity of demand.

In consequence, power spot trading has become of major interest for energy companies and also power intensive industries. Simultaneously, there is increasing research in power price models and mathematical analysis of price risks. Before we deal with price models, let us take a closer look at the three key properties, location, time and quality, for power and renewables.

- **Location**: Transmission of power through high voltage power lines is fast but limited to the transmission capacity. To make power available for consumers after transmission, high voltage power must be transformed into low voltage power by power transformators. Therefore, in most countries with open power markets, a few operators control the complex network of high voltage transmission and low voltage transformation to prevent power outages and network damage. Furthermore, power losses occur proportional to the distance of transmission. Thus, whenever a location for a new power plant needs to be chosen, it is always a trade-off between stable service, environmental issues and efficiency.

- **Time**: Power itself is non-storable and supply & demand have to match exactly at every time. Hydro pump storages can be regarded as power storages but their capacity is limited to the number and size of suitable locations. In Germany for example, estimates showed that maximal 6.7

[4]Also Brazil 1998, Argentina 1991, Republic of South Africa 2000, Australia 1994 and New Zeland 1994. Source: The International Association of Engineering Insurers, http://www.imia.com/.

GW can be provided by hydro pump storage[5]. That is around 0.5% of the daily average power demand (1440 GW[6]) in Germany.

Moreover, an increased installation of renewables is amplifying the problem to match supply & demand. The reason is that renewables are not controllable and hardly predictable since they utilize the forces of nature such as wind and solar beams. On the other hand their operation costs are close to zero so that renewables are must-run units (in Germany by law). Renewable power can be regarded as negative inelastic demand which is highly fluctuating according to weather conditions.

The non-storability of power together with an inelastic seasonal demand and weather dependent renewable power input explain the special shape of power price time series observable in open power markets. Typical patterns are spikes and seasonal variations over the course of the day, week and year. Even negative values occur.

- **Quality**: Electricity itself is a homogeneous good. Nevertheless, power from renewables has a better ecological quality than power from fuel fired power plants. Authorities impose payments for emissions from power generation which is similar to a fee for low quality products. So far authorities have raised taxes, defined emission limits and established auction markets for emission allowances. Furthermore, ecologically aware customers ask for special tariffs to receive power from renewable generation even if prices are higher.

There are a lot of reduced-form models for power prices in the literature, see [Burger, Graeber, Schindlmayr 2007] for an overview. The models are similar to those introduced in Section 1.2 and try to respect the properties above. This leads to names like 'seasonal regime-switching mean-reversion model with jumps'.

2.3 A fundamental market model for power

Fundamental market models consider industrial mechanisms, demand patterns and other drivers to estimate (stochastic) demand & supply curves and derive the market price at the intersection of both. A comprehensive

[5]VDI Nachrichten: http://www.vdi-nachrichten.com/artikel/Die-deutsche-Renaissance-der-Pumpspeicher/55356/2.

[6]Bundesnetzagentur: Auswirkungen des Kernkraftwerk-Moratoriums auf die Übertragungsnetze und die Versorgungssicherheit, Aktualisierung 26. Mai 2011.

introduction into fundamental market models can be found for example in [Burger, Graeber, Schindlmayr 2007] section 4, which we follow here.

2.3.1 Merit order curve and equilibrium price estimation

A merit order curve is a cost based description of the fundamental supply curve. It is constructed by plotting the operation costs of power generation for each technology over the cumulative available capacity in increasing order. Let $a^j_{t_i}$, $j = 1..N^J$, denote the total available capacity of technology j and $c^j_{t_i}$ its operation costs during a certain time interval $[t_i, t_{i+1})$. Without loss of generality, let us assume that $k < j$ whenever $c^k_{t_i} \leq c^j_{t_i}$. Then, the merit order curve $moc_{t_i} : \mathbb{R} \to \mathbb{R}$, $supply \mapsto cost$, during $[t_i, t_{i+1})$ is given by:

$$moc_{t_i}(x) = \sum_{j=1}^{N^j} c^j_{t_i} \mathbf{1}_{\left\{\sum_{l=1}^{j-1} a^l_{t_i} < x \leq \sum_{l=1}^{j} a^l_{t_i}\right\}}. \tag{2.1}$$

In general, operation costs and available capacities change over time and hence the shape of the merit order curve. The index t_i indicates this time dependency. Figure 1 provides an example of a merit order curve. Next to the merit order curve, the price inelastic demand 'curve' at time $t \in [t_i, t_{i+1})$ is plotted in Figure 1 by a strait vertical line at the given demand level d_t. The equilibrium spot price p^P_t at time $t \in [t_i, t_{i+1})$ is the intersection of demand and supply:

$$p^P_t = moc_{t_i}(d_t) \quad \forall t \in [t_i, t_{i+1}). \tag{2.2}$$

An example of a fundamental market model with stochastic demand can be found in [Kramer, 2009].

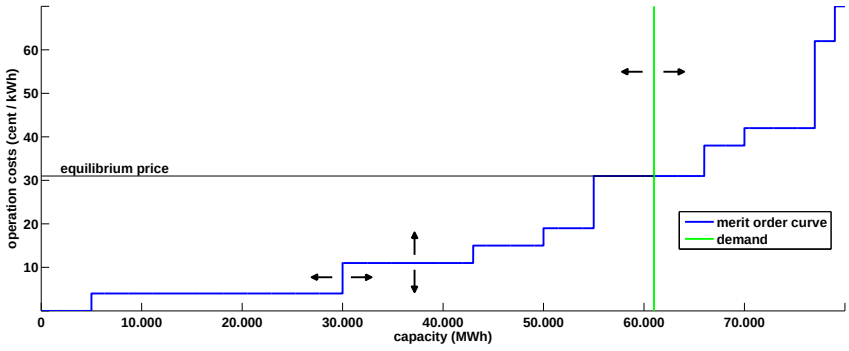

Figure 1: A typical merit order curve in an open power market and an inelastic demand curve. The equilibrium spot price is the intersection of both curves.

A similar approach to determine the market spot price is called *peak load pricing*. In peak load pricing, the lowest bids from the supply side are accepted as long as there is an outstanding demand. The market spot price is fixed then by the lowest bid from the supply side that fulfills the current demand completely. [Coulon, Howison 2009] analyzed the so-called bid-stack curve and proposed a bid stack spot price model.

Note that the differences between the actual operation costs (including a requested profit) on the merit order curve and the real market price from peak load pricing could be interpreted as a risk premium or as scarcity costs, see [Burger, Graeber, Schindlmayr 2007] section 4.3.

2.3.2 Optimal load

The merit order curve does not only provide an equilibrium spot price, but also an optimal choice - optimal in terms of lowest cost - for the amount of generation (load) q_t^j. Power plants of technologies j with $c_{t_i}^j < moc_{t_i}(d_t)$ operate at maximum capacity; plants with $c_{t_i}^j > moc_{t_i}(d_t)$ do not operate at all and plants with $c_{t_i}^j = moc_{t_i}(d_t)$ operate at the level that is required to meet the remaining gap of demand d_t:

$$q_t^j = \begin{cases} a_{t_i}^j & \text{if} \quad c_{t_i}^j < moc_{t_i}(d_t), \\ 0 & \text{if} \quad c_{t_i}^j > moc_{t_i}(d_t), \\ d_t - \sum_j^{N^J} a_{t_i}^j \mathbf{1}_{\left\{\mathbf{c_{t_i}^j < moc_{t_i}(d_t)}\right\}} & \text{if} \quad c_{t_i}^j = moc_{t_i}(d_t). \end{cases} \tag{2.3}$$

In reality, physical constraints need to be taken into account to estimate the optimal load like limited flexibility, limited availability and capacity reserves.

2.3.3 Load and price duration curves

The load duration curve is a re-sorted and aggregated description of the cumulative load $q_t := \sum_{j=1}^{N^J} q_t^j = d_t$ over $[t_i, t_{i+1})$. It can be used whenever the chronological order of load is less relevant. The purpose is to avoid modeling hourly load changes by assuming that the merit order curve stays fixed over a certain time interval (e.g. $[t_i, t_{i+1})$). The load duration curve $ldc_{t_i} : [t_i, t_{i+1}) \to \mathbb{R}$, $t \mapsto q_t$ is constructed out of the load $\{q_t\}_{t \in [t_i, t_{i+1})}$ by sorting the values in descending order such that:

$$\forall t \in [t_i, t_{i+1}) : \quad ldc_{t_i}(t) \geq ldc_{t_i}(T) \quad \forall T \in (t, t_{i+1}) \tag{2.4}$$

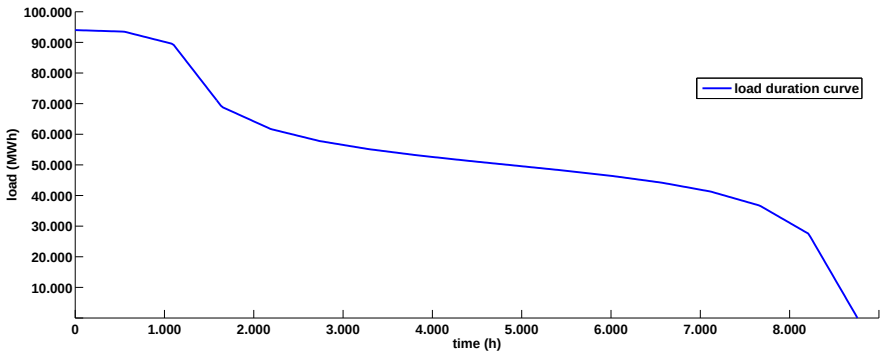

Figure 2: Typical load duration curve for one year

Figure 2 shows a typical load duration curve. The resulting price duration curve $pdc_{t_i} : [t_i, t_{i+1}) \to \mathbb{R},\ t \mapsto p_t^P$ is determined by the equilibrium prices using the fixed merit order curve:

$$pdc_{t_i}(t) = moc_{t_i}(ldc_{t_i}(t)) \quad \forall t \in [t_i, t_{i+1}) \tag{2.5}$$

If the merit order curve is unknown, the price duration curve can be estimated also by an analysis of historic market prices. As mentioned above, fundamental models take the mix of available power plants into account. Therefore, they are especially suitable when it comes to the optimization of power plant portfolios.

2.4 Power plant mechanisms

Power plants have special physical constraints that should be considered in a real option approach. The effect of these constraints depend on the type of facility (e.g. coal, gas, nuclear or biomass) but the constraints themselves are quite similar among all type of fuel fired power plants. The following list provides constraints that are commonly taken into account in a real option valuation of fuel fired power plants according to [Ludwig 2010]. We also show examples of how the constraints could be modeled.

Let us recall the real option approach from Section 1.4.3. Remember that $x_{t_i} \in [a_{min}, a_{max}]$ denotes the operation level during $[t_i, t_{i+1})$ and u_{t_i} denotes the ramping decision at t_i. Additionally, let $y_{t_i} \in \{0, 1\}$ denote the off and on state of the plant during $[t_i, t_{i+1})$ and $\pi_{t_i} \in \{0, 1\}$ the switch-off and switch-on decision, respectively. Then, the NPV of a power plant in state (x_0, y_0) at

time t_0 is given by:

$$V_{t_0}^{RO}(x_0, y_0) = \max_{\{u_{t_i}, \pi_{t_i}\}_{i=1}^{N}} E\left[\sum_{i=1}^{N} e^{-r(t_i - t_0)} g(S_{t_i}, x_{t_i})\, y_{t_i} \,\middle|\, \mathrm{F}_{t_0}\right], \tag{2.6}$$

where g is the gain function from operation and S_t is the price vector of the involved energy commodities. Since we focus on the constraints here, the particular form of the gain function g and the price dynamics dS_t is of minor interest. Possible constraints are:

- Shut-down and Start-up costs: A complete shut-down of a hot power plant and also a start-up from a cold plant produces extra costs. They can be included into the model for example by replacing $g(S_{t_i}, x_{t_i})$ in (2.6) by:

$$\tilde{g}(S_{t_i}, x_{t_i}, y_{t_{i-1}}, \pi_{t_i}) := \left\{ \begin{array}{cc} 0 & \text{if } y_{t_{i-1}} = 0,\ \pi_{t_i} = 0 \\ g(S_{t_i}, x_{t_i}) + c^u & \text{if } y_{t_{i-1}} = 0,\ \pi_{t_i} = 1 \\ c^d & \text{if } y_{t_{i-1}} = 1,\ \pi_{t_i} = 0 \\ g(S_{t_i}, x_{t_i}) & \text{if } y_{t_{i-1}} = 1,\ \pi_{t_i} = 1 \end{array} \right\}, \tag{2.7}$$

  where c^d and c^u are the shut-down and start-up costs, respectively.

- Minimum down time and minimum up time: After a complete shut-down, the plant cannot be restarted again immediately. The plant may need several hours to cool off before it can be restarted. There is an associated (usually shorter) effect after a start-up. Let τ^d and τ^u denote the minimum down time and minimum up time, respectively. Then:

$$\begin{array}{ll} \pi_t = 1 \text{ for } t \in [t_i, t_i + \tau^u] & \text{if } y_{t_{i-1}} = 0,\ \pi_{t_i} = 1, \\ \pi_t = 0 \text{ for } t \in [t_i, t_i + \tau^d] & \text{if } y_{t_{i-1}} = 1,\ \pi_{t_i} = 0. \end{array} \tag{2.8}$$

- Ramp-rates: The speed at which the operation level can be adjusted is called ramp-rate. Let $r^d > 0$ and $r^u > 0$ denote the rate of ramp down and ramp up, respectively. Then for all t_i:

$$\begin{aligned} u_{t_i} &\in [-r^d[t_i, t_{i+1}), r^u[t_i, t_{i+1})], \\ x_{t_{i+1}} &= x_{t_i} + u_{t_i}. \end{aligned} \tag{2.9}$$

  One could also adjust the gain function g to account for the ramping effect.

- Scheduled outages: From time-to-time, a plant needs to be shut down in order to do necessary inspections or repairs. That simply means that $\pi_{t_i} = 0$ is fixed in advance for all t_i in the set of scheduled outage periods:

$$\pi_{t_i} = 0 \text{ when } t_i \in \{\text{scheduled outage periods}\}. \tag{2.10}$$

- Unscheduled outages: Due to operation failures or other circumstances a power plant must be shut down sometimes out of schedule. An example to model unscheduled outages is the following. Suppose that unscheduled outages are exponentially distributed with intensity λ and let $\{\tilde{\pi}_{t_i}\}_{i=0}^{N}$ be a discrete Poisson process. Then the power plant is forced to shut-down for every occurrence of an event:

$$\pi_{t_i} = 0 \text{ when } \tilde{\pi}_{t_i} > \tilde{\pi}_{t_{i-1}}. \tag{2.11}$$

- Minimum operation level: Some power plants must operate at least at a minimum level and cannot operate at lower levels:

$$\text{if } y_{t_i} = 1, \text{ then } x_{t_i} \geq a_{min} > 0. \tag{2.12}$$

- Over-firing: Power plants can operate at higher levels than their normal capacity. But an over-firing event usually requires additional maintenance and repairers. For example, let λ_t denote the time-dependent intensity of the above Poisson process $\tilde{\pi}_{t_i}$ for unscheduled outages. Then $x_t \in [a_{min}, a_{max-over}]$, $a_{max-over} > a_{max}$ and:

$$\lambda_t > \lambda_{t_i} \text{ for } t > t_i, \text{ when } x_{t_i} > a_{max}. \tag{2.13}$$

  Otherwise, we can extend the set {scheduled outage periods} in a particular way whenever $x_{t_i} > a_{max}$.

- Outside temperature: Nuclear power plants, in particular, depend on the outside temperature. The hotter it is, the better is the efficiency. But above a certain maximum temperature T_{max}, the cooling water is too hot and the plant cannot operate at all. For example let T_{t_i} denote the outside temperature at time t_i. T_t is a discrete stochastic process with locally adjusted temperature values and distribution. Then the gain function is dependent on T_t:

$$g = g(S_{t_i}, x_{t_i}, T_{t_i}), \quad \pi_{t_i} = 0 \text{ when } T_{t_i} > T_{max}. \tag{2.14}$$

More details about power plant properties can be found in [Ludwig 2010] or [Burger, Graeber, Schindlmayr 2007].

3 Portfolio theory

In this section we give a brief introduction into portfolio theory. The general aim of portfolio theory is to optimize investment under given utility or risk preferences. We present this framework and show how utility functions describe special risk preferences. Then we introduce the concept of modern portfolio theory and its progression to advanced portfolio theory using different risk measures. We also show how a portfolio model can account for transaction costs and illiquid markets. At the end of the section we have a brief discussion about how the value of market share could be considered in a portfolio model.

3.1 Optimal investment under utility

To clarify the differences between investment and risk-neutral pricing & hedging, let us first recall a basic statement from risk-neutral pricing & hedging of contingent claims.

Definition 3.1 (Perfect market)
A perfect market is frictionless and competitive and there exits continuous trading. A frictionless market is one that has no transaction costs and a competitive market is one where unlimited quantities of each security can be traded without changing the security's price. Trading is continuous when each trade is carried out instantaneously at any time.

Let us assume an arbitrage-free and complete *perfect market*. A contingent claim X is a specified random payoff usually representing wealth. Mathematically speaking, X is a random variable on a probability space (Ω, F, P). In a complete, arbitrage-free market, every contingent claim X can be perfectly replicated (hedged) by a self-financing portfolio of traded assets. Thus, a risk-neutral price of X can be determined by the unique price of the replicating (hedging) portfolio, see Section 1.3.1 for risk-neutral valuation.

3.1.1 Risk preferences

In contrast to pricing & hedging, investment decisions of individual market agents are shaped by preferences over different contingent claims. The investors preferences in algebraic notation $X > Y$ says that claim X is preferred to claim Y. Naturally, the investors preferences are transitive: if $X > Y$ and $Y > Z$ than $X > Z$. A standard example is a risk averse investor who prefers the certain payment of 50 to the 50/50% chance of getting 100 or nothing, even if both claim's expectation is 50. In contrast, a risk-neutral investor would be indifferent.

3.1.2 Utility functions

A basic assumption in portfolio theory is that an investor's preference with respect to wealth can be expressed through a monotone increasing utility function $U : \mathbb{R} \to [-\infty, \infty)$, $X \mapsto U(X)$. So the properties of the utility function must encapsulate the investor's risk preferences.

For example, a risk averse investor has a concave utility. Indeed, for a concave utility function U, Jensen's inequality:

$$E[U(X)] \leq U(E(X)), \tag{3.1}$$

shows that an agent will prefer the certain payment of the claim's expected value to the expectation of the claim's uncertain payment. If U is linear, the agent is risk-neutral and if U is convex, the agent is risk-seeking.

One can prove that the investor's choice stays the same under affine transformations of expected utility function. Therefore [Arrow 1965] and [Pratt 1964] developed two measures of risk aversion that stay constant with respect to these affine transformations:

$$\lambda_1 = -\frac{U''}{U'}, \quad \lambda_2 = -\frac{X\,U''}{U'}. \tag{3.2}$$

λ_1 is the Arrow-Pratt coefficient of absolute risk-aversion (ARA) and λ_2 is the Arrow-Pratt coefficient of relative risk aversion (RRA). A good overview about risk aversion and the corresponding utility functions can be found in [wikipedia, risk aversion].

3.1.3 HARA utility functions

In the literature, the most discussed class of utility functions is Hyperbolic Absolute Risk Aversion (HARA). HARA means that the Arrow-Pratt coefficient of absolute risk aversion λ_1 is positive and hyperbolic in wealth. [Merton 1971] showed that all members of the HARA family can be expressed as:

$$U(X) = \frac{1-\alpha}{\alpha}\left(\frac{\beta X}{1-\alpha} + \nu\right)^{\alpha}, \tag{3.3}$$

with the restrictions:

$$\alpha \neq 1, \quad \beta > 0, \quad \left(\frac{\beta X}{1-\alpha} + \nu\right) > 0, \quad \nu = 1 \text{ if } \alpha = \pm\infty. \tag{3.4}$$

Despite of its name, the class of HARA utility consists of a wide range of utility functions including increasing, decreasing and constant, absolute and relative risk aversion. The most common cases in the literature are:

- Negative exponential utility $U(X) = \frac{1-e^{-\gamma X}}{\gamma}$, with $\gamma > 0$. It has constant absolute risk aversion (CARA) with $\lambda_1 = \gamma$. The parameters are $\beta = \gamma$, $\nu = 1$, $\alpha = \infty$.

- Power utility $\frac{X^\gamma}{\gamma}$, with $\gamma < 1$ and $\gamma \neq 0$. It has constant relative risk aversion (CRRA) with $\lambda_2 = 1 - \gamma$. The parameters are $\beta = (1-\gamma)^{\frac{\gamma-1}{\gamma}}$, $\nu = 0$, $\alpha = \gamma$. CARA utility is sometimes called iso-elastic utility in the literature.

- The logarithmic utility $u(x) = \log(x)$, as the limiting case of power utility when $\gamma \to 0$. It has constant relative risk aversion with $\lambda_2 = 1$.

We will consider these special utility functions in our applications in Part II.

3.1.4 The optimal investment problem

An optimal investment problem under utility is mathematically formulated as follows. Suppose a perfect market of n assets with prices P_t follow an Itô process (continuous-time Markov process):

$$dP_t = \mu(t, P_t)dt + \sigma(t, P_t)dW_t, \quad P_0 = p_0, \tag{3.5}$$

where μ, p_0 are n-dimensional vectors, σ is an $n \times n$ matrix and W_t is a n-dimensional Brownian motion. Furthermore, let $[0, T]$ be the time interval of interest and let $\{\mathrm{F}_t\}_{t\in[0,T]}$ denote the filtration on (Ω, F, P) generated by W_t, which represents the information set at each time $t \in [0, T]$. Let x denote the agent's initial wealth at time 0 and let $a \in A$ denote his investment strategy in the market during $[0, T]$. Here, A is the set of all feasible, F_t-adapted[7] strategies for self-financing[8] portfolios.

By X_T^a we denote the random terminal wealth at time T that is generated using the strategy a. The agent seeks the optimal investment strategy a^* that maximizes his expected utility of terminal wealth. This leads to the following stochastic optimal control problem:

$$\sup_{a\in A} E[U(X_T^a)|\mathrm{F}_0], \tag{3.6}$$

subject to the SDE (3.5). Note that if X_t and a_t are Markov processes, then we can replace F_0 by $X_0 = x$ and $P_0 = p_0$ in the conditional expectation.

[7] For a definition see Appendix A.6.

[8] Self-financing means that there is no other source of income.

3.1.5 The investment-consumption model

In a slightly different approach, an agent additionally derives utility prior to T through consuming wealth. Let χ_t denote the consumption from wealth X_t and let $\delta \geq 0$ denote a discount rate that quantifies impatience. Then the investor's utility is given by:

$$U(X^a) = \int_0^T e^{-\delta t} U_1(\chi_t^a) dt + e^{-\delta T} U_2(X_T^a), \tag{3.7}$$

where U_1, U_2 are utility functions. In the literature, this model is called investment-consumption problem or Merton's problem, see [Merton 1969]. Note that in this approach, the consumption $\{\chi_t, 0 \leq t \leq T\}$ is part of the investment strategy $a \in A$.

Two common hypotheses in perfect markets are, that asset prices P_t are log-normally distributed and that market agents are risk-averse having a strictly concave utility function. Under these assumptions [Merton 1971] showed for the investment-consumption model that "there exists a unique pair of mutual funds constructed from linear combinations of the assets such that, independent of preferences, wealth distribution, or time horizon, individuals will be indifferent between choosing from a linear combination of these two funds or a linear combination of the original n assets." Furthermore, he showed that the prices P_t^* of the mutual funds are log-normally distributed. This is referred to as the *mutual fund theorem* in literature.

3.1.6 A note on utility indifference pricing in incomplete markets

In incomplete markets, the number of uncertainties is, roughly speaking, greater than the number of available underlyings. In particular, there is no risk-free replication for every contingent claim and there exists neither a unique pricing measure nor a unique price. Then, pricing & hedging become similar to investments since the risk preference of an investor is needed to price the non-hedgeable risks. Roughly speaking, the choice of the utility function defines the pricing measure out of all possible risk neutral measures. A good introduction of utility indifference pricing (& hedging) in incomplete markets can be found in [Monoyios 2004] and the recent book [Carmona 2008].

In the literature, investment problems in incomplete markets are commonly analyzed by martingale and duality theory. A general presentation can be found in [Karatzas, Lehoczky, Shreve, Xu 1991].

3.2 Modern portfolio theory

According to the mutual fund theorem (see Section 3.1.5), whenever log-normal distributed prices are assumed, the investment-consumption problem can be reduced to the analysis of two mutual funds. In a complete market, without loss of generality, one of the two funds can be chosen to be the *risk-free* asset, i.e., non-stochastic. This asset is often called *bank account*. Then, the other mutual fund is called the composite risky asset whose prices P_t^* are subject to:

$$dP_t^* = \mu^* P_t^* dt + \sigma^* P_t^* dW_t^*, \tag{3.8}$$

where μ^*, σ^* and dW^* depend only on the distribution parameters of the n assets. So every investor has to deal with only two mutual funds - the risk-free and the composite risky asset- instead of n assets.

[Merton 1971] stated that the proportions of wealth that are invested in the risk-free and the risky asset are derived by "finding the locus of points in the mean - variance space of composite returns which minimize variance for a given mean, and then by finding the point where a line drawn from the risk-free asset is tangent to the locus", see Figure 3. The locus point is often called *market portfolio*. This analysis is analogous to Modern Portfolio Theory (MPT) which was first introduced by [Markowitz 1952].

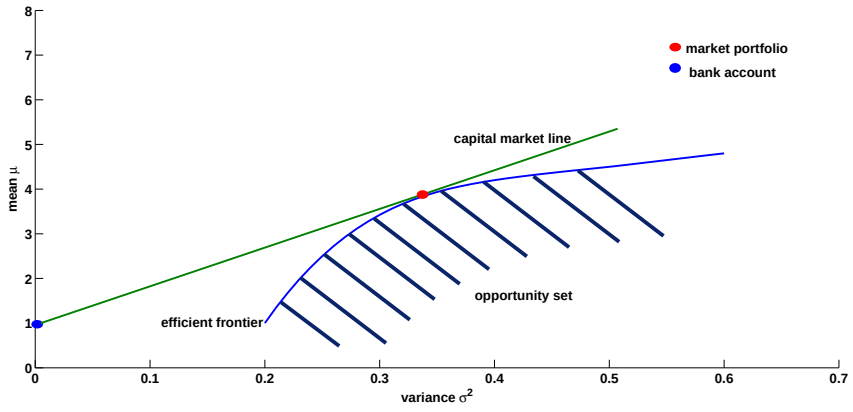

Figure 3: Typical mean-variance diagram of the assets opportunity set with the bank account and the capital market line.

MPT compares the portfolio's means μ and variances σ^2 out of the opportunity set of all possible portfolios that consists of risky assets. The *efficient*

frontier is defined by the portfolios X that maximize the expression:

$$u(X) = u(\mu_X, \sigma_X) = \mu_X - \lambda \sigma_X^2, \tag{3.9}$$

for different λ. Here, λ represents the trade-off between the two conflicting goals of 1) achieving a high mean return μ and 2) minimizing "risk" σ^2. If the market contains a risk-free asset that can be short-sold, the opportunity set is extended by the line through the risk-free asset that is tangent to the efficient frontier in the mean-variance space. Since all optimal portfolios lie on that line, it is called *capital market line* and its slope is the *market price of risk*. The point of contact is the market portfolio, see Figure 3.

MPT provides a neat separation between the capital market opportunities and the agent risk preferences. The former is given by the mean-variance set of possible portfolios and it's efficient frontier or the capital market line whenever there is a risk-free asset in the market. The latter is the choice of risk preference λ.

Transfered to investment under utility in complete markets, the actual optimal portfolio is chosen on the capital market line according to the risk-preferences of the investor. His utility function can be visualized by utility indifference curves (contour lines) in the mean-variance diagram. For concave utility functions the curves are increasing and convex. Therefore, the optimal investment is the point where the indifference curve is tangent to the capital market line, see Figure 4.

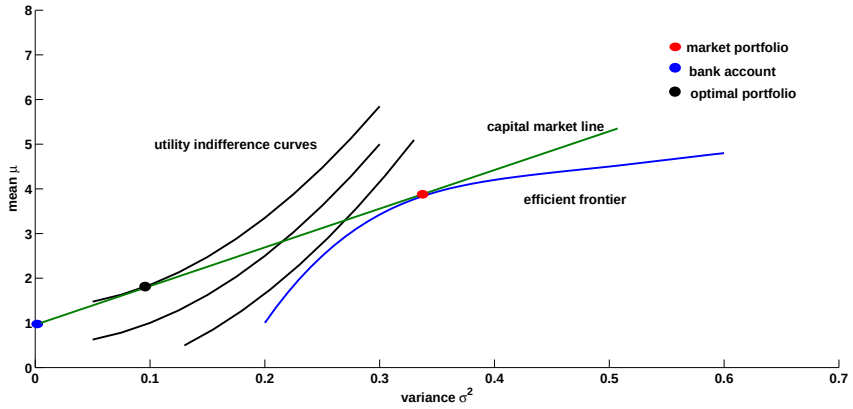

Figure 4: Typical mean-variance diagram of the capital market line with the agent's utility indifference curves and his optimal portfolio.

The concept of a unique market portfolio and a certain market price of risk motivates the well-known capital asset pricing model (CAPM), which was originated by [Treynor 1962] earlier in an unpublished manuscript but published 1999.

In summary, MPT in complete markets has two major aspects:

- A reduction of the portfolio analyis onto two assets, the risky market portfolio and the risk-free asset.
- A separation between market opportunities and the choice of the optimal portfolio according to individual risk preferences.

Therefore, we will concentrate on the two-asset case in our application Part II when we discuss analytical examples in Section 6.

3.3 Advanced portfolio theory

Apart from investment-consumption theory, modern portfolio theory has become of interest for many users in the banking sector - maybe because of its clear separation between the role of the analysts (opportunity set) and the role of traders (risk preferences). A more general framework for MPT arose where the utility function has the following general shape:

$$u(X) = \text{profit}(X) - \lambda\,\text{risk}(X). \tag{3.10}$$

Here, $\text{profit}(X)$ denotes the expected profit of the portfolio X and $\text{risk}(X)$ denotes a measure of risk while λ is a risk-aversion coefficient. To make a clear separation to the original MPT, we call this framework *advanced portfolio theory* here.

Advanced portfolio theory provides more variety to choose the profit function and the risk measure than MPT. Moreover, the portfolio drivers $X \in \mathbb{R}^n$ do not need to be normally distributed. We present several commonly used risk measures after the next paragraph.

Absolute and relative perspectives

If X denotes the absolute wealth of the portfolio, one criticism of the approach (3.10) is that the absolute amount of risk increases with wealth. The parameter λ then denotes absolute risk-aversion. Therefore, especially in a multi-period setting, it is desirable to switch to a relative perspective. Let

$a_t^j = X_t^j / X_t$ denote the proportions of wealth invested in asset j to the composite portfolio X_t. Clearly, $\sum_{j=1}^N a_T^j = 1$ for all t. Then a relative approach is to maximize:

$$u(a) = \text{return}(a) - \lambda \, \text{risk}(a), \tag{3.11}$$

where λ is a relative risk-aversion coefficient.

3.3.1 Risk measures

Now we introduce several risk measures that are commonly used in advanced portfolio theory.

Variance

The classical [Markowitz 1952] portfolio model measures risk through the variance of the returns. Together with the expected mean as return it is also called mean-variance portfolio:

$$\max_{a \in \mathbb{R}^N} \mu^T a - \lambda \, a^T \Sigma^2 a. \tag{3.12}$$

Here, $\mu \in \mathbb{R}^N$ denotes the vector of means of returns and Σ^2 denotes the co-variance matrix of returns. Since Σ^2 is symmetric positive (semi-) definite the utility function is concave and the quadratic optimization problem is well behaved and easy to solve. It is also possible to state an analytical expression for the efficient frontier.

A drawback of the variance approach is, that it is applicable only for normally distributed return structures and does not take into consideration higher moments. Moreover, the measure is symmetric and counts positive outcomes as risk".

Value at Risk

Another risk measure developed by practitioners is the Value at Risk (VaR) that was strongly supported by J.P.Morgan 1994[9]. Let $L(x, w)$ denote the loss of an investment x under the outcome $w \in \Omega$ and let $\Psi(x, \gamma) := P[L(x, w) \leq \gamma]$ be the cumulative distribution function of $L(x, \cdot)$ for a fixed x. Then, for any $\alpha \in [0, 1]$ the VaR on a confidence level α is defined as

$$VaR_\alpha(x) := \min_{\gamma \in \mathbb{R}} \gamma, \tag{3.13}$$

[9] J.P. Morgan launched the RiskMetrics methodology 1992 and outsourced it into the RiskMetrics Group 1998. The Group was acquired by MSCI Inc. 2010.

such that

$$\Psi(x, \gamma) \geq \alpha. \tag{3.14}$$

The merit of the VaR measure is, that it defines a tangible monetary value as risk and that it is a purely down-side risk measure, accounting only for losses. The drawbacks of the VaR are:

- The portfolio optimization becomes a bilevel optimization problem which typically is NP-hard to solve.
- The VaR is often non-smooth and has many local minimizers.
- The VaR does not satisfy the sub-additivity inequality $f(x_1 + x_2) \leq f(x_1) + f(x_2)$, which encapsulates the idea that diversification reduces risk.
- The VaR pays no attention to the magnitude of losses when they occur.

A recent example that includes empirical analysis on U.S. stock and bond markets can be found in [Campbell, Huisman, Koedijk 2001].

Conditional Value at Risk

To overcome the VaR's drawbacks, the concept of Conditional Value at Risk (CVaR), also called Expected Shortfall (ES), has been developed:

$$CVaR_\alpha(x) := \frac{1}{1-\alpha} \int_{w:L(x,w)\geq VaR_\alpha(x)} L(x, w)P[dw]. \tag{3.15}$$

CVaR considers the magnitude of losses, satisfies the subadditivity inequality and is a smooth function. Altogether, CVaR is a coherent[10] risk measure.

Nevertheless, calculating CVaR appears to require the calculation of VaR. To overcome this drawback, [Rockafellar, Uryasev 2000] showed that the optimization problem can be reformulated by introducing the function

$$F_\alpha(x, \gamma) := \gamma + \int_\Omega \frac{(L(x, w) - \gamma)_+}{1-\alpha} P[dw], \tag{3.16}$$

and the fact that $F_\alpha(x, VaR_\alpha(x)) = CVaR_\alpha(x)$. Knowing that VaR is a minimizer of the problem $(\min_\gamma F_\alpha(x, \gamma))$, the portfolio optimization problem $\min_x CVaR_\alpha(x)$ is equal to the problem:

$$\min_{x,\gamma} F_\alpha(x, \gamma). \tag{3.17}$$

[10] A risk measure that satisfies the properties monotonicity, sub-additivity, homogeneity, and translational invariance.

This is not a bilevel optimization anymore and omits the calculation of VaR altogether.

Limits of modern portfolio theory

The original optimal investment model in Section 3.1.4 aims to find an optimal (time-continuous) trading strategy over a specified time horizon $[0, T]$ which leads to a stochastic optimal control problem. The reader may have noticed that in contrast, modern portfolio theory and advanced portfolio theory are single-period models. Both profit and risk are integral functions over the probability space Ω and the actual optimization is non-stochastic.

The connection between Merton's investment-consumption model and MPT was provided by the mutual fund theorem using the assumption of perfect markets. Unfortunately, as soon as we take transaction costs or illiquidity effects (imperfect markets) into account, a single-period analysis (e.g. MPT) is not possible anymore.

3.4 Illiquidity and transaction costs

So far we assumed a perfect market according to Definition 3.1 in Section 3.1. Now we discuss modeling approaches for pricing & hedging and portfolio theory when the frictionless and competitive market hypotheses are relaxed.

3.4.1 Transaction costs

Markets are not frictionless anymore when transaction costs compensate the dealer's effort for trading activities. Usually, transaction costs $f : \mathbb{R} \to \mathbb{R}_-$ are charged proportional to the amount of stock traded $\alpha \in \mathbb{R}$:

$$g(\alpha) = -c|\alpha|, \tag{3.18}$$

where $c > 0$. So the agent's total payment ΔB for the trade α is:

$$\Delta B = -\alpha S - g(\alpha) S, \tag{3.19}$$

where S denotes the asset's price. When transaction costs are taken into account, the problem of pricing & hedging as well as optimal portfolio allocation become more complex. The main issue is that a Brownian motion has infinite variation. So it would be infinitely costly to balance a hedging portfolio for a contingent claim (or an investment portfolio) continuously when the underlying follows an Itó process.

In order to price contingent claims under transaction costs, [Leland 1985] proposed to balance the hedging portfolio only at discrete time points. He proposed a Black-Scholes style formula with an adjusted volatility when using a fixed hedging time scale δt. Another method (known as super-replication) requires that the hedging portfolio dominates, rather than replicates, the claim's payoff at maturity. For incomplete markets, [Monoyios 2003] and [Monoyios 2004b] analyzed marginal utility-based pricing under transaction costs. This approach was first proposed by [Davis 1997] and it is similar to the utility indifference pricing, see Section 3.1.6.

Besides pricing & hedging, optimal portfolio allocation is affected by transaction costs too. [Davis, Norman 1990] analyzed Merton's original investment-consumption problem in continuous time under linear transaction costs. They showed that the space spanned by the amount of risky assets and the amount of risk-less asset held in the portfolio divides into three regions: a buy-, a sell- and a hold-region[11]. The portfolio is only balanced back to the hold-region when it goes outside to the buy- or the sell-region. More recent literature on pricing & hedging and portfolio theory with transaction costs is [Barles, Soner 1998] or [Constantinides, Zariphopoulou 1999] for example.

3.4.2 Illiquidity

The relaxation of the competitive market hypothesis introduces the notion of liquidity. Roughly speaking, imperfect liquidity (illiquidity) is an additional cost due to the timing and size of a trade.

Illiquidity is commonly believed to be a major effect in financial markets. Emperical evidence of temporary illiquidity is provided through the Bid-Ask spreads of standard instruments at financial exchanges. Moreover, permanemt price changes after the anouncement of large trades must be considered as illiquidity effects too, whenever no other new information arrived the market.

So there exist two common phenomena due to illiquidity: temporary illiquidity cost and permanent price impacts. [Rogers, Singh 2010] argue that both can be considered as effects of supply & demand on price, but the temporary illiquidity cost "arises because of the need to clear a market over a short time spell, whereas the permanent price impact comes from the clearing of the market" from large trades "over long periods."

[11] As proposed before by [Magill, Constantinides 1976].

Temporary illiquidity costs

By temporary illiquidity costs we mean that an agent, however small, may face costs in trying to trade rapidly but does not affect the price of the underlying asset. In other words, his trading affects the price at which *he* will trade the asset but not the asset's market price itself.

The following examples show how temporary illiquidity cost are modeled in the literature.

- [Touzi, Astic 2007] introduced a liquidation function that "exhibits a bid-ask spread for each portfolio, which includes any transaction cost that the agent may have to pay to a broker". Thus, their model does not separate the liquidity cost from transaction costs.

- [Cetin, Jarrow, Protter 2004] hypothesize "the existence of a stochastic supply curve for a security's price as a function of trade size, for which agents act as price takers." Their supply curve is independent of the agent's past actions, endowments, risk aversion, or beliefs. An example for a stochastic supply curve S is:

$$S(t, \alpha_t) = f(\alpha_t) S(t, 0), \tag{3.20}$$

  where $S(t,0) \in \mathbb{R}_+$ denotes the asset's market price at current time t, $\alpha_t \in \mathbb{R}$ denotes the current trade size and $f \in \mathbb{R}_+$ is a stochastic illiquidity factor. A positive α_t means buying assets and a negative α_t means selling assets. The function $f \in C^2(\mathbb{R}, \mathbb{R}_+)$ must be positive, convex and increasing w.r.t. α_t, with $f(0) = 1$, $f(\alpha_t) < 1$ for $\alpha_t < 0$, and $f(\alpha_t) > 1$ for $\alpha_t > 0$.

- [Rogers, Singh 2010] argue that "the faster an agent wants to buy (sell) the asset, the deeper into the limit order book he will have to go, and higher (lower) will be the price for the later units of the asset bought (sold). However, once a rapid transaction is completed, we suppose that the limit order book quickly fills up again and that the rapid transaction has no lasting effect on the price of the underlying." Therefore they modeled the stochastic factor $f(\cdot)$ in (3.20) through an integration of the order book:

$$S(t, \alpha_t) = \int_0^{\alpha_t} x \rho(x) dx \;\; S(t, 0), \tag{3.21}$$

  where $\rho(x)$ denotes the density of quotes at relative price x.

[Rogers, Singh 2010] and other authors understand illiquidity as nonlinear transaction costs. Let us briefly show how the supply curve model (3.20) can be interpreted as transaction costs. First, let us consider the agent's total payment ΔB for a trade α_t by:

$$\Delta B = -\alpha_t S(t, \alpha_t) = -\alpha_t f(\alpha_t) S(t, 0). \tag{3.22}$$

Now, let us split this payment into the value of the assets $\alpha_t S(t, 0)$, which the agent holds after his purchase, and the additional cost for illiquidity. Then (3.22) becomes:

$$\begin{aligned} \Delta B = & -\alpha_t S(t,0) - \alpha_t f(\alpha_t) S(t,0) + \alpha_t S(t,0) \\ = & -\alpha_t S(t,0) - (f(\alpha_t) - 1)\, \alpha_t\, S(t,0) \\ = & -\alpha_t S(t,0) - g(\alpha_t) S(t,0), \end{aligned} \tag{3.23}$$

where the function $g(\alpha) = (f(\alpha) - 1)\,\alpha$ denotes the pure transaction costs relative to the asset's price; compare to (3.19). From the properties of f we derive that:

- $g(\cdot) \geq 0$; so g is positive,
- $g(0) = 0$; so there are no transaction cost when there is no trade,
- $\dot{g}(0) = 0$, $\dot{g}(\alpha) < 0$ when $\alpha < 0$ and $\dot{g}(\alpha) > 0$ when $\alpha > 0$; so g has its global minimum at 0,
- $\ddot{g}(\alpha) < 0$ as long as $\alpha > -\frac{\ddot{f}+\dot{f}}{\ddot{f}}$; so g is locally convex around 0.

We see that $g(\cdot)$ fulfills all realistic assumptions of a (nonlinear) transaction cost function.

The Flash Crash as an example of temporary illiquidity effects

The Flash Crash [Wallstreet 2010] was a stock market crash on Thursday May 6, 2010 in which the Dow Jones Industrial Average plunged about 1000 points (about 9%) and recovered most of those losses within about 20 minutes. Many investigators mentioned that the illiquidity of the market (caused through high market fragmentation[12]) made the Flash Crash first possible. For example, using intraday trade data from January 1994 - September 2011, [Madhavan 2011] found "that fragmentation now is at the highest level recorded".

[12]Market fragmentation is defined as "emergence of new segments (in a previously homogeneous market) which have their own distinct needs, requirements, and preferences. These fragments reduce the effectiveness of mass marketing techniques and erode brand loyalty", see [businessdictionary].

Permanent price impacts

By permanent price impacts we mean that an agent's trading activity affect the asset's market price. Since trading decisions are made depending upon the assen's price, but the trade simultaneously influence this price, the phenomenon is called market feedback in literature. [Papanicolaou, Sircar, 1998] for example present a class of pricing models that accounts for the feedback effect from the Black-Scholes dynamic hedging strategies on the price of the asset, and from there back onto the price of the derivative.

An asset price model that accounts for permanent price impacts could have the following form:

$$dS_t = \mu(t, S_t, \alpha_t)dt + \sigma(t, S_t, \alpha_t)dW_t, \tag{3.24}$$

compared to the standard Itô process 3.5 in Section 3.1.4. We will consider the valuation of swing options under permanent price impacts as an example application in Part III, Section 15.

Nevertheless, the proper treatment of permanent price impact is problematic. [Rogers, Singh 2010] mentioned that, "if the actions of a single trader will shift the price, then logically the actions of every trader will shift the price, and in order to understand this effect fully we would have to build a model which accounted for the behavior of all the agents in the market."

For example, [Schönbucher, Wilmott 2000] consider "the free round trip phenomenon, where the large agent rapidly sells and then buys back a large amount of stock, forcing the price instantaneously to drop, and if this round trip is not costly, then the large agent could make profits by selling down-and-out calls and subsequently knocking them out by a round trip."

3.4.3 Other costs

In literature, the concept of transaction costs is transfered to a variety of other costs. [Lobo, Fazel, Boyd 2007] stated for example that "transaction costs can be used to model a number of costs such as brokerage fees, bid-ask spreads, taxes or even fund loads". We will use the concept of transaction cost to account for temporary illiquidity effects in our model for optimal portfolio allocation of commodity related assets in Section 4.4.3.

Note that in practice, transaction costs indirectly keep the investor away from unrealistically frequent changes of his portfolio. Nowadays, the European government discusses about introducing artificial transaction costs

to prohibit high frequency trading of banks. Uncontrolled high frequency trading is supposed to disrupt financial markets.

3.5 Market share

Next to maximizing profits and minimizing risks, gaining market share is also an important goal of a company's decision maker. [investopedia, market share] defines market share as the percentage of an industry or market's total sales that is earned by a particular company over a specified time period. The measure gives a general idea about the relative size of a company in comparison to its competitors.

The true reasons for increasing market share may be the association of stability and bargaining power. However, an economic reason is given by the concept of *economy of scale*. It says that, the greater the operation scale, the more efficient and profitable is the production process. There are many empirical studies that have questioned the validity and generalizability of the relationship between market share and profitability.

[Szymanski, Bharadwaj, Varadarajan 1993] provided a literature overview by performing a meta-analysis on 276 market share-profitability findings from 48 studies. The studies use different factors to describe how market share affects profitability and estimate the correlation between both.

Common models use the following relationship between market share S and profitability P:

$$P = \beta_0 + \beta_1 S + \sum_{i=2}^{N} \beta_i F_i, \tag{3.25}$$

where F_i, $i = 2, ..., N$ denote appropriate economic factors and $\beta_i \in \mathbb{R}$ are coefficients to be estimated by regression. Factors that indicate marketing expenses or the duration of a product's live cycle can influence the change of market share S over time.

On the one hand, a portfolio model could account for market share through simply adding (3.25) to the profit function. On the other hand, increasing market share may be a separate goal for investors next to increasing profitability. Then, a measure of market share (mshare) could be added to the portfolio's objective function:

$$u(X) = \text{profit}(X) - \lambda\, \text{risk}(X) + \gamma\, \text{mshare}(X), \tag{3.26}$$

where γ is a coefficient that indicates the *power addiction* of the investor.

Part II
Optimal portfolio allocation of commodity related assets

In this part of the thesis we develop and analyze our model for optimal portfolio allocation of commodity related assets. By commodity related assets we mean every physical (or financial) assets that involves (or replicates) production, transportation, processing, storage or trading of commodities. We sometimes refer to these assets in a shorter way as production assets.

In Section 4 we develop the general discrete-time model. In Section 5 we reduce the complex discrete-time model to an analyzable continuous-time model. We present analytical solutions for simple cases in Section 6 and numerical solutions for the general one asset case in Section 7. Furthermore, we apply the reduced continuous-time model to a special case from the energy sector, a portfolio of power plants, and present numerical results in Section 8.

4 A discrete-time model for general production assets

We classify first our portfolio approach and place it in the vast portfolio literature. We define the special type of *commodity related assets* that we are dealing with, their *allocation* and the term *portfolio owner*. Further we clarify the type of portfolio optimization that we focus on, the distinction between production asset and commodity trading and the limits of our model approach. We then list a few real-world applications.

4.1 Classification

Let us first explain what we mean by the term production asset.

Definition 4.1 (Production asset)
A production asset is the financial view of a machine, a facility or a production process which transforms several inputs into several outputs. The transformation can be a result of production, transportation, processing, storage or trading. The major inputs and outputs must be homogeneous goods (commodities) which are tradeable at open exchanges in a free market economy.

Production assets differ in their location (e.g. countries, states, nations or continents), their time of availability (e.g. 24-7, at daytime, seasonal or weather dependent) and their technology (e.g. operation flexibility, physical constraints, quality or effectiveness). Note that these three properties of production assets correspond to the the three key properties of location, time and quality which we introduced in the preliminaries in Section 1.1. Typical examples for production assets are power plants that transform energy sources into electricity.

Definition 4.2 (Allocation)
By allocation we mean the following possible actions to change the capacities of production assets:

1. *relocate an asset physically,*
2. *upgrade an asset to a new technology in order to improve operation flexibility or time of availability,*
3. *maintain or fully replace an old assets that shows weak,*
4. *build a new asset from scratch or reassemble an old asset in order to increase or decrease capacities, respectively,*
5. *buy or sell an asset in order to increase or decrease capacities, respectively.*

In reality, each type of allocation is an investment decision of the company. Investment decisions are usually made at regular time intervals, e.g. on a monthly, quarterly or yearly basis.

Definition 4.3 (Portfolio owner)
The portfolio owner is a company (an organization or a person) that owns and operates several production assets. It also holds a bank account and is able to save and borrow money with interest. The company has a specified utility (risk preference) and is able to change its portfolio over time through allocations. We suppose that the organization is small compared to the market, such that these allocations do not influence the market conditions.

For example, a power producer holds a portfolio of power plants and aims to maximize its return through optimal investments while considering possible risks. In another example, the portfolio is the sum of all production assets of a nation or community. Then the portfolio owner is the population, represented by an authority, and aims to minimize costs and pollution while considering energy security.

4.1.1 Type of portfolio optimization

Our goal is to maximize a company's utility through an optimal allocation of production assets over a specified time horizon. For this purpose we utilize the concept of portfolio theory from mathematical finance, see Section 3. Our main questions is:

What is the optimal rate a company should invest in different types of production assets, given an initial capital?

Our portfolio approch has the following properties:

- Utility is gaind from **consumption** of wealth over time that is taken out of the bank account, as well as from terminal wealth.
- The company starts with a **given initial portfolio** at time 0 and searches for an optimal strategy until a fixed time horizon T. So we start with the actual portfolio of the company and **optimize its future investments**. Other approaches search for the portfolio that is optimal today and simply advice the portfolio holder to shift its capacity towards this imaginary optimal portfolio through allocation in the future. It is clear that our approach is much closer to reality since it accounts for the real allocation process over time.
- The capacity allocation over the time interval $[0, T]$ is a **multi-step discrete-time process**. At every time step, a new investment decision is made. This is different to other approaches which only consider a single-step model like modern portfolio theory.
- The uncertainties (risk drivers) are **commodity price changes and changes of the assets' value** over time. All other risk factors like political or economical risks as well as competition are encapsulated in the commodities price and assets value changes.
- We consider **temporary illiquidity costs** for allocations in the market of production assets.

4.1.2 Distinction between production assets and their underlying commodities

The important specialty of our model is that we are not trading commodities but rather the commodity processing production assets, which are operating in the commodity market. The two main differences between trading production assets and their underlying commodities are:

- Owning a production assets one can use the **operation flexibility** to optimize returns.

- One has to account for **essential industrial mechanisms** of the production asset like construction costs, deterioration, variable operation costs, fixed maintenance costs and physical limitations of operation.

Therefore, each production asset is a real option according to our Definition 1.2 in Section 1.4.3.

Figure 5 is an example flow chart that visualizes the portfolio allocation of production assets and their operation in commodity markets. Both markets are uncertain. The blue arrows mean illiquid trading while the red arrows mean liquid trading. The red arrow inside the capacity boxes indicate the operation flexibility.

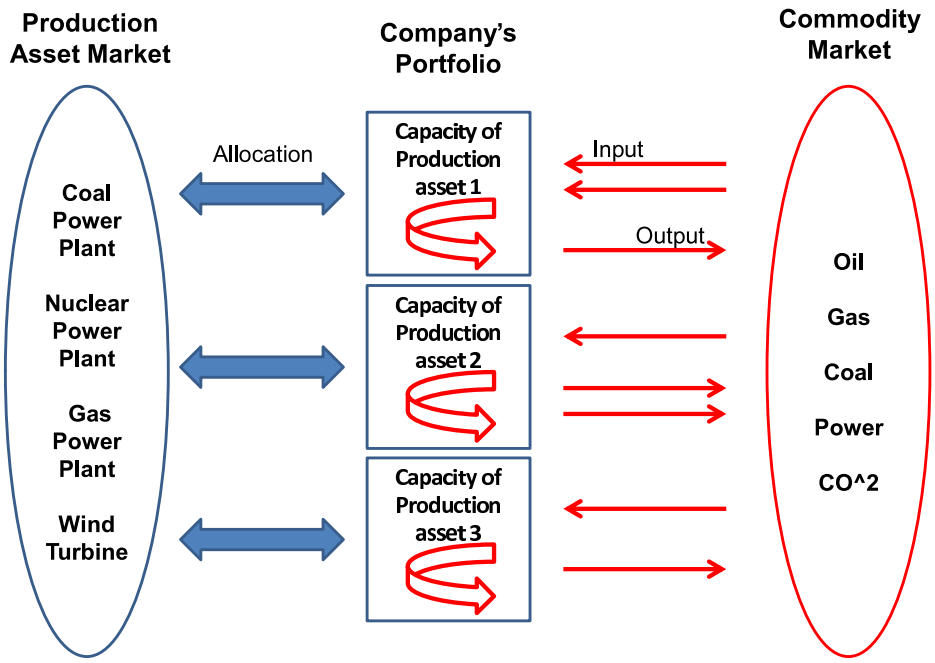

Figure 5: Example flow chart for portfolio allocation of generation assets which operate in the energy and power market.

4.1.3 Limits of our model

Details about pre-processing of inputs or further processing of outputs are out of scope. The corresponding intersections in the commodity chain are modeled by uncertain market prices. Figure 6 shows an example of a power producer. Not considered are 1) mining or pre-refinement of commodities 2) transportation or storage of inputs and outputs 3) the precise physical or chemical reactions and 4) short-term operation scheduling.

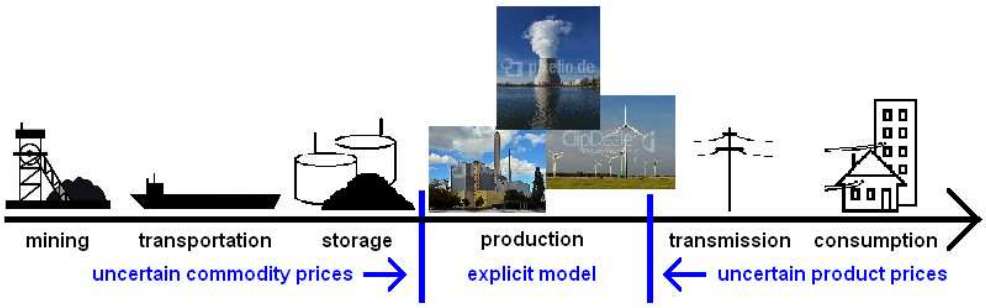

Figure 6: Commodity chain for power generation from mining fossil energy sources to power consumption. The blue markers on the time line indicate the limits for a generation assets portfolio.

Note that we leave aside the issues of foreign exchange rates and insurance fees. However, both financial aspects could be included in an intuitive manner. We also do not consider the investment of money into other industry sectors or markets. Nevertheless, each out-of-scope topic may already be addressed in literature by other research groups, technical engineers or facility operators.

4.1.4 Possible applications

The model is especially applicable for companies in the heavy industry sectors where homogeneous goods are refinement and processed (e.g. metal sublimation, chemical refinement, cracking oil, power generation and food processing). Our business case in Section 8 deals with a portfolio of power plants.

The model can be applied also to companies dealing with commodities transportation. There a portfolio consists of tankers, ships, pipelines, trucks, storages facilities and reloading stations where each asset has its specific capacity

and properties. For example the oil shipping from Saudi Arabia to Europe is a production process. Inputs are oil at the exporting harbors (Saudi Arabia) and outputs are oil at the importing harbors (Europe).

Another suitable application from the service sector is a provider of online computing capacity or storage space. The company's production assets are their hardware servers. The main inputs are local power and local labor while the main output is the globally available server capacity. All three are homogeneous goods. Besides technical properties, assets mainly differ in their location. This implies different power and labor costs, political risks, risk of social stability or power outage risk. A web company like Google[13] needs a tremendous amount of computing power and should optimally choose their server locations in the world.

4.1.5 A comment on modern portfolio theory approaches

We introduced modern portfolio theory (MPT) and advanced portfolio theory in Sections 3.2 and 3.3, respectively. Both approaches are only applicable in perfect markets but some authors applied them to portfolios of production assets, see [Awerbuch, Stirling, Jansen, Beurskens 2005] for a case on power generation assets. In order to do so, they calculated the mean and the variance of the discounted cash flows (DCFs, see Section 1.4.1) of each production asset. Assuming that the DCFs are normally distributed, they treated production assets like standard financial assets and applied the single-step MPT. In contrast we deal explicitly with the specific properties of production assets and the process of portfolio allocation over time. We discuss this aspects in more detail in our Literature review in Section 8.1.

4.2 Definition of variables

Let $[0, T]$ denote the fixed time horizon of interest. Sine investment decisions are made on a regular time frequency, let:

$$\left\{ t_i \in [0,T],\ i = 0, ..., N^T \;\middle|\; t_0 = 0,\ t_{N^T} = T \right\} \tag{4.1}$$

denote an equidistant time grid with step size $\Delta t = T/N^T$.

4.2.1 Capacities and prices of production assets

Let us suppose that an industry sector consists of N^A different kinds of production assets. Let $a_{t_i} \in \mathbb{R}_+^{N^A}$ denote the amount of assets held by the

[13] www.google.com

company during the time interval $[t_i, t_{i+1})$. The amount of assets is measured in operation capacities. Let $p^A_{t_i} \in \mathbb{R}^{N^A}_+$ denote the asset's market price for one unit of capacity. The market price is less related to the construction costs but rather to a fair value for which assets are exchanged, like in mergers & acquisitions. Let $\tilde{a}_{t_i}$ and $\tilde{b}_{t_i} \in \mathbb{R}^{N^A}_+$ denote the amount of capacities purchased and sold through a transaction at time t_i, respectively.

4.2.2 Bank account and consumption

In addition to investment into production assets, the company can put its money into a secure bank account that pays a constant (risk-free) interest rate r. So, let $B_{t_i} \in \mathbb{R}$ denote the amount of money held in the company's bank account. When B_{t_i} is negative, the company is borrowing money. Let $\chi_{t_i} \in \mathbb{R}_+$ denote the consumption of money out of the bank account during $[t_i, t_{i+1})$. The consumption could be a dividend payment to the share holders or a payout to the private owners.

4.2.3 Commodity prices and uncertainties

Let $(\Omega, \mathcal{F}, P)$ denote a probability space endowed with a filtration $\mathbb{F} := \{F_t\}_{t\in[0,T]}$ and let $X^N[0,T]$ denote the space of all time-continuous Markov processes over $[0,T]$, that are N-dimensional, $\mathbb{F}$-measurable and real-valued.[14]

Let us suppose the industry sector deals with N^P different outputs (products) and N^C different inputs (commodities). Let $p^P \in X^{N^P}[0,T]$ and $p^C \in X^{N^C}[0,T]$ denote the price processes of outputs and inputs, respectively. The introduced filtration $\mathbb{F}$ is generated by all uncertain factors of the model.[15] Roughly speaking, F_t represents the information set at time t. Additionally, let $q : [0,T] \to \mathbb{R}^{N^A}_+$ denote the operation level for each asset type. Note that p^P, p^C and q may not be constant over $[t_i, t_{i+1})$.

In order to simplify notation we introduce the following **discrete-time state** sequences:

- the portfolio capacity $a := \left\{a_{t_i} \in \mathbb{R}^{N^A}\right\}_{i=0}^{N^T}$,
- the production asset price $p^A := \left\{p^A_{t_i} \in \mathbb{R}^{N^A}\right\}_{i=0}^{N^T}$

[14] The definitions of probability space, filtration and Markov processes can be found in the appendix A.1.

[15] A brief introduction about the generation of a filtration out of random variables can be found in the appendix A.1.

- the bank account $B \coloneqq \{B_{t_i} \in \mathbb{R}\}_{i=0}^{N^T}$

and the following **discrete-time control** sequences:

- the buying rate $\tilde{a} \coloneqq \left\{\tilde{a}_{t_i} \in \mathbb{R}_+^{N^A}\right\}_{i=0}^{N^T}$,
- the selling rate $\tilde{b} \coloneqq \left\{\tilde{b}_{t_i} \in \mathbb{R}_+^{N^A}\right\}_{i=0}^{N^T}$,
- the consumption rate $\chi \coloneqq \{\chi_{t_i} \in \mathbb{R}_+\}_{i=0}^{N^T}$.

We also have the following **continuous-time state** processes:

- the price processes of outputs $p^P \in X^{N^P}[0,T]$,
- the price processes of inputs $p^C \in X^{N^C}[0,T]$,

and the following **continuous-time control** process:

- the operation level $q : [0,T] \to \mathbb{R}_+^{N^A}$

4.3 The optimization problem

Let us assume that the company starts with an initial portfolio of assets a_0 and bank account B_0, and use the fixed allocation policy χ. Then, the bank account evolves in time by:

$$B_{t_{i+1}} = B_{t_i} + rB_{t_i} - \chi_{t_i} + \text{profit}_{t_i}(a_{t_i}, \tilde{a}_{t_i}, \tilde{b}_{t_i}, p_{t_i}^A, p_{t_i}^P, p_{t_i}^C), \tag{4.2}$$

and the amount of assets evolves in time by:

$$a_{t_{i+1}} = a_{t_i} + \tilde{a}_{t_i} - \tilde{b}_{t_i} - r^{\text{det}} a_{t_i}. \tag{4.3}$$

Here, the function $\text{profit}_{t_i} : \mathbb{R}_+^{N^A} \times \mathbb{R}_+^{N^A} \times \mathbb{R}_+^{N^A} \times \mathbb{R}_+^{N^A} \times \mathbb{R}_+^{N^P} \times \mathbb{R}_+^{N^C} \to \mathbb{R}$ denotes the expected profits over $[t_i, t_{i+1})$ that are gained from the asset's portfolio. The expactation depends on the initial prices $p_{t_i}^A$, $p_{t_i}^P$ and $p_{t_i}^C$ at t_i. The rate $r^{\text{det}} \in [0,1]^{N^A}$ denotes the average depreciation rate of the assets.

The objective is to maximize the portfolio owner's utility. Let $U_1, : \mathbb{R}_+ \to \mathbb{R}$ denote the utility out of consumption χ_{t_i} and let $U_2, : \mathbb{R}_+ \to \mathbb{R}$ denote the terminal utility out of terminal wealth $a_T p_T + B_T$. Then the objective function is given by:

$$J(a_0, B_0, \chi, \tilde{a}, \tilde{b}) = E^{\chi, \tilde{a}, \tilde{b}}\left[\sum_{i=0}^{N^T} e^{-\delta t_i} U_1(\chi_{t_i}) + e^{-\delta T} U_2(a_T p_T + B_T) \middle| \mathrm{F}_0\right], \tag{4.4}$$

where $\delta \geq 0$ denotes the discount rate for deferment and $E^{\chi}[]$ denotes the expectation when using the fixed control policies χ. Note that we use a constant discount factor $e^{-\delta t_i}$ for each $[t_i, t_{i+1})$. The value of the portfolio is given by:

$$V(a_0, B_0) = \sup_{\chi, \tilde{a}, \tilde{b}} J(a_0, B_0, \chi, \tilde{a}, \tilde{b}). \tag{4.5}$$

Note that the estimation of the portfolio's value (4.5) imposes a multi-dimensional stochastic optimal control problem.

4.3.1 The choice of the utility functions

In the preliminaries Section 3, we introduced the concept of utility functions and how they encapsulate the investor's risk preferences. For our portfolio allocation model we consider the class of HARA utility:

$$U(x) = \frac{1-\alpha}{\alpha}\left(\frac{\beta x}{1-\alpha} + \nu\right)^{\alpha}, \tag{4.6}$$

with the restrictions:

$$\alpha \neq 1, \quad \beta > 0, \quad \left(\frac{\beta x}{1-\alpha} + \nu\right) > 0, \quad \nu = 1 \text{ if } \alpha = \pm\infty. \tag{4.7}$$

HARA utility includes a wide range of different risk-aversion cases like increasing, decreasing, and constant, absolute and relative risk aversion, see Section 3.1.3.

Another reason is that HARA utility is widely used in the literature. Therefore we can compare our results for commodity related assets with already analyzed cases for standard financial assets. In particular, [Merton 1971] obtained analytical solutions for the standard investment-consumption problem in case of HARA utility while assuming a perfect market and log-normally distributed prices. We discuss analytical solutions in Section 6.

The actual utility function must be chosen from the HARA class appropriately to the specific business case. In our business example in Section 8 we use the case of negative exponential utility:

$$U_1(x), U_2(x) = \frac{1 - \exp(-\gamma x)}{\gamma}, \quad \gamma > 0. \tag{4.8}$$

Negative exponential utility has constant absolute risk aversion (CARA) where the Arrow-Pratt coefficient of absolute risk aversion is $\lambda_1 = \gamma$, see Section 3.1.2. Other HARA utility functions can be applied in the same manner. Note that in particular business cases the shape of terminal utility U_2 may slightly differ from U_1.

Empirical studies on utility functions

On the one hand, experimental and empirical evidence is mostly consistent with decreasing absolute risk aversion, see [Friend, Blume 1975]. It means that as the investor's wealth increases, his aversion to risk decreases. CRRA implies a decreasing absolute risk aversion for example.

On the other hand, it may not be always true in practice that the risk aversion increases with decreasing wealth; especially not in the limit case. For example, log-utility with $\lambda = 1$ is a CARA utility function but it converges to $-\infty$ when consumption converges to zero.

From empirical studies, [Kallberg, Ziemba 1984] "find that utility functions with similar levels of Arrow-Pratt absolute risk aversion result in similar optimal portfolios irrespective of the functional form of the utility." Therefore, negative exponential utility is a reasonable choice for practical cases when total wealth stays in a reasonable order of magnitude.

4.4 Profit from the asset portfolio

The monetary profit from a portfolio of production assets is the sum of:

1. Gains from operations. Thus, let $\text{gain}_{t_i} : \mathbb{R}_+^{N^A} \times \mathbb{R}_+^{N^P} \times \mathbb{R}_+^{N^C} \to \mathbb{R}$ denote the expected gain from operations during $[t_i, t_{i+1})$.

2. Fixed costs that arise from the pure holding of production assets. Thus, let $\text{cost}_{t_i}^{\text{fix}} : \mathbb{R}_+^{N^A} \to \mathbb{R}$ denote the fixed costs of the portfolio during $[t_i, t_{i+1})$, that are independent from operation.

3. Costs due to the allocation process. Thus let $\text{cost}_{t_i}^{\text{alloc}} : \mathbb{R}_+^{N^A} \times \mathbb{R}_+^{N^A} \times \mathbb{R}_+^{N^A} \to \mathbb{R}$ denote these total costs from buying and selling assets at time t_i.

Summarizing these components, the portfolio profit is given by:

$$\begin{aligned} &\text{profit}_{t_i}(a_{t_i}, \tilde{a}_{t_i}, \tilde{b}_{t_i}, p_{t_i}^A, p_{t_i}^P, p_{t_i}^C) = \\ &\text{gain}_{t_i}(a_{t_i}, p_{t_i}^P, p_{t_i}^C) - \text{cost}_{t_i}^{\text{fix}}(a_{t_i}) - \text{cost}_{t_i}^{\text{alloc}}(\tilde{a}_{t_i}, \tilde{b}_{t_i}, p_{t_i}^A). \end{aligned} \tag{4.9}$$

Note that the function $\text{gain}_{t_i}(a_{t_i}, p_{t_i}^P, p_{t_i}^C)$ depends only on the initial prices $p_{t_i}^P$, $p_{t_i}^C$ at time t_i. This is crucial since the company makes allocation decisions at time t_i without the knowledge of the explicit price paths p_t^P, p_t^C for

$t \in [t_i, t_{i+1})$. Now, we define the three components of (4.9) in more detail.

We state the unit of each variable in the following paragraphs, such that the reader is able to check consistency. Basically, we distinguish between two different types of units, capacity unit [cap] and operation unit [op]. The operation unit is equal to the capacity unit integrated over time, [op] = [cap] [t]. The specific commodity unit is indicated by [unit].

4.4.1 Operation gains

First let us take a closer look at the composition of gains from operation and their cash-flows.

Estimation of earnings

Earnings are made solely from selling goods or services that are produced during operation. We define operation earnings in terms of money per operation unit (/[op]) and assume an instantaneous selling of the outputs at market prices. Thus, the operation earnings at time $\tau \in (t_i, t_{i+1}]$ are given by:

$$\mathrm{e}_\tau^{\text{earning}} := p_\tau^P \, \mathrm{QP}(q_\tau),$$

where $\mathrm{QP} : \mathbb{R}_+^{N^A} \to \mathbb{R}^{N^P \times N^A}$ denotes the output production matrix in terms of the output's unit per operation unit ([unit]/[op]). Note that p^P is given in terms of money per commodity unit (/[unit]). Note also that QP depends on the current operation level q_τ. An additional time dependency of the operation matrix QP_t could account for technology advancements towards more effective production over time. We leave this aspect aside for the sake of simplicity.

Estimation of variable costs

By definition, variable costs occur only through operation. We define variable costs in terms of money per operation unit (/[op]), in the same way as earnings, and assume an instantaneous buying of the inputs at market prices. The variable costs at time $\tau \in (t_i, t_{i+1}]$ are given by:

$$c_\tau^{\text{var}} := p_\tau^C \, \mathrm{QC}(q_\tau) + c^{\text{op}}(q_\tau),$$

where $\mathrm{QC} : \mathbb{R}_+^{N^A} \to \mathbb{R}^{N^C \times N^A}$ denotes the input processing matrix in terms of the input's unit per operation unit ([unit]/[op]). Note that p^C is given in terms of money per commodity unit (/[unit]). The function $c^{\text{op}} : \mathbb{R}_+^{N^A} \to \mathbb{R}_+^{N^A}$ denotes the usual operation costs per operation unit (/[op]). In c_τ^{op}

we include all incremental costs that are not proportional to the input, e.g. repairing costs due to wearing, energy costs, start-up and shut-down costs and labor costs - if these are not considered as an own input variable already. Note also that QP and c^{op} depend on the current operation level q_τ.

The real option approach

The operation gain during $[t_i, t_{i+1})$ is the expected accumulated spread between earnings and variable costs, assuming optimal operation. The expectation is conditional to commodity prices $p^P_{t_i}$ and $p^C_{t_i}$ at t_i. It turns out that the operation gain is a real option according to Definition 1.2 in Section 1.4.3. Using this approach, the optimal gain is given by:

$$\text{gain}_{t_i}(a_{t_i}, p^P_{t_i}, p^C_{t_i}) = \sup_{q \in A(a_{t_i})} E^q_{t_i}\left[\int_{t_i}^{t_{i+1}} \left(p^P_\tau \, \text{QP}(q_\tau) - p^C_\tau \, \text{QC}(q_\tau) + c^{\text{op}}(q_\tau)\right) \, q_\tau \; d\tau \middle| p^P_{t_i}, p^C_{t_i} \right], \tag{4.10}$$

where $A(a_{t_i})$ denotes the set of applicable operation policies:

$$A(a_{t_i}) = \{ \{q_\tau, \tau \in [t_i, t_{i+1})\} \mid \text{ feasible according to } a_{t_i}\} \tag{4.11}$$

Note that q_τ is defined in terms of capacity units [cap] and the function gain_{t_i} has the unit . Note also that we replaced F_{t_i} by $p^P_{t_i}, p^C_{t_i}$ in the conditional expectation, since both stochastic processes are assumed to be Markov and the controls only depend on the current states at t_i.

Physical constraints of operation policies

The control $\{q_\tau, \tau \in [t_i, t_{i+1})\}$ must lay in the set of applicable operation policies $A(a_{t_i})$. In (4.11) the term *feasible* indicates that physical constraints may limit the control space. We already presented common physical constraints of real options in Section 1.4.3 and particular constraints for generation assets in Section 2.4. The proper selection of physical constraints depends on the specific business case.

4.4.2 Fixed costs

Maintenance and investment costs must be payed in advance for availability of capacities, no matter what the operation is. Since these costs are independent of operation q_τ they are called fixed costs. The average fixed costs during time interval $[t_i, t_{i+1})$ are proportional to the capacity a_{t_i} and given by:

$$\text{cost}^{\text{fix}}_{t_i}(a_{t_i}) := \left(c^{\text{main}}_{t_i} + c^{\text{inv}}_{t_i}\right) \, a_{t_i}, \tag{4.12}$$

where $c_t^{\text{main}} \in \mathbb{R}_+^{N^A}$ denotes the fixed maintenance costs and $c_t^{\text{inv}} \in \mathbb{R}_+^{N^A}$ denotes the annuities of the investment costs. Note that $c_{t_i}^{\text{main}}$ and $c_{t_i}^{\text{inv}}$ are given in terms of money per capacity unit (/[cap]) and the capacity a_{t_i} naturally has the unit [cap].

Note that we consider investment costs in the form of annuities in our model. This approach has the following four advantages:

1. Construction costs for a machine or even a huge facility are not paid all at once in operation. They are rather financed through credit with a risk-adjusted credit rate $r^{\text{cred}} \in [0,1]$ - including a potential risk spread - and payed back in annuities over the full lifetime.

2. The commissioning and the decommissioning of single production assets must not to be retraced over time. Therefore, we do not need additional state variables that trace the state of the assets life cycle. Moreover, we do not need to separate similar assets but consider only the sum of their capacities.

3. The investment is independent of former states of the system. This fact ensures that the decision variable a_{t_i} is a Markov control.

4. We need not be concerned with a terminal value or terminal costs of an asset at the time of disposal. All these costs are payed as annuities during each holding period.

The annuities can be derived in the following way. Assume an asset of type j that is commissioned at time t. Let $C_t^{\text{invest},j} \in \mathbb{R}_+$ denote the net present value of investment costs that includes all costs for construction, commissioning and decommissioning. Let $D_t^{\text{cred}} = \frac{1}{(1+r^{\text{cred}})^t}$ be the discount factor for the corresponding credit. Furthermore let T^j denote the average lifetime for an asset of type j. Then, the annuities $c_{t_i}^{\text{inv},j}$ for all t_i are given by[16]:

$$c_{t_i}^{\text{inv},j} = C_{t_i}^{\text{invest},j} \frac{D_{\Delta t}^{\text{cred}} - 1}{\left(D_{T^j}^{\text{cred}} - 1\right) D_{\Delta t}^{\text{cred}}}.$$

Additionally, let us consider the deterioration rate $r^{\text{det},j} \in [0,1]$ from equation (4.3). Then the annuities can account for deterioration by:

$$c_{t_i}^{\text{inv},j} = C_{t_i}^{\text{invest},j} \frac{D_{\Delta t}^{\text{cred}}\left(1 - r^{\text{det},j}\right) - 1}{\left(D_{T^j}^{\text{cred}}\left(1 - r^{\text{det},j}\right)^{T^j} - 1\right) D_{\Delta t}^{\text{cred}}\left(1 - r^{\text{det},j}\right)}. \tag{4.13}$$

[16] Deduced from $C^{\text{invest}} = \sum_{k=1}^{n} c^{\text{inv}} D^k = c^{\text{inv}} \frac{D^n - D}{D-1}$.

Note that the time dependency of $c^{\text{main}}_{t_i}$ and $C^{\text{invest},j}_{t_i}$ expresses possible technology advancements towards cheaper construction and maintenance.

4.4.3 Allocation costs

For the allocation of production assets the company has to pay the asset's market price. Furthermore, we assume that an initiator of a trade faces costs $\text{cost}^{\text{ill}}_{t_i} : \mathbb{R}^{N^A}_+ \times \mathbb{R}^{N^A}_+ \times \mathbb{R}^{N^A}_+ \to \mathbb{R}_+$ due to temporary illiquidity in the market of production asset:

$$\text{cost}^{\text{alloc}}_{t_i}(\tilde{a}_{t_i}, \tilde{b}_{t_i}, p^A_{t_i}) \;=\; (\tilde{a}_{t_i} \,-\, \tilde{b}_{t_i})\, p^A_{t_i} \;+\; \text{cost}^{\text{ill}}_{t_i}(\tilde{a}_{t_i} - \tilde{b}_{t_i}). \tag{4.14}$$

We explain these additional costs and possible other costs in the following paragraphs.

Temporary illiquidity costs

Whenever a production asset is bought from a competitor or sold to him, the counterparties have to negotiate the price. One the one side, we assumed that there exists a "fair" market price (value) p^A of the considered asset. On the other side, commodity markets are usually illiquid, see [Geman 2005], and the market for production assets is naturally less liquid that the market of their underlying commodities. From demand & supply theory it is clear that, the more assets are demanded at the same time, the higher will be the price per asset.

The same price effect occurs when a large amount of the same production assets should be build over a short time period. Here, costs increase with the speed of installation. An example is the realization of renewable energy in Germany where the authorities aim is a fast installation of wind wheels and solar panels. High subvention fees have been paid to increase the willingness to install solar panels in recent years. Also, a fast installation of underwater transmission lines for off-shore wind wheels is expected to be a very costly venture, see [Bloomberg New Energy Finance 2012], section 5.

It is clear that the described costs depend on the total allocation rate $\tilde{a}_{t_i} - \tilde{b}_{t_i}$ of the company while the allocation rate itself does not influence the market value p^A of the assets. Therefore the concept of temporary illiquidity costs, which we introduced in Section 3.4.2 in the preliminaries, is an appropriate way to account for this cost effect.

Let us assume that the costs for the allocation of different assets are independent among each other:

$$\text{cost}^{\text{ill}}_{t_i}(\tilde{a}_{t_i} - \tilde{b}_{t_i}) = \sum_{j=1}^{N^A} \text{cost}^{\text{ill},j}_{t_i}(\tilde{a}^j_{t_i} - \tilde{b}^j_{t_i}), \tag{4.15}$$

Then, according to equation (3.23) in Section 3.4.2, the temporary illiquidity costs have to following properties for each asset $j = 1, ..., N^A$:

- $\text{cost}^{\text{ill},j}_{t_i}$ is smooth, positive and convex,
- $\text{cost}^{\text{ill},j}_{t_i}$ has its global minimum at zero with $\text{cost}^{\text{ill},j}_{t_i}(0) = 0$.

Note that we modeled the temporary illiquidity costs in the form of absolute convex transaction costs (i.e., not relative to the assets price p^A). We believe that a global convexity property is more suitable than a local convexity property as in [Cetin, Jarrow, Protter 2004] or [Rogers, Singh 2010], see Section 3.4.2. Also [Pennanen, Penner, 2010] assumed convex transaction costs for commodity claims with physical delivery in order to model the temporary illiquidity effect.

The assumption of convex transaction costs ensures that the optimal portfolio changes $\tilde{a}_{t_i} - \tilde{b}_{t_i}$ are bounded while they still discourage frequent small changes in the portfolio, in compliance with reality. Moreover, convex *transaction costs* avert practically unrealistic bang-bang controls, which occur for linear transaction costs as shown by [Davis, Norman 1990]. We will discuss this effect in more detail in Section 6.3.

Transaction costs

We introduced the concept of real[17] transaction costs in the preliminaries in Section 3.4.1. Transaction costs are usually proportional to the size of the trade and compensate the broker's effort. Since production assets are not traded at exchanges, an allocation does not impose brokerage fees. Nevertheless, buying and selling production assets is a project for which costs arise from negotiations, location site examination, application for permits, risk analysis, advisory services and other efforts. These costs also arise when a new production asset is built from scratch and could be modeled as proportional transaction costs.

$$\text{cost}^{\text{trans}}(\tilde{a}_{t_i}, \tilde{b}_{t_i}) \;=\; d_1\, \tilde{a}_{t_i} \;+\; d_2\, \tilde{b}_{t_i}, \tag{4.16}$$

[17] In difference to illiquidity costs that are modeled as transaction costs

where $d_1, d_2 > 0$ are constants. Since we have no empirical evidence or literature references we do not include proportional transaction cost into our general model of production assets. However, we include a brief analysis in our business case in Section 7.3 that shows the effect of an additional proportional transaction cost.

4.5 Price dynamics

In our model, the uncertainties are described by random variables. As defined in Section 4.2, the input prices p^C and the output prices p^P are continuous-time stochastic Markov process and their dynamics are given by stochastic differential equations (SDE). The asset prices p^A is a discrete-time stochastic process.

The dynamics of the stochastic processes depend on the asset's industry sector and the corresponding commodities. In the preliminaries Section 1.2, we introduced several commodity price models. In most cases the SDEs are of the form:

$$\begin{aligned} dp_t^P &= \mu^P(t, p_t^P)dt + \sigma^P(t, p_t^P)dW_t^P, \\ dp_t^C &= \mu^C(t, p_t^C)dt + \sigma^C(t, p_t^C)dW_t^C, \end{aligned} \tag{4.17}$$

and the discrete-time processes are of the form:

$$\Delta p_{t_i}^A = \mu^A(t_i, p_{t_i}^A)\Delta t + \sigma^A(t_i, p_{t_i}^A)\Delta W_{t_i}^A, \tag{4.18}$$

where each μ denotes the drift and each σ denotes the volatility of the according process. We choose the particular price models according to our business cases in the corresponding Section 8.

4.6 Bankruptcy constraints

Naturally, an optimal allocation strategy must consider the case of bankruptcy of the company. We say a company is bankrupt when the outstanding credits are equal or higher than the value of the owned assets; i.e., the net equity[18] is equal or less than zero. Thus, the following inequality must hold for all $i = 0, ..., N^T$ to avoid bankruptcy:

$$B_{t_i} - \text{cost}_{t_i}^{\text{alloc}}(0, a_{t_i}, p_{t_i}^A) > 0. \tag{4.19}$$

In case the net equity reaches zero, our exit strategy is to sell all production assets immediately. This strategy prevents unbounded losses in the case of bankruptcy.

[18] Which means Eigenkapital in German.

4.7 The complete optimization problem

Let us summarize all aspects of the above paragraphs in order to get a clear view of the discrete-time model of optimal portfolio allocation for general production assets.

Problem 1 (Discrete-time portfolio allocation of production assets) *Starting with an initial portfolio* (a_0, B_0)*, our goal is to find the value function* V *and the optimal allocation strategies*[19] $\chi^*, \tilde{a}^*, \tilde{b}^*$ *that maximizes:*

$$V(a_0, B_0) = \sup_{\chi, \tilde{a}, \tilde{b}} J(a_0, B_0, \chi, \tilde{a}, \tilde{b}). \tag{4.20}$$

The objective function is given by (4.4):

$$J(a_0, B_0, \chi, \tilde{a}, \tilde{b}) = E^{\chi}\left[\sum_{i=0}^{N^T} e^{-\delta t_i} U_1(\chi_{t_i}) + e^{-\delta T} U_2(a_T p_T + B_T)\middle| \mathrm{F}_0\right], \tag{4.21}$$

and subject to the state dynamics (4.2), (4.3):

$$\begin{aligned} a_{t_{i+1}} &= a_{t_i} + \tilde{a}_{t_i} - \tilde{b}_{t_i} - r^{det} a_{t_i}, \\ B_{t_{i+1}} &= B_{t_i} + r B_{t_i} - \chi_{t_i} + profit_{t_i}(a_{t_i}, \tilde{a}_{t_i}, \tilde{b}_{t_i}, p^A_{t_i}, p^P_{t_i}, p^C_{t_i}), \end{aligned} \tag{4.22}$$

and price dynamics (4.17), (4.18):

$$\begin{aligned} dp^P_t &= \mu^P(t, p^P_t)dt + \sigma^P(t, p^P_t)dW^P_t, \\ dp^C_t &= \mu^C(t, p^C_t)dt + \sigma^C(t, p^C_t)dW^C_t, \\ \Delta p^A_{t_i} &= \mu^A(t_i, p^A_{t_i})\Delta t + \sigma^A(t_i, p^A_{t_i})\Delta W^A_{t_i}, \end{aligned} \tag{4.23}$$

and the bankruptcy constraint (4.19):

$$B_{t_i} - cost^{alloc}_{t_i}(0, a_{t_i}, p^A_{t_i}) > 0, \tag{4.24}$$

given the initial price values p^P_0*,* p^C_0 *and* p^A_0 *at time* t_0*.*

[19] If they exist.

The portfolio profit is given by (4.9):

$$\begin{aligned} &profit_{t_i}(a_{t_i}, \tilde{a}_{t_i}, \tilde{b}_{t_i}, p^A_{t_i}, p^P_{t_i}, p^C_{t_i}) = \\ &gain_{t_i}(a_{t_i}, p^P_{t_i}, p^C_{t_i}) \; - \; cost^{fix}_{t_i}(a_{t_i}) \; - \; cost^{alloc}_{t_i}(\tilde{a}_{t_i}, \tilde{b}_{t_i}, p^A_{t_i}), \end{aligned} \tag{4.25}$$

and the single components are given by (4.10):

$$\begin{aligned} &gain_{t_i}(a_{t_i}, p^P_{t_i}, p^C_{t_i}) = \\ &\sup_{q \in A(a_{t_i})} E^q_{t_i}\left[\int_{t_i}^{t_{t+1}} \left(p^P_\tau \; QP(q_\tau) - p^C_\tau \; QC(q_\tau) + c^{op}(q_\tau)\right) \; q_\tau \; d\tau \middle| p^P_{t_i}, \, p^C_{t_i} \right], \end{aligned} \tag{4.26}$$

and (4.12):

$$cost^{fix}_{t_i}(a_{t_i}) := \left(c^{main}_{t_i} + c^{inv}_{t_i}\right) \; a_{t_i}, \tag{4.27}$$

and (4.14), (4.15):

$$cost^{alloc}_{t_i}(\tilde{a}_{t_i}, \tilde{b}_{t_i}, p^A_{t_i}) \; = \; \left(\tilde{a}_{t_i} - \tilde{b}_{t_i}\right) \; p^A_{t_i} \; + \; \sum_{j=1}^{N^A} cost^{ill,j}_{t_i}(\tilde{a}^j_{t_i} - \tilde{b}^j_{t_i}), \tag{4.28}$$

where $A(a_{t_i})$ denotes the set of applicable operation policies (4.11):

$$A(a_{t_i}) = \left\{ \{q_\tau, \tau \in [t_i, t_{i+1})\} \mid \; \textit{feasible according to } a_{t_i} \right\}, \tag{4.29}$$

and $cost^{ill,j}_{t_i}$ is smooth, positive, convex w.r.t $\tilde{a}^j_{t_i}$, $-\tilde{b}^j_{t_i}$, and has its global minimum at zero with $cost^{ill,j}_{t_i}(0, 0, p^{A,j}_{t_i}) = 0$.

Note that the gain function in (4.26) is an inner optimization problem for each t_i conditional to $p^P_{t_i}$ and $p^C_{t_i}$ and may be subject to physical constraint $A(a_{t_i})$. So our discrete-time optimal portfolio allocation model for general production assets is a bi-level stochastic optimal control problem.

4.8 Possible extensions

The model could be extended in the following way.

- We could allow the utility function to be depend on a measure of market share. A high market share demonstrates power and strength which has a certain utility itself and may be a separate goal for investors, see our discussion in Section 3.5. Thus, a measure of market share $U_m : \mathbb{R}_+^{N^P} \to \mathbb{R}^{N^P}$ could be added to the objective :

$$\tilde{U}_1(\chi_{t_i}, \text{mshare}_{t_i}) = U_1(\chi_{t_i}) + \gamma\, U_m(\nu_{t_i}), \tag{4.30}$$

  where $\gamma \in \mathbb{R}^{N^P}$ denote the power addiction of the company for all N^P products. The market share $\nu_{t_i}^k$ of product $k = 1, ..., N^P$ could be estimated through:

$$\nu_{t_i}^k = E_{t_i}^q \left[\int_{t_i}^{t_{t+1}} \frac{\text{QP}_{t_i}^k q_\tau^k}{d_\tau^k}\, d\tau \right], \quad \forall i = 0, ..., N, \tag{4.31}$$

  where d_τ^k denotes the total market demand.

  In addition to power, higher market share leads to higher profitability due to the economies of scale, bargaining strength and other market advantages. Therefore the profit function could be extended by Equation (3.25) from Section 3.5 in order to account for the market share effect.

- We could introduce an additional supply constraint into the allocation problem such that a certain (or stochastic) peak demand $d_{t_i} \in \mathbb{R}_+$ can be satisfied at all times:

$$\sum_{j=1}^{N^A} a_{t_i}^j \geq d_{t_i}, \quad \forall i = 0, ..., N. \tag{4.32}$$

  This would be necessary when the company is an essential supplier for contractual customers in the market and must provide enough supply.

5 A reduced continuous-time model

Our general discrete-time model (Problem 1) (4.20 - 4.29) of Section 4 provides a template for optimal portfolio allocation. It can be utilized by companies from various industry sectors to estimate their portfolio value of production assets and their future investment strategy.

On the one hand, providing of such a template is one aim of our thesis and a case analysis for a specific company can be performed if the necessary data is available. On the other hand, we are more interested in a quantitative analysis of the model. By quantitative analysis we mean the quantitative understanding of the model's dynamics and its dependence on independent variables and fixed parameters.

In order to get this quantitative understanding, we reduce our discrete-time model to a simpler continuous-time model (Problem 2) (5.13 - 5.16) in this section. In particular and in contrast to other authors, we start from first principles and use only a few transformations to derive an elaborate model for optimal portfolio allocation of commodity related assets that is similar to known investment-consumption models in literature.

We perform a comparative analysis of the known investment-consumption models in Section 6 and analyze our derived model in Section 7.

5.1 Complexity of the problem

The reasons for the model reduction to a simpler continuous-time model are the following.

- In discrete-time we can not use mathematical tools from continuous analysis or utilize derivatives, which are desirable tools for optimization problems.
- Even in continuous-time, it is rather difficult to find an analytical solution so we need to perform numerical approximations. But our model suffers from the so-called *curse of dimensionality* in no less than three ways:
    1. Each state variable a, B, p^A, p^P, p^C adds N^A, 1, N^A, N^P, N^C dimensions to the problem, respectively. When it comes to **numerical sampling**[20], there is an exponential increase in volume

[20] Discretization of a space onto sample points.

associated with adding extra dimensions to a mathematical space.

2. The prices p_t^A, p_t^P, p_t^C are stochastic variables and **stochastic simulations** like Monte Carlo converge with $\frac{1}{\sqrt{M}}$, where M is the number of sample points in each dimension. It means that quadrupling the number of sampled points in each dimension halves the error.
3. To solve a **dynamic optimization** problems by numerical backward induction, the objective function must be computed for each combination of values. Moreover, for each single optimization, the **dimension** of the control space **increases proportionally** to the number of assets N^A. This is a significant obstacle when the dimension of the state variable is large.

To show the exponential increase of the problem's complexity w.r.t. dimensions let us consider the simplest case for which a portfolio analysis makes sense. Imagine only two different types of production assets $N^A = 2$ with one input and one output commodity each $N^P = N^C = 2$. In this case, the problem has nine space dimensions, six random variables, five control dimensions and must be calculated for each point in the time dimension. In a numeric calculation with 100 sample points for each state, 100 simulations for each random variable and 100 test points for each control the problem ends up with 10^{42} evaluations.

- In addition to the above optimization, the problem includes a nested optimization for the operation scheduling at for each evaluation point, due to the real option approach (4.26).

Problem 1 is a high-dimensional, bi-level discrete-time stochastic optimal control problem. In order to perform a quantitative analysis we reduce it to a simpler continuous-time problem in the next sections.

5.2 Model reduction

In the next paragraphs we make the following simplifications that tackle the three difficulties from above:

- We switch from discrete-time to continuous-time, which is a common approach for financial market models in the literature and in practice.
- We make a change of variables by considering the monetary amount of assets $A_t := (a_t \cdot p_t^A)$ and state the model in relative terms w.r.t the assets prices. Thereby we reduce the number of state variables for each amount-price pair from two to one.

- We summarize the buying rate $\tilde{a}$ and the selling rate $\tilde{b}$ into a total investment rate $\alpha = (\tilde{a} - \tilde{b})$. Thereby we reduce the number of control variables for each production asset from two to one.

- We replace the real option gain$_t(a_t, p_t^P, p_t^C)$ in (4.26) by a stochastic variable R_t which models the relative return from operation. In particular, we summarize the input and output commodities - while respecting the production matrices - to one state variable (their relative production spread). Thereby we reduce the number of state variables and stochastic variables for each input-output pair form two to one and simplify the nested optimization.

Beside the above simplifications we specify the dynamics of the assets prices p_t^A in (4.23) and the dynamics of the return R_t which replaces the dynamics for the commodity prices p_t^P, p_t^C in (4.23).

5.2.1 Basic definitions for the continuous-time model

Throughout the following let $\{W_t, 0 \le t \le T\}$ denotes a d-dimensional, standard $\mathbb{F}$-Brownian motions over $[0,T]$ where $\mathbb{F}$ denotes the naturally generated filtration on the probability space $(\Omega, \mathcal{F}, P)$. Furthermore, let $L^2([0,T], \mathbb{R}^n)$ denote the set of all Lebesgue measurable functions $\varphi : [0,T] \to \mathbb{R}^n$ such that $\int_0^T |\varphi(t)|^2 dt < \infty$ and let $L^2_{\mathbb{F}}([0,T], \mathbb{R}^n)$ denote the set of all $\mathbb{F}$-adapted processes $X : [0,T] \times \Omega \to \mathbb{R}^n$ such that $E \int_0^T |X_t|^2 dt < \infty$.[21]

Furthermore let $n := N^A$ and let $(x \cdot y) : \mathbb{R}^{n \times d} \times \mathbb{R}^n \to \mathbb{R}^{n \times d}$ denote a row-wise multiplication of matrix $x \in \mathbb{R}^{n \times d}$ and vector $y \in \mathbb{R}^n$, i.e., a component wise multiplication in case $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^n$ are both vectors.

From now on we switch all discrete-time state variables a_{t_i}, B_{t_i}, $p_{t_i}^A$ and control variables $\tilde{a}_{t_i}$, $\tilde{b}_{t_i}$, χ_{t_i} of the original model - listed in Section 4.2.3 - to continuous-time variables a_t, B_t, p_t^A, $\tilde{a}_t$, $\tilde{b}_t$ and χ_t.

5.2.2 A change of variables

Let us specify the dynamics of the asset prices p_t^A and assume that they are given by the following n-dimensional stochastic process:

$$dp_t^A = (\mu^A \cdot p_t^A)\, dt + (\sigma \cdot p_t^A)\, dW_t, \quad p_0^A = p_0,$$

[21] An introduction into probability theory can be found in the appendix A.1.

where $\mu^A \in \mathbb{R}^n_+$, $p_0 \in \mathbb{R}^n_+$ and $\sigma \in \mathbb{R}^{n\times d}$. Note that this continuous-time SDE is in compliance with our discrete-time price process (4.23) of our original model. Note also that $p^A_t > 0$ for all t.

Now, we make a change of variables by defining $A_t := (a_t \cdot p^A_t)$ as the portfolio owner's monetary amount of assets in continuous-time. According to equation (4.22) the monetary amount of assets evolves by:

$$\begin{aligned} dA_t &= (da_t \cdot p^A_t) \; + \; (dp^A_t \cdot a_t) \\ &= \big((\tilde{a}_t \cdot p^A_t) - (\tilde{b}_t \cdot p^A_t) - (r^{\det} \cdot A_t)\big)\, dt \; + \; (\mu^A \cdot A_t)dt + (\sigma \cdot A_t)\, dW_t \\ &= \big(((\tilde{a}_t - \tilde{b}_t) \cdot p^A_t) + ((\mu^A - r^{\det}) \cdot A_t)\big)\, dt \; + \; (\sigma \cdot A_t)\, dW_t. \end{aligned}$$

5.2.3 Investment rate and deterioration adjusted drift

Since we omit linear transaction costs, our Problem 1 depends not on the single buying and selling rates rather on the total investment rate $\tilde{a}_t - \tilde{b}_t$. Thus we can replace the buying and selling rates by the total monetary amount of investment per time $\alpha_t := (\tilde{a}_t - \tilde{b}_t)p^A_t$. If $\alpha_t > 0$ the company buys assets, if $\alpha_t < 0$ the company sells assets. We also define the deterioration adjusted drift of the assets value by $\mu := \mu^A - r^{\det} \in \mathbb{R}^n$.

Then the monetary amount of assets A_t evolves according to the following forward stochastic differential equation (SDE) for all $t \in [0, T]$:

$$dA_t \; = \; \big((\mu \cdot A_t) + \alpha_t\big)\, dt \; + \; (\sigma \cdot A_t)\, dW_t, \quad A_0 = a, \tag{5.1}$$

where $a := (a_0 \cdot p^A_0) \in \mathbb{R}^n_+$ denotes the initial value of A_0. We may refer to $\tilde{b}_A(t, A_t, \alpha_t, \beta_t) := ((\mu \cdot A_t) + \alpha_t)$ as the drift term and to $\tilde{\sigma}_A(t, A_t) := (\sigma \cdot A_t)$ as the diffusion term of A_t.

Note that the adjusted drift μ may be negative when the assets's value decreases due to depreciation.[22]

5.2.4 Illiquidity costs for investments

Furthermore, let the function $f_t : \mathbb{R}^n \to \mathbb{R}_+$ denote the illiquidity costs for investments α_t at time t. The function f_t is equal to the term $\text{cost}^{\text{ill}}_{t_i}$ in (4.28) but we assume that the costs depend on the absolute amount of investment

[22] According to our Definition 4.2, the term allocation could mean a scheduled maintenance, in order to prevent depreciation.

α_t instead the rates $\tilde{a}_t - \tilde{b}_t$.

According to the required properties of each component $\text{cost}_t^{\text{ill},j}$, see Section 4.4.3, the components f_t^j must be positive, convex, $f_t^j(\cdot) \in C^2(\mathbb{R}, \mathbb{R}_+)$ and f_t^j has its global minimum at zero with $f_t^j(0) = 0$ for all $j = 1, ..., n$.

Now let us approximate f_t^j by its second order Taylor expansion around zero:

$$f_t^j(\alpha^j) = c_{0t}^j \; + \; c_{1t}^j \, \alpha^j \; + \; c_{2t}^j \, \alpha^j \alpha^j.$$

where $c_{0t}^j = f_t^j(0)$, $c_{1t}^j = \partial_{\alpha^j} f_t^j(0)$ and $c_{2t}^j = \frac{1}{2} \partial_{\alpha^j}^2 f_t^j(0)$. In order to satisfy the above properties, $c_{0t}^j = c_{1t}^j = 0$ and $c_{2t}^j \geq 0$. Thus we end up with a quadratic illiquidity costs function, given by:

$$f(\alpha_t) = c_{2t}^\top (\alpha_t \cdot \alpha_t), \quad c_{2t} \geq 0. \tag{5.2}$$

5.2.5 Fixed costs per asset value

Let $M_t \in \mathbb{R}^n$ denote the fixed costs per asset value for holding the corresponding assets at time t. M_t replaces the fixed costs $c_t^{\text{main}} + c_t^{\text{inv}}$ in (4.27), divided[23] by the assets value p_t^A. In this reduced model, we assume that M_t is deterministic, which means that the fixed costs exactly follows the asset's price dynamics.

Then, according to (4.22), the bank account B_t evolves in continuous-time by:

$$\begin{aligned} dB_t = \; & (r B_t - \chi_t + \text{gain}(a_t, p_t^P, p_t^C) \\ & - M_t^\top A_t - 1^\top \alpha_t - c_{2t}^\top (\alpha_t \cdot \alpha_t)) \, dt, \quad B_0 = b, \end{aligned}$$

where b is the initial value in the portfolio and $1\mathbb{R}^n$ denotes the identity vector.

5.2.6 Operation gain per asset value

Let $R_t \in \mathbb{R}^n$ denote the operation gain (return) rates per asset value. Introducing the return rate R_t, we replace the operation's input-output spreads per [op] in equation (4.26), divided by the asset values p_t^A per [op], by a single stochastic variable:

$$R_t \; = \; \frac{p_\tau^{P\top} \text{QP}(q_\tau) - p_\tau^{C\top} \text{QC}(q_\tau) + c^{\text{op}}}{p_t^A}. \tag{5.3}$$

[23] Componentwise.

Since the company optimizes its asset's operations, the gain function in (4.26) becomes:

$$\text{gain}_{t_i}(a_{t_i}, p^P_{t_i}, p^C_{t_i}) = \max(R_t, 0)^\top A_t.$$

Compared to the real option approach in (4.26), the only simplifications we made are:

1. the set of applicable operation policies $A(a_{t_i})$ in (4.29) is only restricted by the capacity:

$$A(a_{t_i}) = \{\, \{q_\tau, \tau \in [t_i, t_{i+1})\} \mid q_t \le a_t \} \,. \tag{5.4}$$

   This means that the production assets are fully flexible.

2. the operation spread is increasing with the operation level q. Then each production asset is shut down instantaneously whenever its operation spread is negative and driven at full capacity otherwise.

Examples for R_t in practice are the so called crack spread for cracking oil or spark spread for power production, see Section 1.1.1 in the preliminaries. These and other spreads are frequently traded through *exchange options* at commodity markets, see Section 1.3.2.

Modeling the return rates

In Section 1.2.1 we stated that commodities prices usually show a mean-reversion behavior and so do their spreads. Moreover, since the return rate of operation reflect the operation gain that must compensate the operation costs, a reversion of the return to a mean level is even more natural. Therefore, we use Vasicek's model for the return rates which we introduced in the preliminaries Section 1.2.3.

Let $\kappa \in \mathbb{R}^n_+$ denote the speed of reversion and let $\bar{R} \in \mathbb{R}^n$ denote the return's mean value. Then R_t evolves according to the following forward SDE with the initial return $c \in \mathbb{R}^n$:

$$dR_t \;=\; \left(\kappa \cdot (\bar{R} - R_t)\right) dt \;+\; \sigma_R \, dW^R_t, \quad R_0 = c, \tag{5.5}$$

where $\sigma_R \in \mathbb{R}^{n\times d}$ is constant. Here, $\tilde{b}_R(t, R_t) = \kappa \cdot (\bar{R} - R_t)$ is the drift term and $\tilde{\sigma}_R(t, R_t) = \sigma_R$ is the diffusion term. Note that R_t can attain negative values.

Since the return rate R_t depends on the asset's price p^A_t in (5.3), we assume in the reduced model that each asset's value A^i_t is correlated to its own spreads

R_t^i. The cross-correlations are negligible. Therefore we use the diagonal matrix ρ to denote the correlations between the Brownian-motions:

$$dW_t dW_t^R = \rho dt, \tag{5.6}$$

where $\rho^{ii} \in [-1, 1]$ and $\rho^{ij} = 0$, for all $i \neq j$.

Summarizing the above statements, the dynamics of the bank account B_t become:

$$\begin{aligned} dB_t = \quad & (rB_t + \max(R_t, 0)^\top A_t - M_t^\top A_t \\ & -\chi_t - 1^\top \alpha_t - c_{2t}^\top (\alpha_t \cdot \alpha_t))\, dt, \quad B_0 = b. \end{aligned} \tag{5.7}$$

Comment

Throughout this thesis, we refer to the term $\max(R_t, 0) - M$ as **optimized returns** in order to indicate its origin, namely the optimization of flexible operation. However note that, considering the term regardless of its origin, the name **optional** return would be more suitable. In particular, the term $\max(R_t, 0) - M$ is the payoff of a standard option with underlying $R_t - M$ and strike $-M$. Moreover, we will see in our numerical results in Section 7.5, that the behavior of the optimized retun model shows some typical features of options.

5.2.7 Constraints in continuous-time

Now we transfer the constraints in Section 4.6 from the discrete-time model to the continuous-time model.

No short-selling of production assets

Short-selling of production assets is unrealistic, see (4.3). The following constraint prohibits the holding of negative amounts of assets in the portfolio:

$$A_t \in \mathbb{R}_+^n, \quad \forall t \in [0, T]. \tag{5.8}$$

Note that there exists always a control policy $\{\alpha_t, 0 \leq t \leq T\}$ such that (5.8) is satisfied. If, for example, we set $\alpha_t = 0$, the process A_t is a geometric Brownian motion which will not approach negative values by itself. Note also that the bank account B_t can take negative values. This feature allows the portfolio holder to assume credit.

Limited allocation and consumption rate

Furthermore, we assume that the controls (α_t, χ_t) must stay in a bounded (convex and compact) cube:

$$\mathcal{A} = [\alpha_{min}, \alpha_{max}] \times [0, \chi_{max}] \subset \mathbb{R}^n \times \mathbb{R}^1_+. \tag{5.9}$$

It means that the allocation and consumption rate are limited.

Bankruptcy constraint

All admissible control policies must satisfy the above state constraints (5.8). Additionally, the portfolio must satisfy a bankruptcy constraint as in (4.24). Imagine that a company declares bankruptcy at time τ and sells all its assets A_τ at an appropriate rate $0 < \alpha_{bry} \leq \alpha_{max}$ to prevent unpredictable further losses. The selling would persist during the time period $\frac{A_\tau}{\alpha_{bry}}$. Assuming the assets price A_τ stays fixed over this sellout period, the earnings from selling would be:

$$A_\tau - f(\alpha_{bry})\frac{A_\tau}{\alpha_{bry}} \;=\; (1 - (c_{2t} \cdot \alpha_{bry}))\, A_t.$$

Following [Davis, Norman 1990], let us define a closed solvency region by:

$$\mathcal{L}_t(\alpha_{bry}) = \left\{ (x, y) \in \mathbb{R}^n_+ \times \mathbb{R}^1 \,\middle|\, y + \sum_{j=1}^{n} \left(1 - (c^j_{2t}\, \alpha^j_{bry})\right) x^j \geq 0 \right\}. \tag{5.10}$$

Whenever the state process (A_t, B_t) leaves the solvency region $\mathcal{L}_t(\alpha_{bry})$ the company is forced to sell its assets at the rate $\alpha_{bry} \in \mathbb{R}^n_+$ and stop consumption $\chi_t = 0$ for $t \in [\tau, T]$. We call this an exit policy.

Note that it is impossible to prevent losses since it is unrealistic to sell all assets instantaneously with an infinite selling rate as [Davis, Norman 1990] suggested. But we believe the the solvency region $\mathcal{L}_t(\alpha_{bry})$ defines a good indicator when to start an exit strategy and prevent uncertain further losses after the sellout period $\frac{A_\tau}{\alpha_{bry}}$. Note, the higher α_{bry}, the smaller the solvency region $\mathcal{L}_t(\alpha_{bry})$. In other words, the shorter the sellout period should be, the 'earlier' the company has to enter the exit policy.

5.2.8 Admissible control policies

Let us denote the bankruptcy time (time entering the exit strategy) $\tau_{bry} = \inf\{t \geq 0 | \, (A_t, B_t) \notin \mathcal{L}_t(\alpha_{bry})\}$. Regarding the above constraints, we define the set of all admissible control policy for investment-consumption up to time T.

Definition 5.1 (Set of admissible control policies)
The set of admissible control policies is defined by:

$$\mathbb{A}_{0,T} = \{(\alpha_t, \chi_t),\ 0 \le t \le T\}, \tag{5.11}$$

such that:

1. (α_t, χ_t) *is* F_t*-adapted,*
2. $(\alpha_t(\omega), \chi_t(\omega)) \in \mathcal{A}$ *for all* $\omega \in \Omega$*, see (5.9),*
3. $A_t \ge 0$ *almost surely,*
4. $$(\alpha_t, \chi_t) = \begin{cases} (-\alpha_{bry}, 0), & \text{if } (A_t, B_t) \notin \mathcal{L}_t(\alpha_{bry}) \wedge A_t > 0, \\ (0,0), & \text{if } (A_t, B_t) \notin \mathcal{L}_t(\alpha_{bry}) \wedge A_t = 0, \end{cases} \quad \forall t \in [\tau_{bry}, T].$$

Starting with an endowment $(a, b) \in \mathcal{L}_t(\alpha_{bry})$ at time $t = 0$, it is easy to show that $\mathbb{A}_{0,T}$ is non-empty, because $\mathbb{A}_{0,\tau_{bry}}$ is non-empty and we use the exit policy for all $t \in [\tau_{bry}, T]$.

5.2.9 Objective function in continuous-time

The portfolio owner's goal is to maximize his expected utility from consumption over a given time horizon $[0, T]$. Thus, let $u_1 \in C^2(\mathbb{R}_+, \mathbb{R})$ and $u_2 \in C^2(\mathbb{R}, \mathbb{R})$ denote his concave utility functions for consumption and terminal wealth, respectively. According to (4.21) the objective function J in continuous-time is given by:

$$J(t, a, b, c, \alpha, \chi) = E_t^{\alpha,\chi}\left[\int_t^T e^{-\delta s} u_1(\chi_s) ds + e^{-\delta T} u_2(A_T + B_T) \,\middle|\, a, b, c\right], \tag{5.12}$$

where $\delta \ge 0$ measure the owner's discount for consumption and $E_t^{\alpha,\chi}[\,\cdot\,|a, b, c]$ denotes the expectation at time t given the initial endowment $A_t = a$, $B_t = b$, $R_t = c$ and using the control policy (α, χ).

5.3 An investment-consumption model with quadratic transaction costs and optimized stochastic returns

Let us summarize the reduced continuous-time model for optimal portfolio allocation of commodity related assets.

Problem 2 (Continuous-time portfolio alloc. of production assets) *We search for the optimal control policy* $(\alpha^*, \chi^*) \in \mathbb{A}_{0,T}$*, (5.11), that maximizes the objective function (5.12):*

$$J(0,a,b,c,\alpha,\chi) = E_0^{\alpha,\chi}\left[\int_0^T e^{-\delta t}u_1(\chi_t)\,dt + e^{-\delta T}u_2(A_T+B_T)\,\middle|\, a,b,c\right], \tag{5.13}$$

under the state dynamics (5.1), (5.5), (5.6) and (5.7):

$$\begin{aligned} dA_t &= \big((\mu\cdot A_t)+\alpha_t\big)\,dt \;+\; (\sigma\cdot A_t)\,dW_t, && A_0 = a,\\ dR_t &= \big(\kappa\cdot(\bar{R}-R_t)\big)\,dt \;+\; \sigma_R\,dW_t^R && R_0 = c,\\ dB_t &= \big(rB_t + \max(R_t,0)^\top A_t - M_t^\top A_t \\ &\quad -\chi_t - 1^\top\alpha_t - c_{2t}^\top(\alpha\cdot\alpha_t)\big)\,dt, && B_0 = b, \end{aligned} \tag{5.14}$$

where $u_1 \in C^2(\mathbb{R}_+,\mathbb{R})$ *and* $u_2 \in C^2(\mathbb{R},\mathbb{R})$ *are concave functions and where:*

$$dW_t\,dW_t^R \;=\; \rho\,dt. \tag{5.15}$$

We define the value function v *for the continuous-time model for all* $t \in [0,T]$ *by:*

$$v(t,a,b,c) = \max_{(\alpha,\chi)\in\mathbb{A}_{t,T}} J(t,a,b,c,\alpha,\chi). \tag{5.16}$$

It turns out that our continuous-time model for optimal portfolio allocation of commodity related assets is an investment-consumption problem (see Section 3.1.5) with convex transaction costs and stochastic returns[24]. The main distinctions between our production asset portfolio and a standard financial equity portfolio are:

1. the assets yield a stochastic, mean-reverting return rate R_t,
2. the returns are optimized by $\max(R_t,0)$ while there is a deterministic holding fee M_t,
3. the allocations yield quadratic transaction costs.

Note again that we derived our model for optimal portfolio allocation of commodity related assets from first principles and used only a few transformations to turn it into an investment-consumption model.

By considering the reduced Problem 2 (5.13 - 5.16), we can compare it to other investment-consumption models in Section 6 and then perform a quantitative analysis in Section 7.

[24] In the literature the returns are sometimes called dividends or income too.

5.3.1 Existence of unique strong solutions

Since Problem 2 is the model we deal with in the following, we briefly show that the objective function (5.13) is well-defined (bounded) and that there exists solutions for the forward SDEs (5.14), (5.15). The existence of solutions for the stochastic optimal control Problem 2, (5.13) - (5.16), is handled in more detail in in Part III of the thesis.

Recall that the time interval $[0,T]$ is fixed and finite and that the values α_t and χ_t of an admissible control policy are bounded. Now let $\pi \in \mathbb{A}_{0,T}$, $\pi_t = (\alpha_t, \chi_t)$, be an arbitrary but fixed admissible control policy. Then we have the following properties:

- The forward SDEs (5.14) are of Markovian type, see Appendix A.10.
- The drift term $\tilde{b}_A(t, A_t, \pi_t)$ and the diffusion term $\tilde{\sigma}_A(t, A_t)$ of the process A_t in (5.14) are Lipschitz continuous w.r.t A_t. Moreover, since $\pi_t \in A$ is bounded for all $t \in [0,T]$, it follows that $|\tilde{b}_A(t,0,\pi_t)|+|\tilde{\sigma}_A(t,0)| \in L^2([0,T],\mathbb{R}^n)$.
- The drift term $\tilde{b}_R(t, R_t)$ and the diffusion term $\tilde{\sigma}_R(t, R_t)$ of the process R_t in (5.14) are Lipschitz continuous w.r.t R_t (the diffusion term σ_R is constant). Moreover, $|b_R(t,0)| + |\sigma_R| \in L^2([0,T],\mathbb{R}^n)$.
- The Ornstein-Uhlenbeck process R_t is asymptotically bounded almost surely, see appendix A.13. Since the drift term has at most linear growth and the diffusion term is constant, R_t is bounded on the finite time interval $[0,T]$ almost surely. Moreover, R_t is independent of A_t, B_t and π_t.

  Thus the drift term $\tilde{b}_B(t, A_t, B_t, R_t, \pi_t)$ of the process B_t in (5.14) is Lipschitz continuous w.r.t A_t and B_t. We also obtain that the drift $|b_B(t,0,0,0,\pi_t)| \in L^2([0,T],\mathbb{R}^n)$.

Theorem 5.2 (Existence of SDEs and the objective function)
Let $\mathcal{X}_t := (A_t, B_t, R_t)^\top$ denote the state processes, $x := (a,b,c)^\top$ the initial values, $\pi := (\alpha, \chi)$ the control policies and let (5.14) be written as:

$$d\mathcal{X}_t = \tilde{b}(t, \mathcal{X}_t, \pi_t)dt + \tilde{\sigma}(t, \mathcal{X}_t)dW_t, \quad X_0 = x. \tag{5.17}$$

Then for any fixed $\pi \in \mathbb{A}_{0,T}$, the process $\mathcal{X}_t$ admits a unique strong solution:

$$\mathcal{X}(t) = x + \int_0^t \tilde{b}(x, \mathcal{X}_s, \pi_t)ds + \int_0^t \tilde{\sigma}(s, \mathcal{X}_s)dW_s, \tag{5.18}$$

and the objective function $J(t,x)$ in (5.13) is bounded for all $t \in [0,T]$.

Proof 5.3
We showed that for any fixed $\pi \in \mathbb{A}_{0,T}$ there exists a constant $L_T > 0$ such that:

$$|\tilde{b}(t,x,\pi_t) - \tilde{b}(t,\hat{x},\pi_t)| + |\tilde{\sigma}(t,x) - \tilde{\sigma}(t,\hat{x})| \leq L_T|x-\hat{x}|,$$

$$\forall t \in [0,T],\ x,\hat{x} \in \mathbb{R}^n,$$

$$|\tilde{b}(\cdot,0,\pi_{(\cdot)})| + |\tilde{\sigma}(\cdot,0)| \in L^2([0,T],\mathbb{R}).$$

Then Theorem 10.3 in Section 10 provide the existence of a unique strong solution (5.18).

Moreover, since u_1 is twice continuously differentiable and the control values χ_t are bounded, the integral $\int_0^T e^{-\delta t}u_1(\chi_t)\,dt$ in (5.13) is bounded. Since u_2 is also twice continuously differentiable and A_t, B_t admit strong solutions, the expectation $E_0^\pi\left[\, e^{-\delta T}u_2(A_T+B_T) \,|\, a,b,c\right] < \infty$. Thus the objective function is bounded, see Theorem 10.2 in Section 10.

6 Analytical solutions for simplified cases

Before we consider numerical solutions for Problem 2 (5.13) - (5.16), we perform a comparative analysis of special cases for which there exists results in the literature or for which we can find analytical solutions. The purpose of analyzing special cases before deploying numerical schemes for the general case is threefold:

1. Analytical solutions of simple cases reveal the basic behavior of the original problem's solution and some of its properties without numerical approximations.

2. They bring forth whether the original problem may have a singular control or a bang-bang control, or whether there are discontinuities in the value function. Therefore, the knowledge of analytical solutions of simple cases can assist in the choice of the numerical algorithm for the complex problem.

3. Analytical solutions of simple cases provide test cases which can be easily used to assess numerical implementations. One needs analytical solutions to measure the error to the exact solution and to estimate convergence.

In Section 6.1 and 6.3 we present known analytical solutions of Problem 2 for the standard Merton case and the proportional transacton cost case, respectively. In Section 6.2 we derive a new analytical solution for the case of stochastic returns and in Section 6.4 we explore the case of quadratic transaction costs.

6.1 Merton's investment-consumption model

We start with [Merton 1969]'s original investment-consumption model. It is the simplest form of our continuous-time model for optimal portfolio allocation of commodity related assets, Problem 2, disregarding the transaction costs and the stochastic returns.

6.1.1 The problem

So let us consider the 1-dimensional case $n = d = 1$. Let the return rate be constant, $\kappa = \sigma_R = 0$, and let the transaction costs be zero, $c_{2t} = 0$. Also, let the control space $\mathcal{A}$ be unbounded. Then we define the new drift $\tilde{\mu} = \mu + c - M$.

Since the utility function in our model is of HARA type, see Section 3.1.3, we consider the following three examples here:

- log utility $u(x) := \log(x)$, having constant relative risk aversion $\lambda_2 = 1$,
- power utility $u(x) = \frac{1}{\gamma}x^\gamma$ with $\gamma < 1$ and $\gamma \neq 0$, having constant relative risk aversion $\lambda_2 = 1$,
- negative exponential utility $u(x) = \frac{1-e^{-\gamma X}}{\gamma}$ with $\gamma > 0$, having constant absolute risk aversion $\lambda_1 = 1$.

In order to obtain analytical solutions, we assume that the well-posedness conditions are satisfied:

$$\begin{aligned} &\delta > 0, && \text{for log utility,} \\ &\delta > \gamma[r + (\tilde{\mu} - r)^2/\sigma^2(1-\gamma)], && \text{for power utility,} \\ &\delta > 0, && \text{for negative exponential utility.} \end{aligned} \tag{6.1}$$

6.1.2 Finding solutions

Following [Merton 1969], we define the total wealth process $X_t := A_t + B_t$, the proportion of wealth invested in assets $\pi := A_t/X_t$ and the proportional consumption $\tilde{\chi}_t = \chi_t/X_t$. Using this change of variables, (5.14) becomes:

$$dX_t = ((\tilde{\mu} - r)\pi_t + (r - \tilde{\chi}_t))X_t dt + \sigma\pi_t X_t dW_t, \quad X_0 = x. \tag{6.2}$$

Note that the allocation rate α_t disappear and the new control is π_t. Note also that the dimension is reduced from two (A_t, B_t) to one X_t dimension. The objective function (5.13) becomes:

$$J(t, x, \pi, \chi) = E_t^\pi \left[\int_t^T e^{-\delta t} u(\tilde{\chi}_t X_t) dt + e^{-\delta T} u(X_T) \Big| X_t = x \right], \tag{6.3}$$

and the value function (5.16) becomes:

$$v(t, x) = \max_{\pi,\chi} J(t, x, \pi, \chi). \tag{6.4}$$

[Merton 1971] constructed optimal control policies π^*, χ^* and a general analytical solution $v(t, x) \in C^{1,2}([0, T] \times \mathbb{R}_+, \mathbb{R})$ for the class of HARA utility functions. He used the well-known Hamilton-Jacobi-Bellman (HJB) theory and the fact that the value function v is homogeneous in x:

$$\begin{aligned} &v(t, p\,x) = \tfrac{1}{\delta}\log(p) + v(t, x), && \text{for } u(\chi) = \log(\chi), && \forall p > 0, \\ &v(t, p\,x) = p^\gamma v(t, x), && \text{for } u(\chi) = \tfrac{\chi^\gamma}{\gamma}, && \forall p > 0, \\ &v(t, p + x) = e^{-d(t)p}\, v(t, x), && \text{for } u(\chi) = \tfrac{e^{-\gamma\chi}}{\gamma}, && \forall p > 0, \end{aligned} \tag{6.5}$$

where $d(t)$ is a function only depending on time and known parameters. The HJB equation of the investment-consumption problem is given by:

$$\begin{aligned} -\partial_t v(t,x) = &\sup_{c\in\mathbb{R}_+}\left[e^{-\delta t}u(cx) + (r-c)x\,\partial_x v\right] \\ &+\sup_{\pi\in\mathbb{R}}\left[(\mu-r)\pi x\,\partial_x v + \frac{\sigma^2}{2}\pi^2 x^2\,\partial_{xx} v\right], \qquad (6.6) \\ v(T,x) = &\; e^{-\delta t}u(x). \end{aligned}$$

Now let us compare the solutions of the three examples: log, power and negative exponential utility.

6.1.3 Analytical solutions

We define:

$$K := (r-\delta) + \frac{(\tilde{\mu}-r)^2}{2\sigma^2},$$

$$\bar{\delta} := \frac{1}{1-\gamma}\left(\delta - \gamma\left(\frac{1}{2}\frac{(\tilde{\mu}-r)^2}{(1-\gamma)\sigma^2} + r\right)\right),$$

$$\beta(t) := \frac{1+(\delta-1)e^{-\delta(T-t)}}{\delta}, \quad \bar{\beta}(t) := \frac{1+(\bar{\delta}-1)e^{-\bar{\delta}(T-t)}}{\bar{\delta}}, \quad \beta(T) = \bar{\beta}(T) = 1,$$

$$b(t) := \left(\frac{\delta-1}{\delta}(T-t) - \frac{1}{\delta^2}\right)e^{-\delta T}, \quad b(T) = \frac{1}{\delta^2}e^{-\delta T},$$

$$a := \frac{(\delta-r) + \frac{1}{2}\frac{(\mu-r)^2}{\sigma^2}}{r\gamma},$$

$$d(t) := \frac{\gamma r}{1-(1-r)e^{-r(T-t)}}, \quad d(T) = \gamma,$$

$$\bar{a}(t) := \frac{\gamma}{d(t)}\exp\left(-a\gamma(1-e^{-r(T-t)})\right), \quad a(T) = 1.$$

Table 1 shows the solutions of the three examples for the infinite-time horizon case, $T = \infty$, and the finite-time horizon case, $T < \infty$, in a compact form. We derived the finite-time solutions by hand in consideration of the solution for general HARA utility provided in [Merton 1971]. As an example, we present the derivation of the solution for negative exponential utility in Appendix A.2.

	Infinite-time horizon	Finite-time horizon
Log utility	$v(x) = \frac{1}{\delta}\ln(\delta x) + \frac{1}{\delta^2}K,$ $\pi^* = \frac{\tilde{\mu}-r}{\sigma^2}, \quad \tilde{\chi}^* = \delta.$	$v(t,x) = e^{-\delta t}\left(\beta(t)\ln(\frac{x}{\beta(t)}) + \frac{1}{\delta^2}K\right)$ $+b(t)K,$ $\pi^* = \frac{\tilde{\mu}-r}{\sigma^2}, \quad \tilde{\chi}_t^* = \beta(t)^{-1}.$
Power utility	$v(x) = \frac{1}{\gamma}x^\gamma\bar{\delta}^{\gamma-1},$ $\pi^* = \frac{(\tilde{\mu}-r)}{(1-\gamma)\sigma^2}, \quad \tilde{\chi}^* = \bar{\delta}.$	$v(t,x) = e^{-\delta t}\frac{1}{\gamma}x^\gamma\bar{\beta}(t)^{1-\gamma},$ $\pi^* = \frac{(\tilde{\mu}-r)}{(1-\gamma)\sigma^2}, \quad \tilde{\chi}_t^* = \bar{\beta}(t)^{-1}.$
Neg. expo. utility	$v(x) = \frac{1}{\delta\gamma} - \frac{1}{r\gamma}e^{-a\gamma}e^{-r\gamma x},$ $\pi^* x = \frac{\mu-r}{r\gamma\sigma^2}, \quad \tilde{\chi}^* x = rx + a.$	$v(t,x) = \frac{e^{-\delta t}}{\gamma}\left[\frac{1}{\delta} - \bar{a}(t)e^{-d(t)x}\right]$ $+\frac{1}{\gamma}(1-\frac{1}{\delta})e^{-\delta T},$ $\pi^* x = \frac{\mu-r}{d(t)\sigma^2},$ $\tilde{\chi}_t^* x = \frac{d(t)x}{\gamma} + a(1-e^{-r(T-t)}).$

Table 1: Solutions of Merton's investment-consumption problem for three HARA utility functions and different time horizons.

6.1.4 Conclusions

Constant relative risk aversion

Comparing the **value functions** v of the log and power utility case (CRRA) with infinite-time and finite-time horizon, we notice that the structure of the solutions are the same. Since the latter case is time dependent, the discounting factor $e^{-\delta t}$ appears and δ is replaced by $\beta^{-1}(t)$. Furthermore, an additional time dependent term $b(t)K$ is needed in order to match the specific terminal value $u(T)$ at T. In the power utility case δ is replaced by $\bar{\delta}$. Note that the log utility case is the limit of the power utility case for $\gamma \to 0$.

The **optimal control** π^*, which denotes the proportion of wealth invested in the risky asset, is constant over time. Starting at a non-optimal point (A_t, B_t), $X_t = A_t + B_t$, it is optimal to immediately shift the portfolio such that $A_t^* = \pi^* X_t$, $B_t^* = (1-\pi^*)X_t$. When the optimal point (A_t^*, B_t^*) is reached, the allocation stops. This is called a bang-bang control:

$$\alpha_t^* = \begin{cases} \infty, & \text{if } \frac{A_t}{A_t+B_t} < \pi^*, \\ 0, & \text{if } \frac{A_t}{A_t+B_t} = \pi^*, \\ -\infty, & \text{if } \frac{A_t}{A_t+B_t} > \pi^*. \end{cases} \tag{6.7}$$

Note that the optimal portfolios $(\pi^* X_t, (1-\pi^*)X_t)$ for all $X_t \in \mathbb{R}_+$ is a line in the (A, B) space. For later purpose we call it *Merton's CRRA line* $\bar{\pi}$.

The **optimal consumption** $\chi_t^* X_t$ is proportional to the current wealth X_t in all four cases. In the infinite-time horizon case the rate is constant, in the finite-time horizon case the rate increases with time and reaches 1 at terminal time T.

Constant absolute risk aversion

In case of negative exponential utility (CARA), not the optimal control π_t^* but rather the **optimal monetary amount of wealth invested in the risky asset** $\pi_t^* X_t$ is independent of X_t. In particular, it only depends on time $T-t$ when $T < \infty$. This means that as an investor becomes wealthier, the proportion of his wealth invested in the risky asset falls.

In contrast to CRRA utility, the optimal portfolios $(\pi_t^* X_t, (1-\pi_t^*)X_t) = (A_t^*, X_t - A_t^*)$ for negative exponential utility have a fixed optimal amount of assets A_t^* for all $X_t \in \mathbb{R}_+$. This is a vertical line in the (A, B) space, see Figure 7. In particular, A_t^* only depends on $T-t$ when $T < \infty$. For later purpose we call this optimal point *Merton's CARA portfolio* A_t^*.

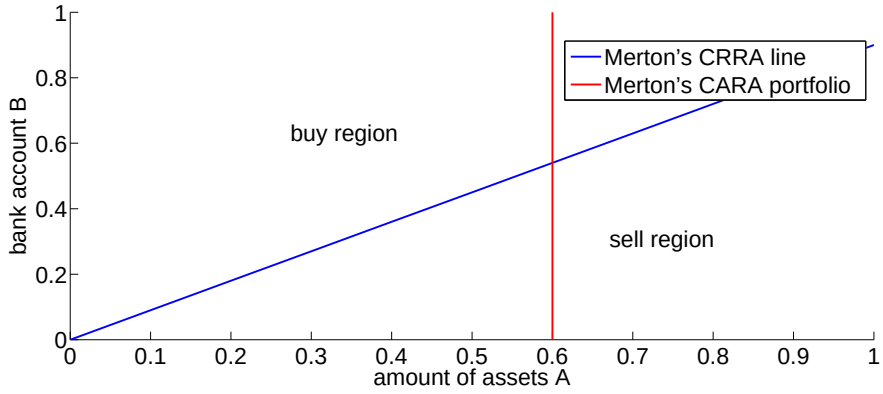

Figure 7: Examples of Merton's CRRA line and Merton's CARA portfolio in the (A, B) space at a fixed time point.

The **optimal consumption** $\chi_t^* X_t$ is no longer proportional to the current wealth X_t but still linear in wealth. In the finite-time horizon case, the additional term decrease with time while the overall rate χ_t^* increases and

reaches 1 at terminal time T. This is similar to the log and power utility cases.

Comment

An outstanding feature of Merton's investment-consumption model is the constant[25] allocation ratio π^* that occurs in all cases. It means a continiously rebalancing of the portfolio to maintain the optimal ratio.

6.2 Investment-consumption model with stochastic returns

Now we consider the investment-consumption Problem 2 with the stochastic return rate R_t but without transaction costs.

6.2.1 The problem

Again, let us consider the 1-dimensional case $n = d = 1$ of Problem 2 and let the transaction costs be zero, $c_{2t} = 0$. But this time, let the mean-reversion speed κ, the mean-reversion level $\bar{R}$, the volatility σ_R and the initial rate z be positive constant.

Since we do not optimize the returns here, we replace $\max(R_t, 0)$ with R_t in (5.14). Also, let the control space $\mathcal{A}$ be unbounded and let the well-posedness condition (6.1) be satisfied. We re-define the drift by $\tilde{\mu} = \mu - M$.

6.2.2 Finding solutions

Using the same change of variables as in the previous subsection, our SDE system (5.14) becomes:

$$\begin{aligned} dX_t &= \big((R_t + \tilde{\mu} - r)\pi_t + (r - \tilde{\chi}_t)\big)X_t dt + \sigma\pi_t X_t dW_t, \qquad X_0 = x, \\ dR_t &= \kappa(\bar{R} - R_t)dt + \sigma_R dW_t^R, \qquad R_0 = z, \end{aligned} \tag{6.8}$$

where $\rho = dW_t dW_t^R$ denotes the correlation between W_t and W_t^R. The value function becomes:

$$v(t, x, z) = \max_{\pi,\chi} J(t, x, z, \pi, \chi). \tag{6.9}$$

[25] Time dependent in case of negative exponential utility.

Note that the SDE for X_t (6.8) is linear w.r.t X_t and thus the value function v is homogeneous in x (6.5), as in the case of no returns in the previous Section 6.1.2. Note also that we use z (instead of c) as the initial value of R_0 here.

In the following paragraphs we investigate finding an analytical solution for the log-utility case. We believe that there exist also analytical solutions for the power and the negative exponential utility cases. The problem is to find a good guess of the value function's shape which could be inserted into the HJB equation.

Infinite-time horizon case

We start with the infinite-time horizon case. This is the easier case because we do not have to deal with time. So let us assume that there exists a smooth value function $v(x,z) \in C^2(\mathbb{R}_+ \times \mathbb{R}, \mathbb{R})$. Using Bellman's principle of optimality, it is well known that v must satisfy the HJB equation:

$$-\delta v(x,z) + \sup_{\pi,\chi \in \mathbb{R}\times\mathbb{R}_+} \left[u(\chi x) + A^{\pi,\chi} v(x,z)\right] = 0, \tag{6.10}$$

where the generator $A^{\pi,\chi}$ is given by:

$$\begin{aligned} A^{\pi,\chi}\, v := \; & \left[(r-\chi) + (z+\tilde{\mu}-r)\pi\right] x\, \partial_x v + \tfrac{\sigma^2}{2}\pi^2 x^2\, \partial_{xx} v \\ & + \kappa(\bar{R}-z)\, \partial_z v + \tfrac{\sigma_R^2}{2}\, \partial_{zz} v + \rho\, \sigma_R \sigma\, \pi x\, \partial_{xz} v. \end{aligned} \tag{6.11}$$

In the log utility case, $u(x) = \log(x)$, we make a change of variables $y := \log(x)$. Then x must be replaced by e^y in the HJB equation (6.20) and the generator becomes:

$$\begin{aligned} A^{\pi,\chi} v := \; & \left[(r-\chi) + (z+\tilde{\mu}-r)\pi - \tfrac{\sigma^2}{2}\pi^2\right] \partial_y v + \tfrac{\sigma^2}{2}\pi^2\, \partial_{yy} v \\ & + \kappa(\bar{R}-z)\, \partial_z v + \tfrac{\sigma_R^2}{2}\, \partial_{zz} v + \rho\, \sigma_R \sigma\, \pi\, \partial_{yz} v. \end{aligned} \tag{6.12}$$

In view of the solutions in Table 1 and by considering the fact that the value function v is homogeneous in x (6.5), we suppose that v is of the following form:

$$\begin{aligned} v(y,z) &= \beta\,(y - \log(\beta)) + f(z), \\ \partial_y v(y,z) &= \beta, \\ \partial_z v(y,z) &= \partial_z f(z). \end{aligned} \tag{6.13}$$

Plugging this Ansatz into the HJB equation, we get:

$$\begin{aligned} 0 = \; & -\delta\beta\,(y - \log(\beta)) - \delta f(z) + \sup_{(\pi,\chi)\in\mathbb{R}\times\mathbb{R}_+} \big[\log(\chi) + y \\ & \Big((r-\chi) + (z+\tilde{\mu}-r)\pi - \tfrac{\sigma^2}{2}\pi^2\Big)\,\beta \\ & + \kappa(\bar{R}-z)\,\partial_z f(z) + \tfrac{\sigma_R^2}{2}\,\partial_{zz} f(z)\Big]. \end{aligned} \tag{6.14}$$

Since this equation must hold for arbitrary y, we determine $\beta = \frac{1}{\delta}$. The optimal controls π^*, χ^* can be found by considering the first derivatives w.r.t. π and χ:

$$\begin{aligned} 0 = \; & \tfrac{1}{\chi} - \tfrac{1}{\delta}, \\ 0 = \; & (z+\tilde{\mu}-r) - \sigma^2\pi. \end{aligned}$$

Since $\chi, \sigma > 0$, the second derivatives are less than zero and the following controls are optimal:

$$\pi^* = \frac{z+\tilde{\mu}-r}{\sigma^2}, \quad \chi^* = \delta. \tag{6.15}$$

Plugging the optimal controls π^* and χ^* back into (6.14), we get the following algebraic equation:

$$0 = \frac{\sigma_R^2}{2}\,\partial_{zz} f(z) + \kappa(\bar{R}-z)\,\partial_z f(z) - \delta f(z) + \frac{(r-\delta)}{\delta} + \frac{(z+\tilde{\mu}-r)^2}{2\delta\sigma^2}. \tag{6.16}$$

The variable z appears only in the first and second power in (6.16). Therefore, we suppose $f(z)$ to be of the following form:

$$\begin{aligned} f(z) = \; & az^2 + bz + c, \\ \partial_z f(z) = \; & 2az + b, \\ \partial_{zz} f(z) = \; & 2a. \end{aligned} \tag{6.17}$$

We plug this Ansatz into (6.16) and sort by the terms of same power:

$$\begin{aligned} 0 = \; & \Big(-\delta a + \tfrac{1}{2\sigma^2\delta} - 2\kappa a\Big)\,z^2 + \Big(2\kappa\bar{R}a - (\delta+1)b + \tfrac{(\tilde{\mu}-r)}{\sigma^2\delta}\Big)\,z \\ & + \sigma_R^2 a + \kappa\bar{R}b - \delta c + \tfrac{(r-\delta)}{\delta} + \tfrac{(\tilde{\mu}-r)^2}{2\delta\sigma^2}. \end{aligned}$$

Since this equation must hold for arbitrary z, we determine a, b, c as:

$$\begin{aligned} a = \; & \tfrac{1}{2\delta(\delta+2\kappa)\sigma^2}, \\ b = \; & \tfrac{2\kappa\bar{R}a}{\delta+1} + \tfrac{(\tilde{\mu}-r)}{\delta(\delta+1)\sigma^2}, \\ c = \; & \tfrac{\sigma_R^2 a}{\delta} + \tfrac{\kappa\bar{R}b}{\delta} + \tfrac{(r-\delta)}{\delta^2} + \tfrac{(\tilde{\mu}-r)^2}{2\delta^2\sigma^2}. \end{aligned} \tag{6.18}$$

Combining (6.13), (6.15), (6.17) together with the definition of (a, b, c) in (6.18), we obtain the following solution of the investment-consumption model with stochastic return:

$$\begin{aligned} v(x,z) &= \tfrac{1}{\delta}\log(\delta x) + az^2 + bz + c, \\ \pi^* &= \tfrac{z+\tilde{\mu}-r}{\sigma^2}, \quad \chi^* = \delta. \end{aligned} \tag{6.19}$$

Finite-time horizon case

Let us assume that there exists a smooth value function $v(t,x,c) \in C^{1,2}([0,T]\times \mathbb{R}_+ \times \mathbb{R}, \mathbb{R})$. Using Bellman's principle of optimality, it is well known that v must satisfy the HJB equation:

$$\begin{aligned} &\partial_t v(t,x,z) + \sup_{(\pi,\chi)\in\mathbb{R}\times\mathbb{R}_+} \left[e^{-\delta t}u(\chi x) + A^{\pi,\chi}\, v(t,x,z)\right] = 0, \\ &v(T,x,z) = e^{-\delta T}u(x), \end{aligned} \tag{6.20}$$

where the generator $A^{\pi,\chi}$ is given by:

$$\begin{aligned} A^{\pi,\chi}v := \; &\left[(r-\chi) + (z+\tilde{\mu}-r)\pi\right] x\,\partial_x v + \tfrac{\sigma^2}{2}\pi^2 x^2\,\partial_{xx} v \\ &+\kappa(\bar{R}-z)\,\partial_z v + \tfrac{\sigma_R^2}{2}\,\partial_{zz} v + \rho\,\sigma_R\sigma\,\pi x\,\partial_{xz} v. \end{aligned} \tag{6.21}$$

In the log utility case $u(x) = \log(x)$ we make a change of variables $y := \log(x)$. Then x must be replaced by e^y in the HJB equation (6.20) and the generator becomes:

$$\begin{aligned} A^{\pi,\chi}v := \; &\left[(r-\chi) + (z+\tilde{\mu}-r)\pi - \tfrac{\sigma^2}{2}\pi^2\right]\partial_y v + \tfrac{\sigma^2}{2}\pi^2\,\partial_{yy} v \\ &+\kappa(\bar{R}-z)\,\partial_z v + \tfrac{\sigma_R^2}{2}\,\partial_{zz} v + \rho\,\sigma_R\sigma\,\pi\,\partial_{yz} v. \end{aligned} \tag{6.22}$$

In view of the solutions in Table 1 and by considering the fact that the value function v is homogeneous in x (6.5), we suppose that v is of the following form:

$$\begin{aligned} &v(t,y,z) = e^{-\delta t}\left[\beta(t)(y - \log(\beta(t))) + f(t,z)\right], \\ &\beta(T) = 1, \quad f(T,z) = 0. \end{aligned} \tag{6.23}$$

Plugging this Ansatz into the HJB equation, we get:

$$
\begin{aligned}
0 = \quad & \dot{\beta}(t)e^{-\delta t}y - \delta\beta(t)e^{-\delta t}y \\
& + \delta e^{-\delta t}\beta(t)\log(\beta(t)) - e^{-\delta t}\dot{\beta}(t)\log(\beta(t)) - e^{-\delta t}\beta(t)\tfrac{\dot{\beta}}{\beta(t)} \\
& + \partial_t f(t,z)e^{-\delta t} - \delta f(t,z)e^{-\delta t} \\
& + \kappa(\bar{R}-z)\,\partial_z f(t,z)e^{-\delta t} + \tfrac{\sigma_R^2}{2}\,\partial_{zz} f(t,z)e^{-\delta t}, \\
& + \sup_\chi \left[\log(\chi e^y) + (r-\chi)\beta(t)\right] e^{-\delta t} \\
& + \sup_\pi \left[(z+\tilde{\mu}-r)\pi - \tfrac{\sigma^2}{2}\pi^2\right] \beta(t)e^{-\delta t}.
\end{aligned}
\tag{6.24}
$$

The optimal controls π_t^*, χ_t^* can be found by considering the first derivatives w.r.t. π and χ:

$$
\begin{aligned}
0 = \quad & \tfrac{1}{\chi} - \beta(t), \\
0 = \quad & (z+\tilde{\mu}-r) - \sigma^2\pi.
\end{aligned}
$$

Since $\chi, \sigma > 0$, the second derivatives are less than zero and the following controls are optimal:

$$
\pi_t^* = \frac{z+\tilde{\mu}-r}{\sigma^2}, \quad \chi_t^* = \frac{1}{\beta(t)}. \tag{6.25}
$$

Plugging the optimal controls π^* and χ^* back into (6.24) and divide by $e^{-\delta t}$, we get:

$$
\begin{aligned}
0 = \quad & \left(\dot{\beta}(t) - \delta\beta(t) + 1\right)\left(y - \log(\beta(t))\right) \\
& - \left(\dot{\beta} - r\beta(t) + 1\right) \\
& + \partial_t f(t,z) - \delta f(t,z) \\
& + \kappa(\bar{R}-z)\,\partial_z f(t,z) + \tfrac{\sigma_R^2}{2}\,\partial_{zz} f(t,z), \\
& + \beta(t)\tfrac{(z+\tilde{\mu}-r)^2}{2\sigma^2}.
\end{aligned}
\tag{6.26}
$$

Since this must be true for arbitrary y, $\beta(t)$ must satisfy the ODE:

$$
\dot{\beta}(t) = \delta\beta(t) - 1, \quad \beta(T) = 1,
$$

with the solution[26]:

$$
\beta(t) := \frac{1 + (\delta-1)e^{-\delta(T-t)}}{\delta}. \tag{6.27}
$$

(6.26) becomes:

$$
\begin{aligned}
0 = \quad & (r-\delta)\beta(t) + \beta(t)\tfrac{(z+\tilde{\mu}-r)^2}{2\sigma^2} \\
& + \partial_t f(t,z) - \delta f(t,z) + \kappa(\bar{R}-z)\,\partial_z f(t,z) + \tfrac{\sigma_R^2}{2}\,\partial_{zz} f(t,z).
\end{aligned}
\tag{6.28}
$$

[26]Same $\beta(t)$ as in Merton's original problem.

The variable z appears only in the first and second power in (6.28). Remembering that $f(T,z) = 0$, we suppose $f(t,z)$ to be of the following form:

$$f(t,z) = a(t)z^2 + b(t)z + c(t), \quad a(T), b(T), c(T) = 0, \tag{6.29}$$

with

$$\begin{aligned} \partial_t f(t,z) &= \dot{a}(t)z^2 + \dot{b}(t)z + \dot{c}(t), \\ \partial_z f(t,z) &= 2a(t)z + b(t), \\ \partial_{zz} f(t,z) &= 2a(t). \end{aligned}$$

We plug this Ansatz into (6.28), compare terms of same power and obtain that $a(t), b(t), c(t)$ must satisfy the ODE system:

$$\begin{aligned} \dot{a}(t) &= (\delta + 2\kappa)\, a(t) - \tfrac{\beta(t)}{2\sigma^2}, && a(T) = 0, \\ \dot{b}(t) &= (\delta + 1)\, b(t) - 2\kappa\bar{R}a(t) - \tfrac{\beta(t)(\tilde{\mu}-r)}{\sigma^2}, && b(T) = 0, \\ \dot{c}(t) &= \delta c(t) - \sigma_R^2 a(t) - \kappa\bar{R}b(t) - \beta(t)\left(\tfrac{(\tilde{\mu}-r)^2}{2\sigma^2} + r - \delta\right), && c(T) = 0. \end{aligned} \tag{6.30}$$

The derivation of the solution of the ODE system is provided in Appendix A.3. Combining (6.23), (6.25), (6.27), (6.29) together with the solutions of the ODE system (6.30) in Appendix (A.22 - A.27) we obtain the following solution of the investment-consumption model with stochastic return:

$$\begin{aligned} v(t,x,z) &= e^{-\delta t}\left[\beta(t)\left(\log(\tfrac{x}{\beta(t)}\right) + a(t)z^2 + b(t)z + c(t)\right], \\ \pi^*(z) &= \tfrac{z+\tilde{\mu}-r}{\sigma^2}, \quad \chi_t^* = \tfrac{1}{\beta(t)}, \end{aligned} \tag{6.31}$$

where:

$$\begin{aligned} a(t) &= \left(-(k_a + h_a)e^{-2\kappa(T-t)} + k_a\right) e^{-\delta(T-t)} + h_a, \\ b(t) &= \left[-(l_b + k_b + h_b)e^{-(T-t)} + l_b e^{-2\kappa(T-t)} + k_b\right] e^{-\delta(T-t)} + h_b, \\ c(t) &= \left(m_c\, e^{-(T-t)} + l_c\, e^{-2\kappa(T-t)} + k_c\,(T-t) - (m_c + l_c + +h_c)\right) e^{-\delta(T-t)} + h_c, \end{aligned} \tag{6.32}$$

and the parameters are defined by:

$$
\begin{aligned}
h_a &:= \tfrac{1}{2\delta(\delta+2\kappa)\sigma^2}, \quad k_a := \tfrac{\delta-1}{2\kappa\, 2\delta\sigma^2}, \\
h_b &:= \tfrac{2\kappa\bar{R}h_a}{\delta+1} + \tfrac{(\tilde{\mu}-r)}{\delta(\delta+1)\sigma^2}, \quad k_b := 2\kappa\bar{R}k_a + \tfrac{(\tilde{\mu}-r)}{\sigma^2}\tfrac{(\delta-1)}{\delta}, \quad l_b := \tfrac{2\kappa\bar{R}(k_a+h_a)}{2\kappa-1}, \\
h_c &:= \tfrac{\sigma_R^2 h_a}{\delta} + \tfrac{\kappa\bar{R}h_b}{\delta} + \tfrac{1}{\delta^2}\left(\tfrac{(\tilde{\mu}-r)^2}{2\sigma^2} + r - \delta\right), \\
k_c &:= \sigma_R^2 k_a + \kappa\bar{R}k_b + \tfrac{(\delta-1)}{\delta}\left(\tfrac{(\tilde{\mu}-r)^2}{2\sigma^2} + r - \delta\right), \\
l_c &:= \tfrac{\sigma_R^2(k_a+h_a)}{2\kappa} - \tfrac{\kappa\bar{R}l_b}{2\kappa}, \quad m_c := \kappa\bar{R}(l_b + k_b + h_b).
\end{aligned}
\tag{6.33}
$$

We observe that the solution of the finite-time horizon case (6.31) converges to the solution of the infinite-time horizon case (6.19) when T goes to infinity. In detail, we have $h_a = a$, $h_b = b$, $h_c = c$ and thus $(a(t), b(t), c(t)) \to (a, b, c)$ when $T \to \infty$.

It is reasonable that the value function v is quadratic with respect to the returns R_t. Plugging the optimal control π_t^* into the SDE for X_t, we obtain that the optimal wealth process X_t has a quadratic drift w.r.t. R_t:

$$
dX_t = \frac{(R_t + \tilde{\mu} - r)^2}{\sigma^2} + (r - \tilde{\chi}_t))X_t dt + \frac{(R_t + \tilde{\mu} - r)}{\sigma} X_t dW_t.
$$

6.2.3 Optimized stochastic returns

Let us consider the case of optimized stochastic returns with the term $\max(R_t, 0)$ in (6.8). By using the same Ansatz (6.29) as above, on one hand, the optimal controls become:

$$
\pi^*(z) = \frac{\max(z, 0) + \tilde{\mu} - r}{\sigma^2}, \quad \chi_t^* = \frac{1}{\beta(t)}.
$$

On the other hand, the parameters $a(t), b(t)$ and $c(t)$ have to satisfy the following ODE system whenever $z < 0$:

$$
\begin{aligned}
\dot{a}(t) &= (\delta + 2\kappa)\, a(t), && a(T) = 0, \\
\dot{b}(t) &= (\delta + 1)\, b(t) - 2\kappa\bar{R}a(t), && b(T) = 0, \\
\dot{c}(t) &= \delta c(t) - \sigma_R^2 a(t) - \kappa\bar{R}b(t) - \beta(t)\left(\tfrac{(\tilde{\mu}-r)^2}{2\sigma^2} + r - \delta\right), && c(T) = 0.
\end{aligned}
\tag{6.34}
$$

It is easy to verify that the only solutions are $a(t) = b(t) \equiv 0$ and the solution for $c(t)$ is given by:

$$c(t) = \left(\frac{(\delta-1)}{\delta}(T-t) - \frac{1}{\delta^2}\right) K e^{-\delta(T-t)} + \frac{K}{\delta^2}, \quad \forall t \in [0,T]. \tag{6.35}$$

Plugging (6.35) into (6.31) we see that the value function v is equal to the value function for the finite-time horizon case without returns in Table 1. So for $z < 0$ we get:

$$\begin{aligned} v(t,x) = &\quad e^{-\delta t}\left(\beta(t)\ln(\tfrac{x}{\beta(t)}) + \tfrac{1}{\delta^2}K\right) + b(t)K, \\ \pi^* = &\quad \tfrac{\tilde{\mu}-r}{\sigma^2}, \quad \tilde{\chi}_t^* = \beta(t)^{-1}. \end{aligned} \tag{6.36}$$

When we combine the solutions (6.31) for $z > 0$ and (6.36) for $z < 0$, the value function $v(t,x,z)$ is not continuous at $z = 0$ and not differentiable w.r.t. z at $z = 0$. It follows that we can not use the Ansatz (6.23) to construct a closed form solution for the value function v in the case of optimized stochastic returns.

Instead, if $v \in C^2$, the optimal control π_t^* would have the following form:

$$\pi_t^*(z) = \left[\frac{\max(z,0) + \tilde{\mu} - r}{\sigma^2}\,\partial_y v + \rho\frac{\sigma_R}{\sigma}\,\partial_{yz}^2 v\right] / \left[\partial_y v - \partial_{yy}^2\right]. \tag{6.37}$$

Note that the optimal control $\pi_t^*(z)$ may not be differentiable w.r.t. z at $z = 0$.

6.2.4 Analytical solutions

Table 2 shows the analytical solutions for the stochastic return case. We believe that it is not possible to find a closed form solution for the optimized return case.

	Infinite-time horizon	Finite-time horizon
Log utility	$v(x,z) = \frac{1}{\delta}\log(\delta x) + az^2 + bz + c,$ $\pi^*(z) = \frac{z+\tilde{\mu}-r}{\sigma^2}, \quad \chi^* = \delta.$	$v(t,x,z) = e^{-\delta t}\left[\beta(t)\left(\log(\frac{x}{\beta(t)}\right) + a(t)z^2 + b(t)z + c(t)\right],$ $\pi^*(z) = \frac{z+\tilde{\mu}-r}{\sigma^2}, \quad \chi_t^* = \frac{1}{\beta(t)}.$

Table 2: Solutions of the investment-consumption problem with stochastic returns for log utility and different time horizons.

6.2.5 Conclusions

In the case of **stochastic returns**, the first terms of the **value functions** v in Table 2 are equal to the original Merton (constant return) case in Table 1. The additional second term is a quadratic polynomial in z and independent of x. The splitting of x and z in two separate terms is only the case for log utility. Consequently, the correlation ρ disappears in the solution because $\partial_{xz} v(x,z) = 0$. No solutions for other utilities are known yet.

The **optimal proportion of wealth invested in assets** $\pi^*(z)$, depends linearly on the return rate R_t and therefore it becomes stochastic:

$$\pi^*(R_t) = \frac{R_t + \tilde{\mu} - r}{\sigma^2}.$$

In contrast to the fixed *Merton line* $\bar{\pi}$, the transition from investment $\alpha_t^* > 0$, to de-investment, $\alpha_t^* < 0$ now depends linearly on the stochastic return R_t. Note that $\pi^*(R_t)$ depends only on the current state of R_t, but not on the return's volatility σ_R or on its mean-reversion parameters κ and $\bar{R}$.

The **optimal consumption** χ_t^* is equal to the constant return case and does not depend on the stochastic returns at all. Moreover, the infinite-time horizon case is the limit of the finite-time horizon case when $T \to \infty$.

In the case of **optimized stochastic returns** we are not able to derive a closed form solution for the optimized return case. We analyze this case in our numerical studies in Section 7.5. There we use our controlled forward-backward algorithm from Part III, which utilizes the Stochastic Maximum Principle (instead of the HJB theory) and therefore requires less smoothness of v.

6.3 Investment-consumption model with proportional transaction costs

Now we consider the investment-consumption Problem 2 with proportional transaction costs while returns are neglected. It is well-known that there exists no analytical solution for this problem. Instead, this section presents the known analytical insights and numerical results. The section serves as a preparation for our numerical studies on quadratic transaction costs in the following Section 7.

6.3.1 The problem

Again, let us consider the 1-dimensional case $n = d = 1$ of Problem 2 and let the return rate R_t and the fixed costs M_t be zero. This implies $\kappa, \bar{R}, \sigma_R, c = 0$. But this time, let the quadratic transaction costs term $c_{2t}\alpha^2$ in (5.14) be replaced by a proportional transaction costs term $c_1|\,\alpha_t|$. Also, let the control space $\mathcal{A}$ be unbounded and let the well-posedness condition (6.1) be satisfied. Then the state system (5.14) becomes:

$$\begin{aligned} dA_t &= \left(\mu A_t + \alpha_t\right) dt \;+\; \sigma A_t\, dW_t, \qquad A_0 = a, \\ dB_t &= \left(\, rB_t - \chi_t - \alpha_t - c_1|\,\alpha_t|\,\right) dt, \qquad B_0 = b, \end{aligned} \tag{6.38}$$

where $c_1 > 0$. In the case of proportional transaction costs, transaction charges are imposed equal to a constant fraction of the amount transacted. We discussed the relevance of proportional transaction costs for our discrete-time model in Section 4.4.3.

Facing transaction costs, we can not eliminate the allocation rate α_t by making the usual change of variables $A_t + B_t \mapsto X_t$, $A_t/X_t \mapsto \pi_t$. On one hand, we have to deal with both dimensions A_t and B_t separately instead of one dimension X_t. On the other hand, the control variables α_t and χ_t appear only in the drift terms in the SDEs (6.38) but not in the diffusion terms.

6.3.2 Infinite-time horizon

Assuming that there exists a smooth value function $v(a,b) \in C^2(\mathbb{R}\times\mathbb{R},\mathbb{R})$, the HJB equation reads:

$$-\delta v(a,b) + \sup_{\pi,\chi\in\mathbb{R}\times\mathbb{R}_+} \left[u(\chi x) + A^{\alpha,\chi} v(a,b)\right] = 0, \tag{6.39}$$

where the generator $A^{\alpha,\chi}$ is given by:

$$\begin{aligned} A^{\alpha,\chi} v(a,b) \coloneqq\; & \left[\mu a + \alpha_t\right] \partial_a v + \tfrac{\sigma^2}{2} a^2\, \partial_{aa} v \\ & + \left(rb - \chi - \alpha_t - c_1\,|\,\alpha_t\,|\right)\, \partial_b v. \end{aligned} \tag{6.40}$$

Since the HJB equation is linear in α for both regions $\alpha > 0$ and $\alpha < 0$, the optimal control is of bang-bang type. Moreover, since both derivative $\partial_a v$ and $\partial_b v$ are naturally positive, the optimal allocation rate α^* is given as follows:

$$\begin{aligned} &\alpha^* = \infty, && \text{if } \partial_a v \geq (1+c_1)\partial_b v, \\ &\alpha^* = 0, && \text{if } (1+c_1)\partial_b v > \partial_a v > (1-c_1)\partial_b v, \\ &\alpha^* = -\infty, && \text{if } \partial_a v \leq (1-c_1)\partial_b v. \end{aligned} \tag{6.41}$$

Note that the bang-bang control splits the state space (a, b) into three regions: a buy, a hold and a sell region. The allocation policy (6.41) states that an investor immediately shifts his portfolio to the boundary of the hold region where $\partial_a v = (1+c_1)\partial_b v$ or $\partial_a v = (1-c_1)\partial_b v$ when his portfolio is in the sell or the buy region, respectively.

It turns out that the proportional transaction costs model is a free-boundary problem. Let $X = A + B$ denote the total wealth of the investor and let π denote the proportion of wealth invested in the risky asset A. Then, the problem is to find the proportions π_b^* and π_s^* such that the boundaries of the buy and the sell regions are given by $(\pi_b^* X, (1-\pi_b^*)X)$ and $(\pi_s^* X, (1-\pi_s^*)X)$ in the (a, b) space, respectively.

[Davis, Norman 1990] showed for the case of constant relative risk aversion (CRRA) that the boundaries are straight lines in the (a, b) space and the hold region is a wedge containing *Merton's CRRA line* $\pi^* = \frac{\mu - r}{\sigma^2}$, see Figure 8.

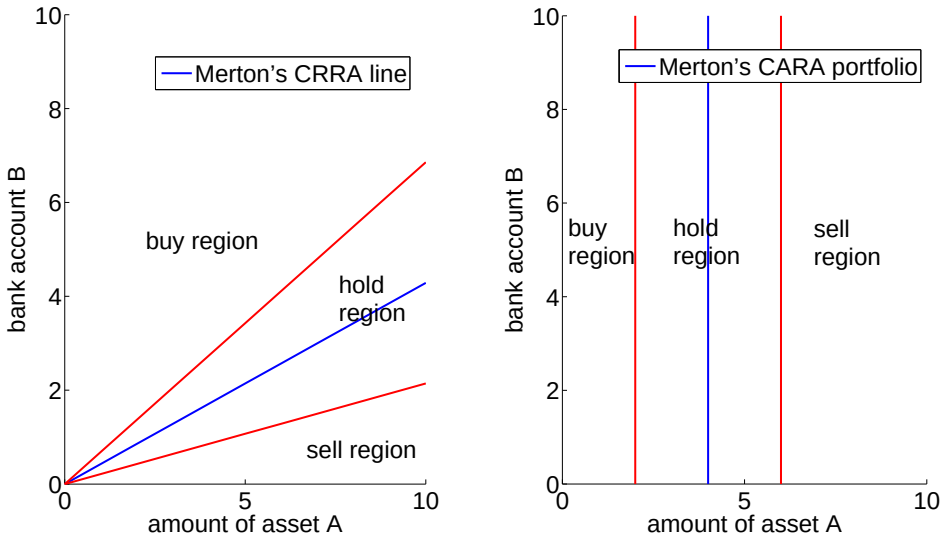

Figure 8: Examples of buy, hold and sell regions in the (A, B) space for CRRA (power) utility and CARA (negative exponential) utility.

They exploited the fact that the value function v is concave and homogeneous

in (a, b), $\forall p > 0$:

$$\begin{aligned} v(pa, pb) &= p^{\gamma} v(a, b), && \text{for } u(\chi) = \tfrac{\chi^{\gamma}}{\gamma}, \\ v(pa, pb) &= \tfrac{1}{\delta} \log(p) + v(a, b), && \text{for } u(\chi) = \log(\chi), \end{aligned} \tag{6.42}$$

and showed that there exists a smooth solution for the value function $v(a, b) \in C^2(\mathbb{R} \times \mathbb{R})$ and that the control policy α^* in (6.41) is optimal.

The most noticeable feature of the model's behavior according to the transaction cost factor c_1 is that:

1. π_s is strictly increasing w.r.t c_1 but the rate of increase is decreasing with higher c_1. π_b is decreasing w.r.t c_1 accordingly, see Figure 9.

2. *Merton's CRRA line* π^* is not the center of π_s and π_b (green line in Figure 9). In contrast, the lower (buy) boundary π_b is decreasing faster than the upper (sell) boundary π_x is increasing for higher c_1.

[Davis, Norman 1990] state that "this is probably due to the asymmetry in the model: all consumption takes place from the bank, so stock must be sold (and transaction charges paid) before it can be realized for consumption." So it may not be worthwhile to invest in stocks and then reselling them if transaction costs are high. This is why π_b is decreasing faster, see Figure 9.

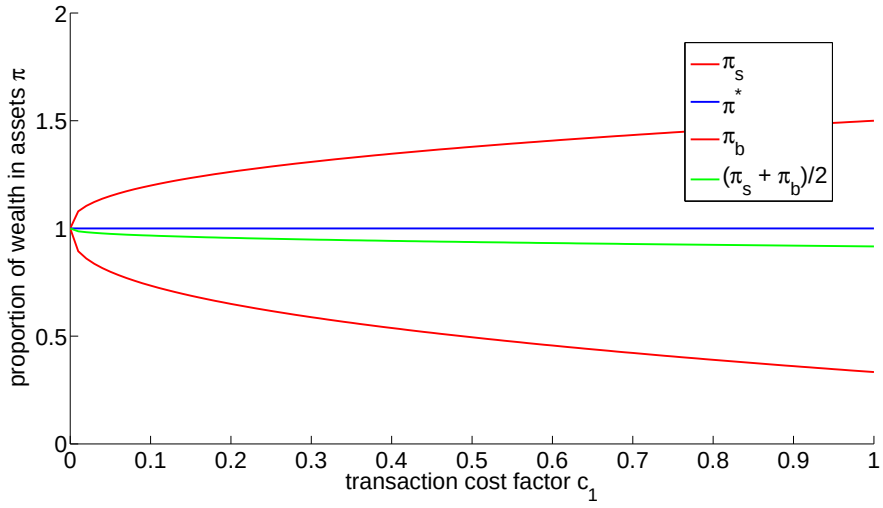

Figure 9: Example of the sell and the buy boundaries, π_s and π_b, plus *Merton's CARA line* π^* for different $c_1 > 0$.

Besides [Davis, Norman 1990], many authors studied numerical schemes to calculate the free boundaries. [Muthuraman, Kumar 2006] for example adapt a finite element method and use an iterative procedure that converts the free boundary problem into a sequence of fixed boundary problems.
[Tourin, Zariphopoulou 1997] considered viscosity solutions and used finite differences and dynamic programming to obtain numerical results while more recent papers like [Janecek, Shreve 2005] consider "an heuristic derivation of the asymptotic expansion of the value function".

[Akian, Sequier, Sulem 1995] showed that the hold region converges to *Merton's CRRA line* when $c_1 \to 0$, see Figure 9.

6.3.3 Finite-time horizon

Constant relative risk aversion

The finite-time horizon case with constant relative risk aversion CRRA as been studied only recently for example by [Dai, Jiang, Li, Yi 2009]. They present an analytical approach to analyze the behaviors of the free boundaries. Note that the finite-time horizon is more challenging since the corresponding free boundaries vary with time. They used the terminal utility:

$$e^{-\delta T} u\left((1-c_1)A_T + B_T \right) \tag{6.43}$$

and found the following results. Let $\tau = T - t$ denote the time until T, and let $x = \frac{b}{a}$ denote the proportion between the bank account b and the monetary amount of assets a. Let $x_s(\tau)$ and $x_b(\tau)$ denote the boundaries of the sell and the buy region, respectively and let x_M denote the point on *Merton's CRRA line*. Then for all $\tau > 0$:

$$\begin{aligned} x_s(\tau) &\leq (1-c_1)x_M, \\ x_b(\tau) &\geq (1+c_1)x_M, \\ x_b(\tau) &= \infty \text{ if and only if } \tau \leq \tau_0, \end{aligned} \tag{6.44}$$

where $\tau_0 = \frac{1}{\mu - r} \log\left(\frac{1+c_1}{1-c_1}\right)$. The strict distance between x_s, x_b and x_M even for τ close to zero is due to the choise of the terminal utility in (6.43). [Dai, Jiang, Li, Yi 2009] state that the third equation in (6.44) "indicates that there is a critical time τ_0 after which it is never optimal to purchase stocks." "If the investor does not have a long enough expected horizon to recover at least the transaction costs, then s/he should not purchase any additional stock."

Neglecting consumption, [Gennotte, Jung 1992] numerically show that the no transaction region narrows and converges rapidly to the infinite horizon limit as the time horizon increases. They used a standard dynamic programming approach on the binominal model of asset returns. [Oksendal, Sulem 2002] analysed the case that additionally involves a fixed transaction cost $c_0 > 0$ by considering viscosity solutions.

Constant absolute risk aversion

[Liu 2004] considers the optimal intertemporal consumption and investment policy of a constant absolute risk aversion (CARA) investor who faces proportional and also fixed transaction costs when trading multiple risky assets. In the case of porportional transaction costs, he showed that the boundaries of the hold region A_s and A_b are independent of the holdings in the bank account, see the second chart in Figure 8.

He also observes that *Merton's CARA portfolio* A^* lies inside the hold region and that the boundaries A_s, A_b show the same increasing, decreasing effect for higher c_1 as in the CRRA case, see Figure 9.

6.3.4 Conclusions

In the case of proportional transaction costs $c_1|\alpha|$, the allocation rate α_t can not be eliminated through a change of variables. On one hand, analytical solutions cannot be derived anymore. On the other hand, the value function is still **homogeneous** (6.42) and therefore the problem can be turned into a free boundary problem. **Numerical methods** are available to compute solutions.

The (A, B) space splits into three region: a buy region, a hold region and a sell region, see Figure 8. The **optimal allocation rate** α^* is ∞ in the buy region, 0 in the hold region and $-\infty$ in the sell region. This is a **bang-bang control**. Roughly speaking, trading only takes place at the boundary of the hold region. Whenever the portfolio leaves the hold region, the infinite allocation rate immediately brings it back to the hold region's boundary.

In case of CRRA (CARA) utility, the boundaries π_s and π_b (A_s and A_b) of the sell and buy region diverges from *Merton's CRRA line* π^* (*Merton's CARA portfolio* A^*) with **increasing transaction cost factor** c_1, see Figure 9. The **divergence slows down** for larger c_1 and it is not symmetric due to the **asymmetry** of the problem: the consumption is only taken out of the

bank account.

The **optimal consumption** $\chi^* = \dot{u}^{-1}\left(e^{\delta t}\partial_b v\right)$ has a similar functional form as the optimal consumption in Merton's original model but it depends on $\partial_b v$ instead of $\partial_x v$.

The **finite-time horizon** case is more challenging since the corresponding free boundaries varies with time. Numerical results show that the model behaves similar to it's infinite-time horizon version and converges to it when $T \to \infty$.

6.4 Investment-consumption model with quadratic transaction costs

The motivation for quadratic transaction costs comes from the second order Taylor expansion of convex temporary illiquidity effects in Section 5.2.4. In addition to commodity markets, the illiquidiy of stock markets has become of more interest in recent years after the financial crices in 2008.

Again, let us consider the 1-dimensional case $n = d = 1$ of Problem 2 and neglect the return rate and the fixed costs $(\kappa, \bar{R}, \sigma_R, c, M = 0)$ in order to focus on the quadratic transaction costs $c_2 > 0$. Also, let the control space $\mathcal{A}$ be unbounded and let the well-posedness condition (6.1) be satisfied. Then the state system (5.14) becomes:

$$\begin{aligned} dA_t &= \left(\mu A_t + \alpha_t\right) dt \; + \; \sigma A_t \, dW_t, \qquad A_0 = a, \\ dB_t &= \left(r B_t - \chi_t - \alpha_t - c_2 \alpha^2 \right) dt, \qquad B_0 = b. \end{aligned} \tag{6.45}$$

Let us assume that there exists a smooth value function $v(t,a,b) \in C^{1,2}([0,T] \times \mathbb{R} \times \mathbb{R}, \mathbb{R})$. Then the generator $A^{\alpha,\chi}$ of the HJB equation (6.39) becomes:

$$\begin{aligned} A^{\alpha,\chi} v(t,a,b) := \; & \left[\mu a + \alpha_t\right] \partial_a v + \tfrac{\sigma^2}{2} a^2 \, \partial_{aa} v \\ & + \left(rb - \chi_t - \alpha_t - c_2 \alpha_t^2\right) \partial_b v. \end{aligned} \tag{6.46}$$

The optimal controls α_t^*, χ_t^* can be found by considering the first derivatives of the HJB equation w.r.t. α and χ:

$$\begin{aligned} 0 &= \partial_a v - \partial_b v - 2 c_2 \alpha \, \partial_b v, \\ 0 &= e^{-\delta t} \dot{u}(\chi_t) - \partial_b v. \end{aligned} \tag{6.47}$$

Since $\partial_a v, \partial_b v$ are naturally positive and the utility function u is strictly concave the following controls are optimal:

$$\begin{aligned} \alpha_t^* &= \tfrac{1}{2c_2}\tfrac{\partial_a v - \partial_b v}{\partial_b v}, \\ \chi_t^* &= \dot{u}^{-1}\left(e^{\delta t}\partial_b v\right). \end{aligned} \tag{6.48}$$

From (6.48) we observe that the optimal allocation rate α_t^* is not of bang-bang type but rather smooth and differentiable when $v \in C^2$. We also observe that the optimal allocation is zero when both partial derivatives attain the same value. So the optimal portfolio is the set of points $\{A_t^*, B_t^*\}$ in the (A, B) space where $\alpha_t^* = 0$:

$$\{A_t^*, B_t^*\} = \{(A, B) \in \mathbb{R}_+ \times \mathbb{R} \mid \alpha_t^* = 0 \ (\partial_a v = \partial_b v)\}. \tag{6.49}$$

On one hand, even though we gained a lot of experience on finding an appropriate Ansatz of a possible value function, we are not able to construct an analytical solution in the quadratic transaction costs case. The main reason is that the value function is no longer homogeneous (6.5) because of the quadratic term.

On the other hand, the quadratic (convex transaction costs) term provides a convexity property to the problem. Therefore, our CFB algorithm of Section 12 can be applied to it. Actually, besides the illiquidity issue in commodity markets, the convexity property of the problem is one reason why we choose the convex transaction costs model.

6.4.1 Adding stochastic returns

Now, let us add the stochastic return rate R_t to the investment-consumption model with quadratic transaction costs ($\kappa, \bar{R}, \sigma_R, M > 0$). Then the state system (5.14) becomes:

$$\begin{aligned} dA_t &= (\mu A_t + \alpha_t)\, dt \ + \ \sigma A_t\, dW_t, && A_0 = a, \\ dB_t &= (\, rB_t + R_t A_t - MA_t - \chi_t - \alpha_t - c_2\alpha^2\,)\, dt, && B_0 = b, \\ dR_t &= \kappa\left(\bar{R} - R_t\right) dt \ + \ \sigma_R\, dW_t^R, && R_0 = c. \end{aligned} \tag{6.50}$$

Let us assume that there exists a smooth value function $v(t, a, b, c) \in C^{1,2}([0, T] \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}, \mathbb{R})$. Then the generator $A^{\alpha,\chi}$ of the HJB equation (6.39) becomes:

$$\begin{aligned} A^{\alpha,\chi} v(t, a, b, c) := \ & [\mu a + \alpha_t]\, \partial_a v + \tfrac{\sigma^2}{2} a^2\, \partial_{aa} v \\ & + (rb + ca - Ma - \chi_t - \alpha_t - c_2\alpha_t^2)\, \partial_b v \\ & + \kappa(\bar{R} - c)\, \partial_c v + \tfrac{\sigma_R^2}{2}\, \partial_{cc} v \ + \ \rho\sigma\sigma_R a\, \partial_{ac} v. \end{aligned} \tag{6.51}$$

The optimal controls α_t^*, χ_t^* can be found by considering the first derivatives of the HJB equation w.r.t. α and χ:

$$\begin{aligned} \alpha_t^* &= \frac{1}{2c_2} \frac{\partial_a v - \partial_b v}{\partial_b v}, \\ \chi_t^* &= \dot{u}^{-1} \left(e^{\delta t} \partial_b v \right). \end{aligned} \tag{6.52}$$

Note that (6.52) is equal to (6.48) since the return rate is a non-controlled diffusion.

We present numerical results in the following Section 7.

7 Numerical results for the general one asset case

In this section we present numerical results of our reduced continuous-time model (Problem 2) (5.13)-(5.16) in the case of one asset. The single asset could be interpreted as the market portfolio in Merton's mutual fund theorem, see Section 3.1.5 in the preliminaries.

Considering quadratic transaction costs is much more challenging than proportional transaction costs which we present in Section 6.3. The reason is that the value function **loses its homogeneity** (6.5) and therefore the model cannot be simplified by a reduction of variables[27]. Moreover, since there are no up-front analytical insights about the solution, the model must be analyzed fully by **numerical simulations**.

As far as we know, the case of convex (quadratic) transaction costs has not been studied in the literature. A reason may be the numerical complexity of the multi-dimensional stochastic optimal control problem. In the case of just one asset, we are actually dealing with three dimensions: the asset, the bank account and the return rate.

Our **controlled forward-backward (CFB) stochastic algorithm** from Part III gives us the ability (in terms of performance) to study this convex problem numerically. All following results are generated with this method. The CFB algorithm is developed in Part III and all details about the implementation are given in Part IV of this thesis.

The outline is as follows. We define one asset case in Section 7.1 and analyze the pure quadratic transaction cost case in Section 7.2 in detail. After a short aside on proportional transaction costs in Section 7.3, we add stochastic returns in Section 7.4 and then consider optimized returns in Section 7.5.

[27] In the case of proportional transaction costs the reduction is made through $(a, b) \mapsto \frac{a}{b}$.

7.1 Definition of the problem for one asset

Throughout the section we use the negative exponential (CARA) utility in formulas and examples for simplicity and numerical convenience[28]:

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

We leave other HARA utility function to future research. Now, let us recall Problem 2 (5.13)-(5.16) with negative exponential utility for one asset.

Problem 3 (Continuous-time problem for one asset)
Let the objective function be given by:

$$J(t,a,b,c,\alpha,\chi) = E_t^{\alpha,\chi}\left[\int_t^T e^{-\delta s}\frac{1-e^{-\gamma\chi_s}}{\gamma}\,ds \;+\; e^{-\delta T}\frac{1-e^{\gamma(A_T+B_T)}}{\gamma}\,\middle|\, a,b,c\right], \tag{7.1}$$

and let the states evolve according to:

$$\begin{aligned} dA_t &= (\mu A_t + \alpha_t)\,dt \;+\; \sigma A_t\,dW_t, && A_0 = a,\\ dR_t &= \left(\kappa(\bar{R} - R_t)\right)\,dt \;+\; \sigma_R\,dW_t^R && R_0 = c,\\ dB_t &= (rB_t + \max(R_t,0)A_t - MA_t && \\ &\quad -\chi_t - \alpha_t - c_2\alpha^2)\,dt, && B_0 = b, \end{aligned} \tag{7.2}$$

where:

$$dW_t\,dW_t^R \;=\; \rho\,dt. \tag{7.3}$$

We search for the optimal allocation policy $\alpha^*(t,a,b,c)$ and the optimal consumption policy $\chi^*(t,a,b,c)$:

$$(\alpha^*(t,a,b,c),\chi^*(t,a,b,c)) = \arg\max_{(\alpha,\chi)} J(t,a,b,c,\alpha,\chi), \tag{7.4}$$

and the value function $v(t,a,b,c)$:

$$\begin{aligned} v(t,a,b,c) &= \max_{(\alpha,\chi)} J(t,a,b,c,\alpha,\chi),\\ v(T,a,b,c) &= g(a,b,c), \end{aligned} \tag{7.5}$$

for any given initial time $t \in [0,T]$ with initial monetary amounts of assets $a \in [A_{min}, A_{max}] \subset \mathbb{R}_+$, initial money in the bank account $b \in [B_{min}, B_{max}] \subset \mathbb{R}$

[28]Negative exponential utility is defined on the whole $\mathbb{R}$ space, in contrast to log utility which is only defined on $\mathbb{R}_+$.

and initial return rates $c \in [R_{min}, R_{max}] \subset \mathbb{R}$.

The state and control constraints are:

$$A_t \geq 0, \quad \alpha_t \in [\alpha_{min}, \alpha_{max}], \quad \chi_t \in [0, \chi_{max}], \tag{7.6}$$

and the exit strategy in case of bankruptcy is given by:

$$(\alpha_t, \chi_t) = \begin{cases} (-\alpha_{bry}, 0), & \text{if } (A_t, B_t) \notin \mathcal{L}_t(\alpha_{bry}) \wedge A_t > 0, \\ (0, 0), & \text{if } (A_t, B_t) \notin \mathcal{L}_t(\alpha_{bry}) \wedge A_t = 0, \end{cases} \tag{7.7}$$

where:

$$\mathcal{L}(\alpha_{bry}) = \{(x, y) \in \mathbb{R}_+ \times \mathbb{R} \mid y + (1 - c_2\, \alpha_{bry})\, x \geq 0\}. \tag{7.8}$$

In Problem 3, all parameters are given as positive constants: $T, \delta, \gamma, \mu, \sigma, c_2 > 0$ and $\kappa, \bar{R}, \sigma_R, r, M, \rho, \alpha_{bry} \geq 0$. In order to avoid short-selling naturally, we assume that μ is sufficiently larger than r.

The region of interest defined by A_{min}, A_{max}, B_{min}, B_{max}, R_{min} and R_{max} is part of the problem's statement with the only restriction that $A_{min} \geq 0$. The bounds of the control space $\alpha_{min}, \alpha_{max}$ and χ_{max} are not part of the problem's statement but rather of technical nature to indicate that the controls are bounded.

7.2 Quadratic transaction costs

Let us at first drop all state and control constrains. Also let us neglect the return rate and the fixed costs ($\kappa, \bar{R}, \sigma_R, c, M = 0$) in order to focus on the quadratic transaction costs $c_2 > 0$. Then the state system (7.2) becomes:

$$\begin{aligned} dA_t &= (\mu A_t + \alpha_t)\, dt + \sigma A_t\, dW_t, \qquad A_0 = a, \\ dB_t &= (\, r B_t - \chi_t - \alpha_t - c_2 \alpha^2 \,)\, dt, \qquad B_0 = b. \end{aligned} \tag{7.9}$$

From our analyses in Section 6.4, we assume that we have a smooth value function and the optimal controls are given by equation (6.48):

$$\begin{aligned} \alpha_t^* &= \tfrac{1}{2c_2} \tfrac{\partial_a v - \partial_b v}{\partial_b v}, \\ \chi_t^* &= \dot{u}^{-1}\left(e^{\delta t} \partial_b v\right). \end{aligned} \tag{7.10}$$

In particular, the optimal allocation is not singular (bang-bang) but rather smooth due to the convexity of the problem: the gain function is concave in χ and the costs are convex (quadratic) in α. The following numerical results support this assumption.

7.2.1 Basic parameter set up

Let us fix the parameters for a basic set up to start with our study. All further analyses are variations of this basic set up. The region of interest is:

$$\begin{aligned} &a \in [0,2], && \text{initial monetary amount of asset in €},\\ &b \in [-1,1], && \text{initial monetary amount in bank account in €},\\ &c \in [-0.3,0.3], && \text{initial return of asset in } 100\%. \end{aligned} \tag{7.11}$$

The model parameters of the basic set up are defined in Table 3. Note that all parameters are constant over time $[0,T]$ and that the return R_t and the fixed costs M are zero in the basic set up.

T	γ	δ	r	μ	σ	c_2	M	$\bar{R}$	κ	σ_R	ρ
2	1	0.01	0.02	0.06	0.2	0.1	0.0	0.0	0.0	0.0	0.0

Table 3: Parameters of the basic set up for the one asset case.

The market in Table 3 has a low interest rate r while the asset's drift and volatility are high. Gamma is chosen such that the utility has a significant curvature in the region of interest $(a,b) \in [0,2] \times [-1,1]$. We choose values that are observable at current equity exchanges and that are commonly used in the literature.

7.2.2 Approximation parameters

Let h denote our approximation parameter. For our calculations, we use the time step size $\Delta t = h$. All other grid step sizes A, B, R and the number of simulations of the Brownian increments $\Delta W, \Delta W^R$ are derived from h. In particular, they are optimized according to equation (13.12) in Section 13.2 in Part III, in order to provide linear convergence of our CFB algorithm when $h \to 0$. In the following we use the superscript h to indicate a numerical approximation χ^h of the opitmal solution χ^*.

7.2.3 Shape of solutions at time $t = 0$

In order to get a first impression of the solution's behavior in the (a,b)-space, we present the surfaces of v^h, χ^h and α^h at time $t = 0$. Figure 10 shows the **value function** $v^h(0,a,b)$ and the optimal **consumption rate** $\chi_0^{*h}(a,b)$.

Consistent with the results of Section 6.3, the value function is increasing and concave w.r.t. the amount of assets a and the bank account b. We also observe that χ_0^h is almost linear in $a+b$.

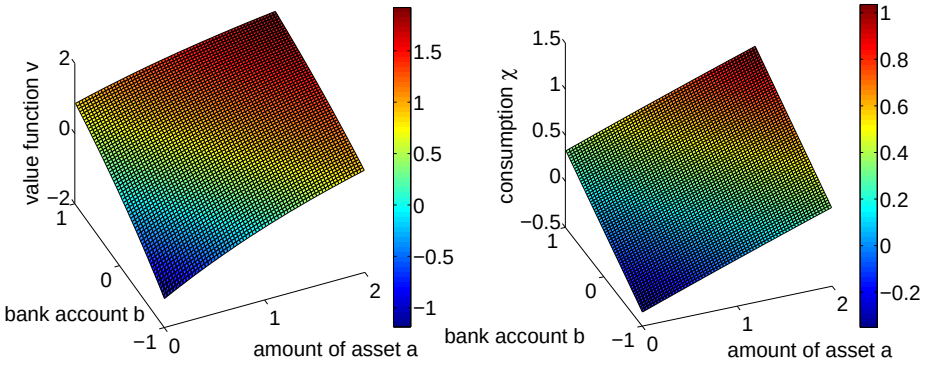

Figure 10: Value function $v^h(0,a,b)$ and optimal consumption rate $\chi_0^h(a,b)$ at initial time 0, for $h = 0.1$.

Figure 11 shows the optimal **allocation rate** $\alpha_0^h(a,b)$. It appears that $\alpha_0^h(a,b)$ is independent of b and decreases almost linearly w.r.t. a, at least in the neighbourhood of the optimal portfolios $\{A_0^h, B_0^h\}$ ($\alpha_0^h = 0$). We observe that $A_0^h = 1.70$ is smaller than *Merton's CARA portfolio* $A_0^* = 2.92$.

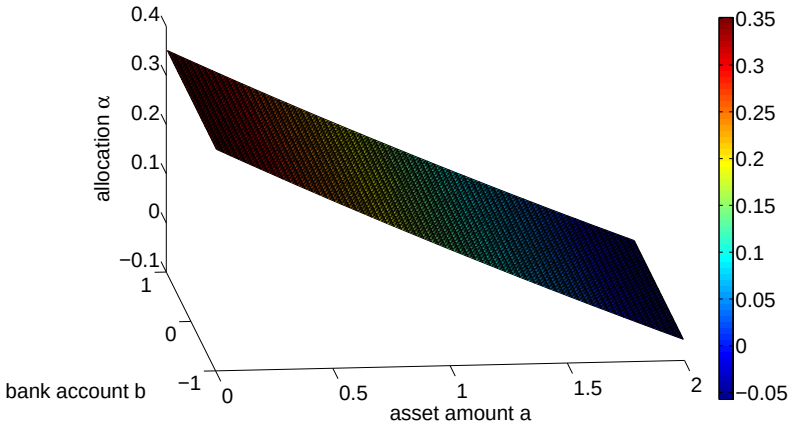

Figure 11: Optimal allocation rate $\alpha_0^{*h}(a,b)$ for all $(a,b) \in [0,2] \times [-1,1]$ at initial time 0, for $h = 0.1$.

7.2.4 Behavior of solutions over time $t \in [0, T]$

In order to get an impression of the solution's behavior over time, we present v^h, χ^h and α^h at five selected (fixed) space points $(a, b) \in [0, 2] \times [-1, 1]$ over $t \in [0, T]$. For this paragraph we set $T = 4$ in order to observe the evolution over a longer time period.

Figure 12 shows the evolution of the **value function** $v^h(t, a, b)$. Naturally, v^h is strictly decreasing with time t since the remaining time $[t, T]$ becomes shorter. We also observe that the level of the value function is strongly related to the total wealth $a + b$.

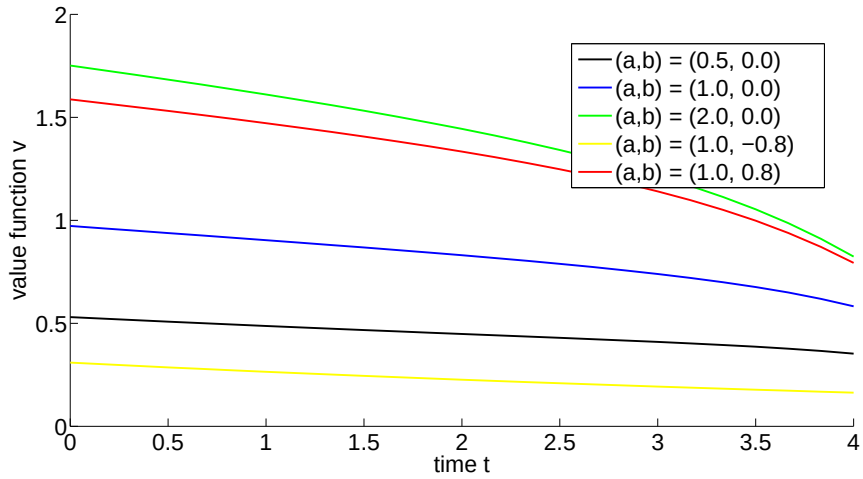

Figure 12: Evolution of the value function $v^h(t, a, b)$ for selected $(a, b) \in [0, 2] \times [-1, 1]$ over $t \in [0, 4]$, for $h = 0.2$.

Figure 13 shows the evolution of the optimal **relative consumption rate** $\frac{\chi_t^{*h}(a,b)}{a+b}$. All rates are greater than zero and approach the value 1 when $t \to T$. Also, the relative consumption rates seems to be almost equal for the different points.

Figure 14 shows the evolution of the optimal **allocation rate** $\alpha_t^{*h}(a, b)$. We observe that $\alpha_t^h \to 0$ when $t \to T$. Again, the allocation rate seems to be almost equal for equal amounts of assets a and independent of b.

The most varying curve in Figure 14 is the green one where $(a, b) = (2, 0)$. We observe that on one hand, an investor with this portfolio ($a = 2.0$) should buy

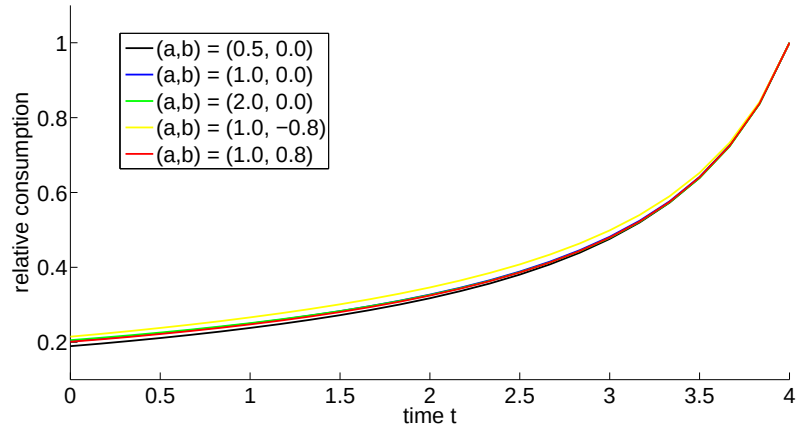

Figure 13: Evolution of the optimal relative consumption rate $\frac{\chi_t^{*h}(a,b)}{a+b}$ for selected $(a,b) \in [0,2] \times [-1,1]$ over $t \in [0,4]$, for $h = 0.1$.

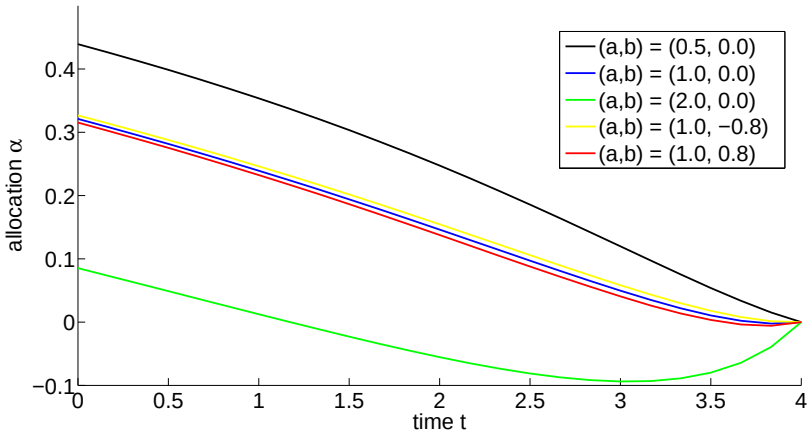

Figure 14: Evolution of the optimal absolute allocation rate $\alpha_t^{*h}(a,b)$ for selected $(a,b) \in [0,2] \times [-1,1]$ over $t \in [0,4]$, for $h = 0.2$.

assets to profit from the drift $\mu > r$ when the time horizon T is sufficiently far away. On the other hand, an investor with the same portfolio but closer to T should sell assets in order to reduce the risk through σ. This is an effect of the negative exponential utility and consistent with the decreasing *Merton's CARA portfolio*:

$$A_t^* = \frac{\mu - r}{\gamma r \sigma^2}(1 - (1 - r)e^{-r(T-0)}),$$

see Table 1 in Section 6.1. When $a < A_t^*$ assets are bought, when $a > A_t^*$ assets are sold. Very close to T the optimal allocation goes back to zero,

since the terminal value $u(A_T + B_T)$ does not treat money in the assets A_T and money in the bank B_T differently.

7.2.5 Reduction of consumption proportional to quadratic transaction costs

From the first spot, it seems that the consumption rate χ_t^h is almost equal to the analytical solution χ_t^* of Merton's original model without transaction costs:

$$\bar{\chi}_t = \frac{d(t)(a+b)}{\gamma} + \bar{a}(1 - e^{-r(T-t)}), \tag{7.12}$$

see Table 1 in Section 6.1. In particular, the observed linearity of χ^h in $a+b$ in Figure 10 and the similarity of the relatives rates $\frac{\chi_t^{*h}(a,b)}{a+b}$, as well as their convergence to one at $t = T$ in Figure 13, would be consistent with (7.12).

So, in order to analyze the difference between the consumption in the case of quadratic transaction costs and consumption in Merton's original model, we look at the differences between $\chi_t^h(a,b)$ and $\chi_t^*(a,b)$ at time $t = 0$ for several points in the (a,b) space:

1. $\epsilon_0^\chi(1,0) = |\chi_0^h(1,0) - \chi_0^*(1,0)|$, in the buy region,
2. $\epsilon_0^\chi(2,0) = |\chi_0^h(2,0) - \chi_0^*(2,0)|$, close to the optimal portfolio,
3. $\epsilon_0^\chi(3,0) = |\chi_0^h(3,0) - \chi_0^*(3,0)|$, in the sell region,
4. $\epsilon_0^\chi(max) = \max_{a,b} |\chi_0^h(a,b) - \chi_0^*(a,b)|$, maximum over region.

Table 4 shows the four values ϵ_0^χ for different h. We observe that $\epsilon_0^\chi(max)$ gets smaller for smaller h but the single $\chi_0^h(a,b)$ do not converge[29] exactly to $\chi_0^*(a,b)$ in all 3 cases. In other words, $\epsilon_0^\chi > 0$ is not just an approximation error.

Figure 15 shows the differences $\chi_0^*(a,b) - \chi_0^h(a,b)$ for all $(a,b) \in [1,3] \times [-1,1]$. We observe that $\chi_0^* - \chi_0^h$ is always positive and smallest near the optimal portfolio $\{A_0^h, B_0^h\}$, $(\alpha_0^h = 0)$. For higher or lower amounts of assets a, $(\alpha_0^h < 0$ or $\alpha_0^h > 0)$, the difference $\chi_0^* - \chi_0^h$ becomes larger but stays independent of the bank account b. We can say that the optimal consumption χ^h is reduced compared to χ^* when the allocation $|\alpha^h|$ is raised.

[29] The algorithm has linear convergences in h, see Section 13.

$\epsilon_0^\chi(1,0)$, buy region				
$h = 1/4$	$h = 1/8$	$h = 1/12$	$h = 1/16$	$h = 1/20$
0.0093	0.0054	0.0048	0.0046	0.0045
$\epsilon_0^\chi(2,0)$, close to the opt. portfolio				
$h = 1/4$	$h = 1/8$	$h = 1/12$	$h = 1/16$	$h = 1/20$
0.0140	0.0068	0.0060	0.0057	0.0055
$\epsilon_0^\chi(3,0)$, sell region				
$h = 1/4$	$h = 1/8$	$h = 1/12$	$h = 1/16$	$h = 1/20$
0.0326	0.0217	0.0205	0.0198	0.0195
$\epsilon_0^\chi(max)$, maximum over region				
$h = 1/4$	$h = 1/8$	$h = 1/12$	$h = 1/16$	$h = 1/20$
0.0342	0.0218	0.0206	0.0199	0.0198

Table 4: Differences between the consumptions χ_0^h (with quadratic transaction costs) and χ_0^* (without transaction costs) for different h.

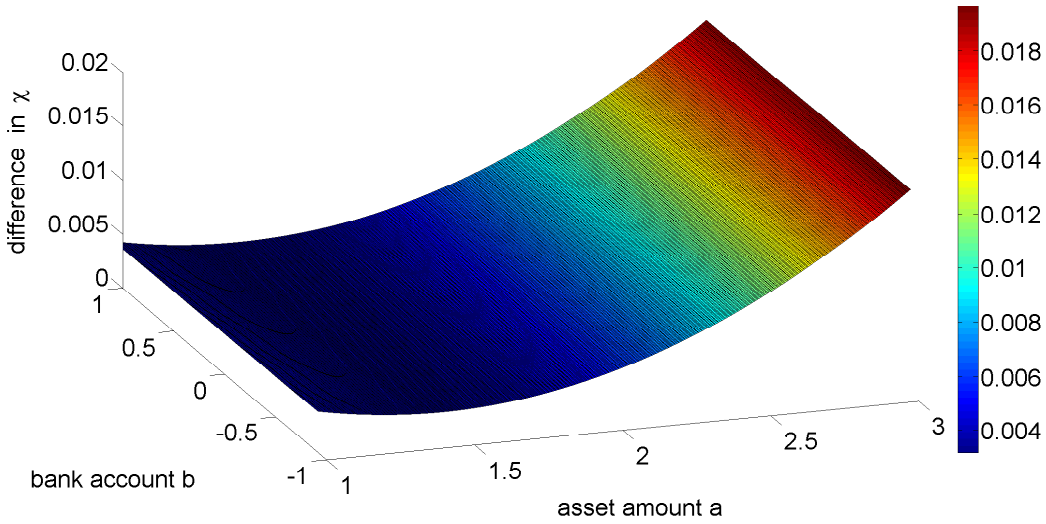

Figure 15: Reduction of consumption $\chi_0^*(a,b) - \chi_0^h(a,b)$ for all $(a,b) \in [1,3] \times [-1,1]$ at initial time 0, for $h = 0.05$.

Let us go one step further and compare the actual transaction costs $c_2(\alpha_0^h)^2$ and the reduction of consumption $\chi_0^* - \chi_0^h$. Figure 16 shows that the reduction of consumption is almost proportional to the quadratic transaction costs. The consumption is also shifted since the optimal portfolio $A^h \neq A^*$ is shifted too, see next paragraph.

It turns out that **the optimal consumption χ_t^h decreases (almost) proportional to the quadratic transaction costs $(\alpha^h)_t^2$ compared to the optimal consumption χ_0^* with no transaction costs**.

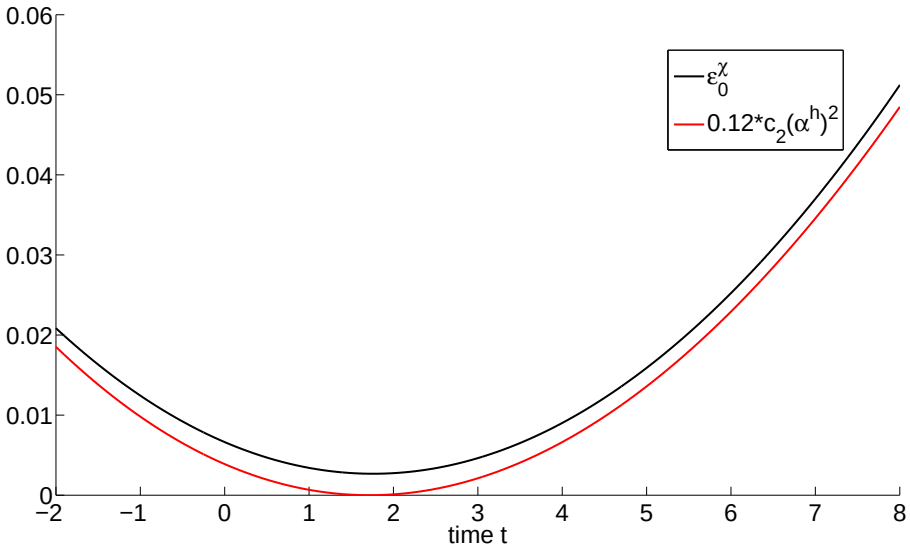

Figure 16: Reduction of consumption $\chi_0^*(a,0) - \chi_0^h(a,0)$ compared to transaction costs $c_2(\alpha^h)_t^2$ scaleld by a factor $0,12$ for $a \in [-2,8]$ at initial time 0; with $c_2 = 5e-3$ and $h = 0.03$.

Note that the small differences between the five curves (optimal relative consumption rate) in Figure 13 are not due to the different values of b. They rather occur due to the different allocation rates (compare to Figure 14) and due to the second term in (7.12), which is not proportional to the total wealth $a+b$ but goes to zero when $t \to T$. The latter argument explains why the portfolio with the smallest total wealth $a+b$ (yellow curve) shows the highest relative consumption $\frac{\chi_t^{*h}(a,b)}{a+b}$.

7.2.6 Discounted continuous allocation towards a smaller amount of assets

We observe two main effects for the optimal allocation rate α_t^h due to quadratic transactin costs.

First, we already observed that α_t^h is smooth in (a, b) and also smooth in t, see Figure 11 and 14. In particular, it is not of bang-bang type as in the proportional transaction costs case but rather decreases almost linearly w.r.t. a. Roughly speaking, having a non-optimal portfolio, the investor still trades towards the optimal portfolio, but the shift is not made immediately rather decelerated due to the quadratic transaction costs (temporary illiquidity costs).

In the case of negative exponential (CARA) utility, $\alpha_t^h(a, b)$ is independent of b and therefore the optimal portfolio ($\alpha_0^h = 0$) is given at a fixed point $a = A_t^h$ regardless of b, for any fixed time point $t \in [0, T]$. This is consistent with *Merton's CARA portfolio*:

$$A_t^* = \frac{\mu - r}{\gamma r \sigma^2}(1 - (1 - r)e^{-r(T-0)}), \tag{7.13}$$

see Table 1 in Section 6.1.

Second, the optimal portfolio A_t^h is smaller than A_t^*. For example, using the parameters of Table 3, *Merton's CARA portfolio* $A_0^* \approx 2.92$ but $A_0^h \approx 1.70$. This effect is probably due to the following two reasons:

1. The asymmetry in the model: all consumption takes place from the bank account, so stock must be sold (and transaction charges paid) before it can be realized for consumption. We observed this effect for proportional transaction costs too.

2. The combination of facts where the shift towards the optimal portfolio is decelerated and needs some time, and the optimal portfolio A_t^* in (7.13) decreases with time. Roughly speaking, the investor keeps his optimal amount of assets A slightly smaller in order to save future transaction costs, since the optimal portfolio decreases in time.

7.2.7 Model behavior w.r.t. the transaction costs factor c_2

Let us analyze the model's behavior with respect to the quadratic transaction costs factor c_2.

Figure 17 shows the optimal portfolio A_t^h at time $t = 0$ for several transaction costs factors c_2. The parameters are taken from Table 3 but with $T = 1$. We observe that the decrease of the optimal portfolio A^h w.r.t c_2 has a similar form as the (green) center line of the hold region for proportional transaction costs in Figure 9 in Section 6.3.

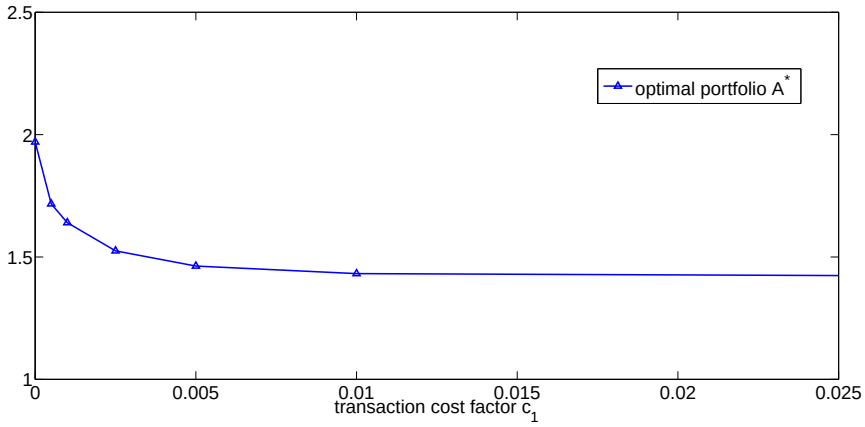

Figure 17: Optimal portfolio A_t^h for different transaction costs factors c_2 at $b = 0$ and time $t = 0$ with $h = 0.05$.

Figure 18 shows the optimal portfolio A_t^h over the full time horizon $t \in [0, 1]$ for several transaction costs factors c_2. Again, the parameters are taken from Table 3 but with $T = 1$. The optimal portfolio A_t^h drifts away from $A_t^\star$ backwards in time, or: $A_t^h \to A_t^*$ when $t \to T$. We already mentioned that this effect occurs because our terminal value $u(A_T + B_T)$ does not treat money in the assets A_T and money in the bank B_T differently. Moreover, on one hand, $A_t^h \to A_t^*$ when $c_2 \to 0$. On the other hand, A^h decreases very slowly when we increase $c_2 > 0.01$, see Table 5.

c_2	0.0	0.0005	0.001	0.0025	0.005	0.01	0.1	1.0
A_0^h	1.97	1.72	1.64	1.525	1.46	1.43	1.39	1.38

Table 5: Values of Figure 17.

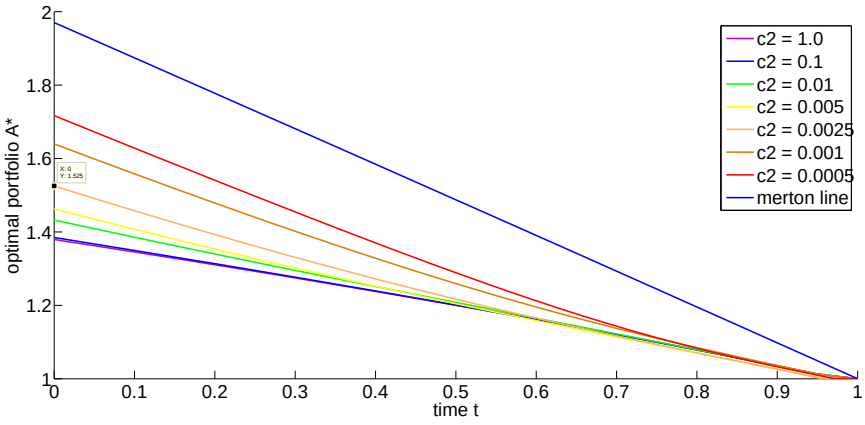

Figure 18: Evolution of the optimal portfolio A_t^* for different transaction costs factors c_2 over $t \in [0,1]$, for $h = 0.05$.

7.2.8 Model behavior w.r.t. other parameters

We also performed several studies in order to analyze the model's behavior due to changes on its parameters. We observed that the optimal consumption χ_t^h and the optimal portfolio A_t^h behave accordingly to their solutions for Merton's original model χ_t^* and A_t^*, respectively:

$$\begin{aligned} \chi_t^*(a,b) &= \tfrac{d(t)(a+b)}{\gamma} + \bar{a}(1 - e^{-r(T-t)}), \\ A_t^* &= \tfrac{\mu - r}{\gamma r \sigma^2}(1 - (1-r)e^{-r(T-0)}). \end{aligned} \tag{7.14}$$

Since we do not observe other effects, the results are not shown here.

7.2.9 Conclusions

In the pure quadratic transaction costs case:

- The value function $v(t,a,b)$ is increasing and concave w.r.t. a and b for all $t \in [0,T]$.
- The optimal consumption $\chi_t^h(a,b)$ behaves similar to Merton's original case (7.14). But, in comparison to χ_t^*, the **optimal consumption χ_t^h shows a reduction that is almost proportional to the quadratic transaction costs**. It is also shifted toward the optimal portfolio $A_t^h \leq A_t^*$.
- **The optimal allocation rate $\alpha_t^h(a,b)$ is continuous in (t,a,b)** and directs towards the optimal portfolio A_t^h. In the case of negative

exponential (CARA) utility, $\alpha_t^h(a,b)$ is independent of b and almost linearly w.r.t. a. Compared to the proportional transaction costs case, α_t^h **is discounted due to the temporary illiquidity costs**.

- The **optimal portfolio** A_t^h **is smaller** than *Merton's CARA portfolio* (7.14). The reasons are probably the **asymmetry in the model** (all consumption takes place from the bank account) and the decrease of the optimal portfolio A_t^h in time, combined with **discounted allocation** (future transaction costs are lowered).

- When c_2 is increased, the optimal portfolio A^h decreases in a similar form as the center line of the hold region for proportional transaction costs. In particular, $A_t^h \to A_t^*$ **when** $c_2 \to 0$ **and** A^h **decreases slower when** c_2 **is increased to higher values**. With respect to time, **the optimal portfolio** $A_t^h \to A_t^*$ **when** $t \to T$.

- The **parameter dependency** of the optimal consumption χ_t^h and the optimal portfolio A_t^h is **similar to Merton's original model** χ_t^* and A_t^* in (7.14), respectively.

7.3 Quadratic and proportional transaction costs

Let us briefly consider the previous quadratic transaction costs case but also add proportional transaction costs. Then the state system (7.2) becomes:

$$\begin{aligned} dA_t &= \left(\mu A_t + \alpha_t\right) dt \; + \; \sigma A_t \, dW_t, && A_0 = a, \\ dB_t &= \left(r B_t - \chi_t - \alpha_t - c_1 |\, \alpha| - c_2 \alpha^2 \right) dt, && B_0 = b. \end{aligned} \tag{7.15}$$

From our analyses in Section 6.3 and 6.4, we assume a smooth value function and the optimal controls to be given by:

$$\begin{aligned} \alpha^* &= \begin{cases} \frac{1}{2c_2} \frac{\partial_a v - (1+c_1)\partial_b v}{\partial_b v}, & \text{if } \partial_a v \geq (1+c_1)\partial_b v, \\ 0, & \text{if } (1+c_1)\partial_b v > \partial_a v > (1-c_1)\partial_b v, \\ \frac{1}{2c_2} \frac{\partial_a v - (1-c_1)\partial_b v}{\partial_b v}, & \text{if } \partial_a v \leq (1-c_1)\partial_b v. \end{cases} \\ \chi_t^* &= \dot{u}^{-1}\left(e^{\delta t} \partial_b v\right). \end{aligned} \tag{7.16}$$

Figure 19 shows the optimal allocation rate $\alpha_t^h(a,b)$ for $(a,b) \in [-2,3]\times[-2,2]$ at time $t = 0$. We observe the following two effects:

- On one side, there is a hold region where $\alpha_t^h = 0$ as in the case of pure proportional transaction costs. In particular, the optimal portfolio A_t^h from the case of pure quadratic transaction costs lies inside this hold region.

- On the other hand, in the sell and buy region, the allocation $\alpha_t^h > 0$ is not of bang-bang type as in the case of pure proportional transaction costs. It is rather discounted, continuous and almost linear w.r.t. a as in the case of pure quadratic transaction costs. At the boundaries of the sell and buy region, α_t^h goes to zero.

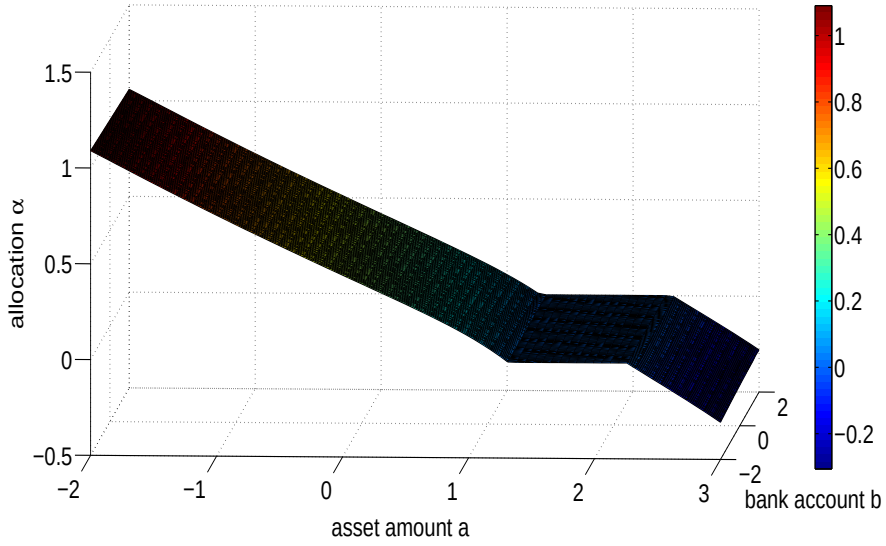

Figure 19: Optimal allocation rate $\alpha_t^h(a, b)$ for $(a, b) \in [-2, 3] \times [-2, 2]$ at time $t = 0$, considering linear and quadratic transaction costs $c_1 = 0.1$ and $c_2 = 0.1$.

7.3.1 Conclusions

The quadratic and proportional transaction costs case combines the segmentation into a buy, hold and sell region from pure proportional transaction costs and the smooth allocation rate from quadratic transaction costs.

7.4 Quadratic transaction costs with stochastic returns

Now let us add the stochastic return rate R_t to the problem. We still neglect the state and control constrains. Then the state system (7.2) becomes:

$$\begin{aligned} dA_t &= (\mu A_t + \alpha_t)\, dt \;+\; \sigma A_t\, dW_t, & A_0 &= a, \\ dR_t &= \big(\kappa(\bar{R} - R_t)\big)\, dt \;+\; \sigma_R\, dW_t^R & R_0 &= c, \\ dB_t &= (\, rB_t + R_t A_t - \chi_t - \alpha_t - c_2\alpha^2\,)\, dt, & B_0 &= b. \end{aligned} \tag{7.17}$$

According to our analyses in Section 6.4, we assume a smooth value function and the optimal controls to be given by equation (6.48):

$$\begin{aligned} \alpha_t^* &= \tfrac{1}{2c_2}\tfrac{\partial_a v - \partial_b v}{\partial_b v}, \\ \chi_t^* &= \dot{u}^{-1}\left(e^{\delta t}\partial_b v\right). \end{aligned} \tag{7.18}$$

Since we studied the pure quadratic transaction costs case extensively in Section 7.2, we only analyze the main differences to it at this time. We also use the same parameters as in Section 7.2.1, Table 3 but add the parameters for the return rate R_t. Table 6 shows these parameters.

T	γ	δ	r	μ	σ	c_2	M	$\bar{R}$	κ	σ_R	ρ
2	1	0.01	0.02	0.06	0.2	0.1	0.0	0.05	0.5	0.05	0.0

Table 6: Parameters of the basic set up for the one asset case.

7.4.1 Shape of solutions at time $t = 0$

Figure 20 shows the optimal consumption $\chi_0^h(a, b, 0)$ and the optimal allocation $\alpha_0^h(a, b, 0)$ for $(a, b) \in [1.6, 2.3] \times [-1, 1]$ and $c = 0$ at time $t = 0$. Overall, the shape of consumption and allocation do not change when adding returns. However, we observe a slightly b dependency in the the allocation rate α^h.

Figure 21 shows the optimal consumption $\chi_0^h(a, 0, c)$ and the optimal allocation $\alpha_0^h(a, 0, c)$ for $(a, c) \in [1.6, 2.3] \times [-0.1, 0.1]$ and $b = 0$ at time $t = 0$. We observe that the optimal consumption χ^h and allocation α^h increase with higher (initial) returns c. In particular, α_t^h seems to depend linearly on the initial return level c.

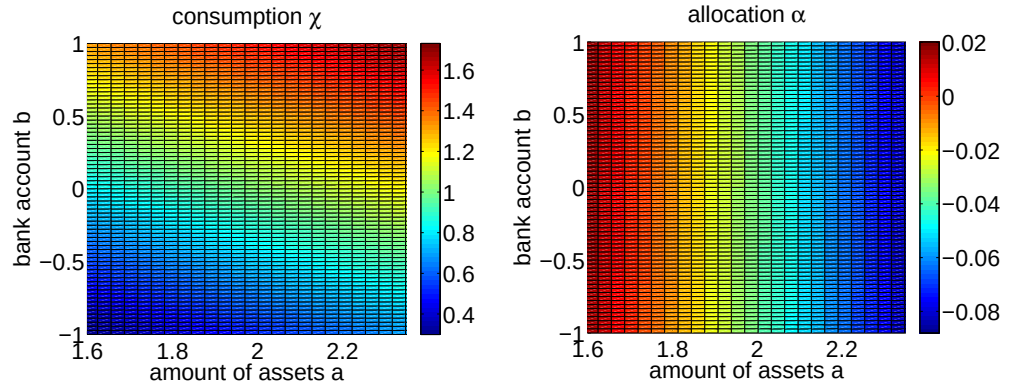

Figure 20: Optimal consumption $\chi_0^h(a,b,0)$ and the optimal allocation $\alpha_0^h(a,b,0)$ at $c=0$ and time $t=0$, for $h=0.1$.

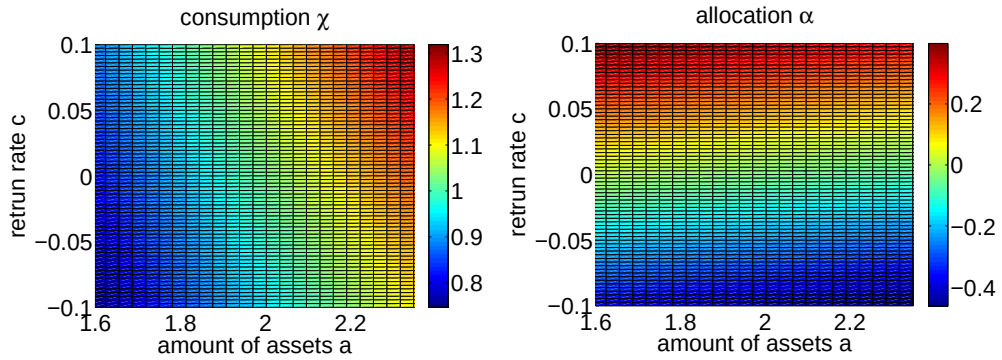

Figure 21: Optimal consumption $\chi_0^h(a,0,c)$ and the optimal allocation $\alpha_0^h(a,0,c)$ at $b=0$ and time $t=0$, for $h=0.1$.

However, it is interesting to note that the first derivative of the value function w.r.t. the amount of assets $\partial_a v^h$ does not vary much in c (initial return rate) but the first derivative of the value function w.r.t. the bank account $\partial_b v^h$ do vary in c, see Figure 22. Note that $\alpha^h \sim \partial_a v^h / \partial_b v^h$ from (7.18).

It means that the sensitivity of the value function w.r.t. the bank account depends on the return rates, while the sensitivity w.r.t. the amount of assets is less dependent on c. A possible reason for this counter intuitive phenomenon may be the asymmetry of the problem, namely that the consumption is taken out of the bank account only.

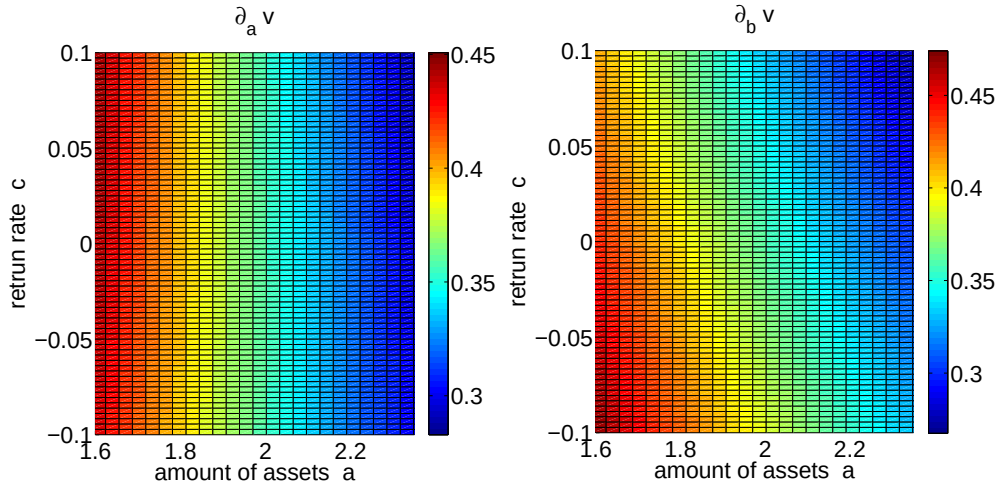

Figure 22: First derivatives $\partial_a v^h(a,0,c)$ and $\partial_b v^h(a,0,c)$ at $b = 0$ and time $t = 0$, for $h = 0.1$.

7.4.2 Model behavior w.r.t. the return's volatility σ_R

In this study we set the return's drift $\kappa(\bar{R} - R_t)$ to zero and variate only the volatility $\sigma_R > 0$. All other parameters are kept constant according to Table 6. Note that σ_R is the absolute volatility of the returns R_t in (7.17). In particular, it is not a relative volatility ($\sigma_R R_t$) as in the SDE for A_t.

Figure 23 shows the optimal allocation $\alpha_t^h(a,0,0)$ for all $a \in [0.5, 3.0]$, $b = 0$ and $c = 0$ for several values of σ at time $t = T/2$. We observe that a change of the return's volatility σ_R leads to a parallel shift of the optimal allocation rate α_0^h. A higher volatility σ_R lowers the allocation α_t^h and the optimal portfolio A_t^h ($\alpha_t^h = 0$). This behaviour of the model occur due to the risk aversion of the investor (concave utility function).

Figure 24 shows the optimal allocation $\alpha_t^h(A_t^*,0,c)$ for all $c \in [-0.15, 0.15]$, $a = A_t^*$ and $b = 0$ for several values of σ at time $t = T/2$. A_0^* is *Merton's CARA portfolio* at time $t = 0$. So it is the same situation as in Figure 23 but regarded from the c-dimension while a and b are fixed. We again observe that a change of σ_R leads to a parallel shift of the allocation rate α_0^h, also in this dimension. Moreovere, it seems that α_0^h in linear in c.

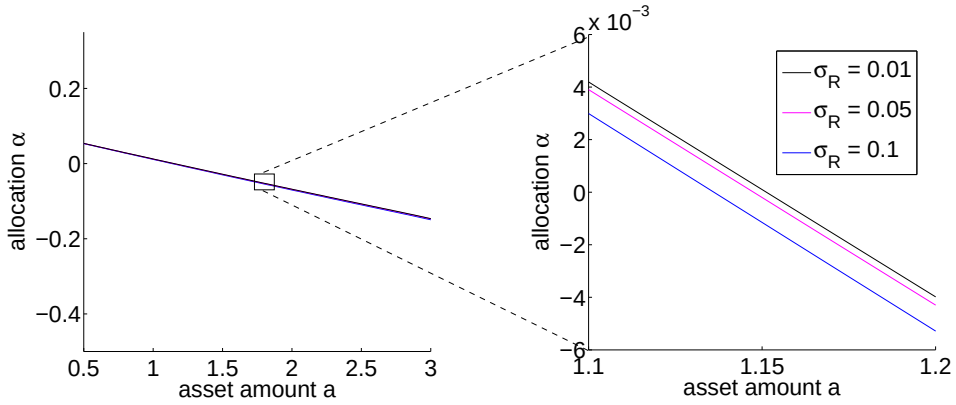

Figure 23: Optimal allocation $\alpha_t^h(a,0,0)$ for all $a \in [0.5, 3.0]$, $b = 0$ and $c = 0$ at time $t = T/2$, for $h = 0.1$.

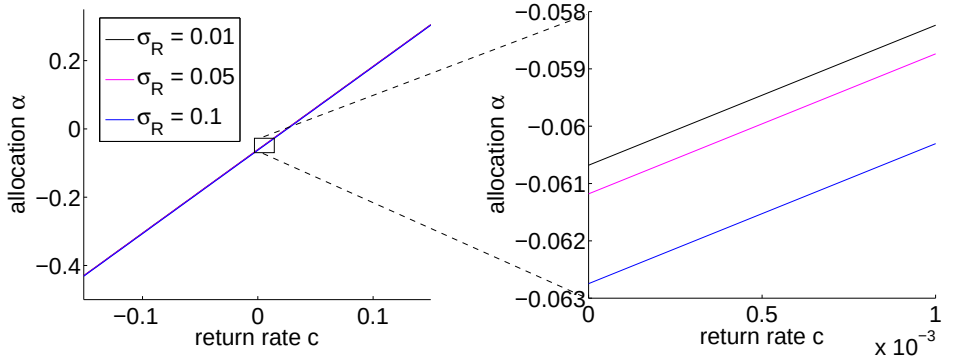

Figure 24: Optimal allocation $\alpha_t^h(A_t^*,0,c)$ for all $c \in [-0.15, 0.15]$, $a = A_t^*$ and $b = 0$ at time $t = T/2$, for $h = 0.1$.

7.4.3 Model behavior w.r.t. the return's mean-reversion level $\bar{R}$

In this study we variate the mean-reversion level $\bar{R} > 0$. All other parameters are kept constant according to Table 6. In particular we set the mean-reversion speed $\kappa = 0.5$ and the volatility $\sigma_R = 0.05$.

Figure 25 shows the optimal allocation $\alpha_t^h(a,0,0)$ for all $a \in [0, 2.5]$,$b = 0$, $c = 0$ and $\alpha_t^h(A_t^*,0,c)$ for all $c \in [-0.15, 0.15]$, $a = A_t^*$, $b = 0$ at time $t = T/2$. We observe that a change of the mean-reversion level $\bar{R}$ leads to a parallel shift of the allocation rate α_t^h. A higher mean-reversion level $\bar{R}$ leads to a higher allocation α_t^h and a higher optimal portfolio A_t^h ($\alpha_t^h = 0$).

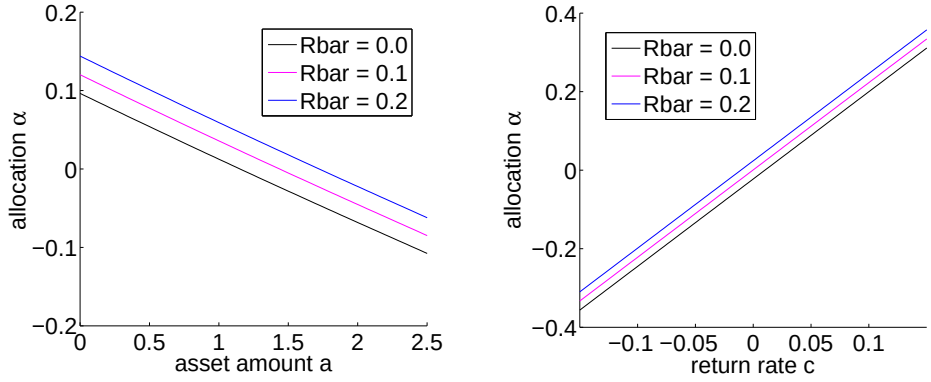

Figure 25: Optimal allocation $\alpha_t^h(a,0,0)$ for all $a \in [0,2.5]$ and $\alpha_t^h(A_t^*,0,c)$ for all $c \in [-0.15,0.15]$ at time $t = T/2$, for $h = 0.1$.

7.4.4 Model behavior w.r.t. the return's mean-reversion speed κ

In this study we vary the mean-reversion speed κ. All other parameters are kept constant according to Table 6. In particular we set the mean-reversion level $\bar{R} = 0.05$ and the volatility $\sigma_R = 0.05$.

Figure 26 shows the optimal allocation $\alpha_t^h(a,0,0)$ for all $a \in [0.5,2.5]$,$b = 0$, $c = 0$ and $\alpha_t^h(A_t^*,0,c)$ for all $c \in [-0.15,0.15]$, $a = A_t^*$, $b = 0$ at time $t = T/2$. On one hand, we observe that a change of the mean-reversion speed κ leads to a parallel shift of the allocation rate α_t^h when c is fixed. In particular, starting at $c = 0 < \bar{R}$ here, a higher mean-reversion speed κ leads to a higher allocation rate α_t^h and a higher optimal portfolio A_t^h.

On the other hand, we observe that a change of the mean-reversion speed κ leads to a tilt in the allocation rate α_t^h regarded from the c-dimension, see the right chart in Figure 26. The truning point is the mean-reversion level $\bar{R} = 0.05$ here.

Naturally, when starting at a $c > \bar{R}$, a higher mean-reversion speed κ leads to a faster drop of the expected returns towards $\bar{R}$ and therefore to a lower consumption. The other way round, when starting at a $c < \bar{R}$, a higher mean-reversion speed κ leads to a faster rise of the expected returns towards $\bar{R}$ and therefore to a higher consumption. This affects the allocation rates α_t^h and the optimal portfolio A_t^h in a synchron manner.

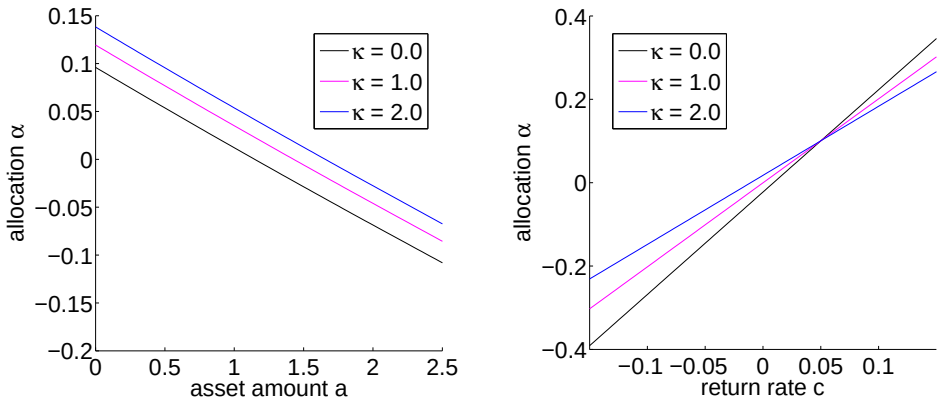

Figure 26: Optimal allocation $\alpha_t^h(a, 0, 0)$ for all $a \in [0, 2.5]$ and $\alpha_t^h(A_t^*, 0, c)$ for all $c \in [-0.15, 0.15]$ at time $t = T/2$, for $h = 0.1$.

7.4.5 Conclusions

In the case of quadratic transaction costs with stochastic returns:

- The optimal consumption χ^h and allocation α^h increase with higher returns c. In particular, α^h seems to depend linearly on the (initial) return level c.
- The **optimal allocation α_t^h and optimal portfolio A_t^h decreases with higher return volatility** σ_R. This is intuitively due to the risk aversion of the investor (concave utility function).
- The value function v, the optimal consumption χ_t^h, the optimal allocation α_t^h and the optimal portfolio A_t^h increases with higher mean-reversion level $\bar{R}$.
- A change of the mean-reversion speed κ leads to a tilt in the allocation rate α_t^h. The turning point is the mean-reversion level $\bar{R}$.

7.5 Quadratic transaction costs with optimized returns

By *optimized returns* we refer to the term $\max(R_t, 0) - M$ as a simplified model of optimized flexible operation of production assets. The derivation from first principles was provided in Section 5.2.6. Note also our discussion about the terminology *optional* returns there.

Optimized returns basically means that the total losses from an asset are bounded below at the asset's fixed costs $-M$. To make this clearer, let us split the total return (loss) R_{total} into the return from operation R_{op} and the fixed cost M: $R_{total} = R_{op} - M$. Then, the optimized outcome of flexible operation is: $\max(R_{op}, 0) - M = \max(R_{total}, -M)$. So the total return (loss) is bounded below at $-M$.

Now let us consider the optimized stochastic return case. Let $M \geq 0$ denote the proportional fixed costs and let R_t denote the return from operation. We still neglect the state and control constrains. Then the state system (7.2) becomes:

$$\begin{aligned} dA_t &= \left(\mu A_t + \alpha_t\right) dt \; + \; \sigma A_t \, dW_t, && A_0 = a, \\ dR_t &= \left(\kappa(\bar{R} - R_t)\right) dt \; + \; \sigma_R \, dW_t^R && R_0 = c, \\ dB_t &= \left(r B_t + \max(R_t, 0) A_t - M A_t - \chi_t - \alpha_t - c_2 \alpha^2 \right) dt, && B_0 = b. \end{aligned} \tag{7.19}$$

In the case of optimized returns, we can not assume that the value function $v(t, a, b, c)$ is continuously differentiable at the critical point $c = 0$, see our discussion about this issue in Section 6.2.3. In consequence, the second order HJB equation can not be utilized here. Nevertheless, from our analyses in Section 6.4 and 6.4.1, we assume that the optimal controls are still given by equation (6.48) as in the non-optimized stochastic return case:

$$\begin{aligned} \alpha_t^* &= \tfrac{1}{2c_2} \tfrac{\partial_a v - \partial_b v}{\partial_b v}, \\ \chi_t^* &= \dot{u}^{-1} \left(e^{\delta t} \partial_b v \right). \end{aligned} \tag{7.20}$$

The reason is that the return rate R_t is not controlled (the SDE for R_t does not depend on α nor on χ) and therefore R_t plays no role in the optimization of α and χ in (7.20). Moreover, the violation of the HJB condition (the term $\max(c, 0)$ is not differentable at $c = 0$) does not harm the CFB algorithm, which is based on the Stochastic Maximum Principle. In particular, the CFB algorithm does not rely on any partial derivatives. So it does not struggle with the non-differentiability of the value function at $c = 0$. The assumption (7.20) is also supported by the numerical results that we obtained.

Since we studied the pure quadratic transaction costs case extensively in Section 7.2 and the stochastic return case in Section 7.4, we present only the main differences to these cases here. We use the same parameters as before, see Section 7.4 Table 6. In particular, we set the fixed costs $M = 0$ for the first analyses.

7.5.1 Higher consumption and portfolio value

Figure 27 shows the difference of the value functions $v(0, a, 0, c)$ and the optimal consumption rates $\chi_0(a, 0, c)$ between the optimized return case and the stochastic return case, at time $t = 0$. Note that $M = 0$ in both cases. We observe that the consumption (and therefore the value function) is higher in the optimized return case. The differences are small for $c >> 0$ but increase when c decreases below zero.

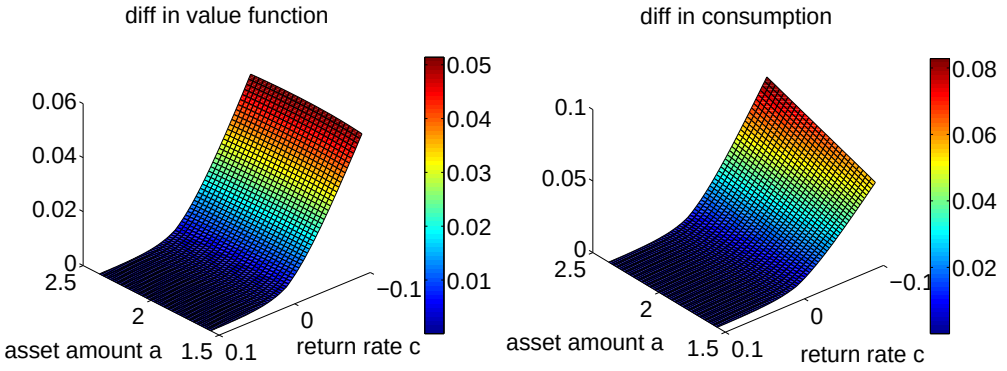

Figure 27: Difference of the value functions $v(0, a, 0, c)$ and the optimal consumption rates $\chi_0(a, 0, c)$ between the optimized return case and the stochastic return case at $t = 0$, for $h = 0.1$.

7.5.2 Flattened allocation and a discontinuity in $\partial_c v$

Figure 28 shows the optimal allocation rate $\alpha_0^h(a, 0, c)$ and the first derivative of the value function w.r.t. the return rate $\partial_c v(0, a, 0, c)$, at time $t = 0$.

We observe that for initial return rates c close to zero, the allocation rate α_0^h is not linear in a (amount of assets) anymore, as it was in the stochastic return case in the previous section, see Figure 21. Instead, α_0^h flattens out when the initial return rate $c < 0$.

We also observe that there is a discontinuity in $\partial_c v$ at $c = 0$. The reason is that the evolution of the bank account in (7.19) is not differentiable w.r.t. R_t at $R_t = 0$. Nevertheless, the other partial derivatives $\partial_a v$ and $\partial_b v$ and both controls α_t^h and χ_t^h are still smooth as in the previous cases. This fact supports our assumption (7.20) that the controls do not depend on $\partial_c v$.

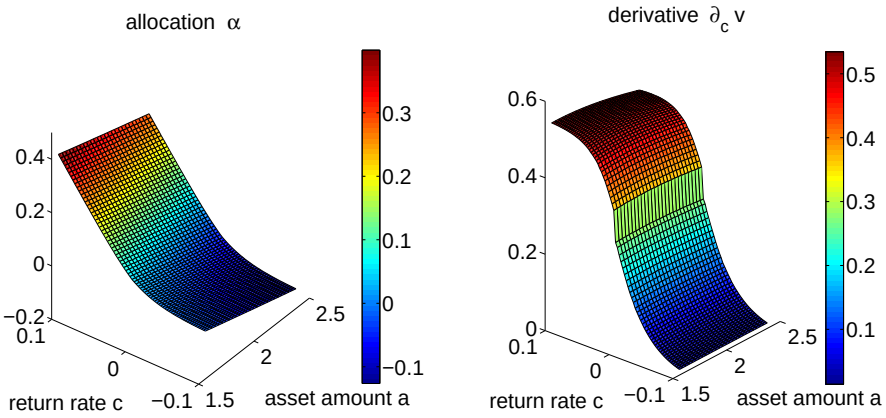

Figure 28: Optimal allocation rate $\alpha_0^h(a, 0, c)$ and the first derivative $\partial_c v(0, a, 0, c)$, at $b = 0$ and time $t = 0$, for $h = 0.05$.

Figure 29 shows the optimal allocation rate $\alpha_0^h(A_0^\star, b, c)$ and the first derivative of the value function w.r.t. the return rate $\partial_c v(0, A_0^\star, b, c)$, at time $t = 0$. Here, the amount of assets a is fixed at *Merton's CARA portfolio*. So it shows the same situation as Figure 28 but in the dimensions (b, c). Again, the allocation rate α_0^h flattens out when the initial return rate $c < 0$ and we observe a discontinuity in $\partial_c v$ at $c = 0$.

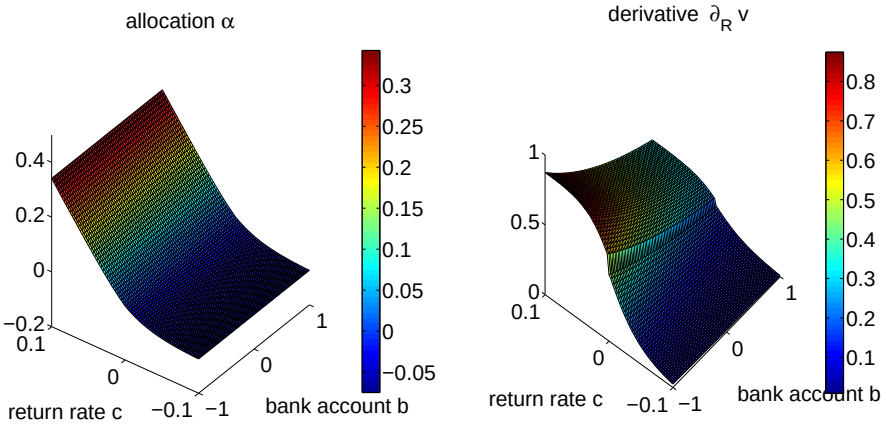

Figure 29: Optimal allocation rate $\alpha_0^h(A_0^\star, b, c)$ and the first derivative $\partial_c v(0, A_0^\star, b, c)$, at $a = A_0^\star$ and time $t = 0$, for $h = 0.05$.

7.5.3 Model behavior w.r.t. the return's volatility σ_R

In this study we variate the volatility $\sigma_R > 0$. We set the return's drift $\kappa(\bar{R} - R_t)$ to zero in order to focus on the parameter σ_R and get clearer results. All other parameters are kept constant according to Table 6. Note that σ_R is the absolute volatility of the returns R_t in (7.19). In particular, it is not a relative volatility ($\sigma_R R_t$) as in the SDE for A_t. In consequence, assuming an average return of 0.05, a volatility of 0.05 would be 100%.

The left chart of Figure 30 shows the optimal allocation $\alpha_t^h(a, 0, 0)$ for several values of σ_R, for $a \in [0, 2.5]$, $b = 0$ and $c = 0$ at time $t = T/2$. We observe that a change of the return's volatility σ_R leads to a parallel shift of the allocation rate α_0^h in the a-dimension. In contrast to the non-optimized stochastic return case (see Figure 24 in Section 7.4.2), a higher volatility σ_R increases the allocation α_t^h and the optimal portfolio A_t^h (and also the optimal consumption χ_t^h and value function v). This may be counter intuitive considering the risk aversion of the investor.

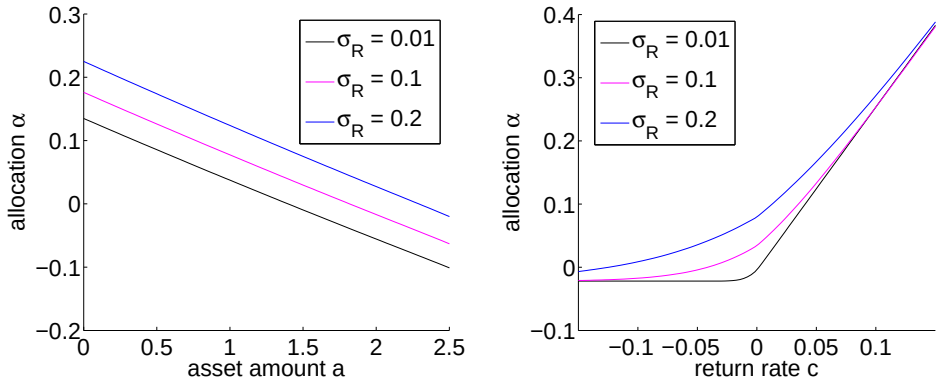

Figure 30: Optimal allocation $\alpha_t^h(a, 0, 0)$ for all $a \in [0.5, 3.0]$, $b = 0$ and $c = 0$ at time $t = T/2$, for $h = 0.1$.

The right chart of Figure 30 shows the optimal allocation $\alpha_t^h(A_t^*, 0, c)$ for several values of σ_R, for $c \in [-0.15, 0.15]$, $a = A_t^*$ and $b = 0$ at time $t = T/2$. Again, a higher volatility σ_R increases the allocation α_t^h.

Taking a closer look at the chart in the c-dimension, we observe that the curves are comparable to the value of a call option on R_t with strike zero. So the char reveals the intrinsic option $\max(R_t, 0)$ on the asset's returns.

Remember, this feature originates from the flexible operation of production assets, see Section 5.2.6. Note also our discussion about the terminology *optional* returns there.

An interesting fact is, that this feature does not only occur in the shape the value function v (value of the portfolio/option), rather it is transfered to the shape of the optimal allocation rate α_t too.

Adding fixed costs

Now we perform the same study on σ_R but change $M = 0.0 \mapsto 0.03$ and $\bar{R} = 0.05 \mapsto 0.08$. So the average total return $R_t - M$ does not change rather the loss boundary changed from $0.0 \mapsto -0.03$. This case represents a production asset with (high) fixed costs M.

Figure 31 shows the optimal allocation $\alpha_t^h(a, 0, 0)$ and $\alpha_t^h(A_t^\star, 0, c)$ for several values of σ_R at time $t = T/2$. On one hand, we observe the same effects as we did for the case without fixed costs (Figure 30). On the other hand, the values of $\alpha_t^h(a, 0, 0)$ and $\alpha_t^h(A_t^\star, 0, c)$ are shifted downwards. The shift seems to be almost parallel. The reason is that the value of the intrinsic option is less worth since losses are possible up to an amount of M.

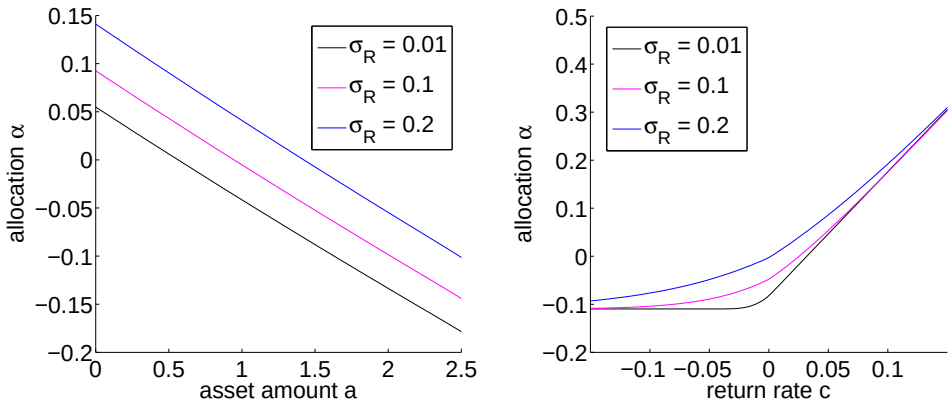

Figure 31: Optimal allocation $\alpha_t^h(a, 0, 0)$ for all $a \in [0.5, 3.0]$, $b = 0$ and $c = 0$ at time $t = T/2$, for $h = 0.1$.

In order to compare the right charts of Figure 30 and Figure 31, we need to shift the latter chart from operation returns R_t to total returns $R_t - M$:

$c \mapsto c - M = -0.03$. Then, roughly speaking, the strike of the intrinsic option is at $c = -M$ and its payoff for $c < -M$ is negative since money is lost.

7.5.4 Conclusions

In the case of quadratic transaction costs with optimized returns:

- The optimal consumption χ^h and allocation α^h is higher than in the non-optimized stochastic return case.
- The optimal allocation α_0^h flattens out when the initial return rate $c < 0$ and there is a discontinuity in $\partial_c v$ at $c = 0$.
- **A higher volatility σ_R increases the the optimal consumption χ^h, the value function v, the optimal allocation α_t^h and the optimal portfolio A_t^h.** The reason is the operation flexibility (intrinsic option $\max(R_t, 0)$) of the asset.
- **The positive effect of higher volatility is reduced for assets with (high) fixed costs M.**

7.6 Considering constraints

Last but not leat, we analyse the constrainted case of Problem 3 (7.6) - (7.8). In order to focus on the constraints, we consider the pure quadratic transaction costs case which we studied extensively in Section 7.2. So let us neglect the return rate and the fixed costs $(\kappa, \bar{R}, \sigma_R, c, M = 0)$. Then the state system (7.2) becomes:

$$\begin{aligned} dA_t &= (\mu A_t + \alpha_t)\, dt \;+\; \sigma A_t\, dW_t, \qquad A_0 = a, \\ dB_t &= (\, rB_t - \chi_t - \alpha_t - c_2\alpha^2\,)\, dt, \qquad B_0 = b. \end{aligned} \tag{7.21}$$

According to our analyses in Section 6.4, we assume a smooth value function and the optimal controls to be given by equation (6.48):

$$\begin{aligned} \alpha_t^* &= \tfrac{1}{2c_2}\tfrac{\partial_a v - \partial_b v}{\partial_b v}, \\ \chi_t^* &= \dot{u}^{-1}\left(e^{\delta t}\partial_b v\right). \end{aligned} \tag{7.22}$$

7.6.1 State constraints

The state constraint of Problem 3 is the short-selling prohibition:

$$A_t \geq 0. \tag{7.23}$$

This constraint is automatically satisfied whenever the optimal holdings in the asset $A_t^* \geq 0$. The reason is the following. When the investor do not allocate assets ($\alpha = 0$), then the SDE (7.21) is log-normally distributed and never attains negative values. So short-selling will not occur when the portfolio is left unchanged. Additionally, as we saw in all previous studies, the optimal allocation α_t^* shifts the amount of assets A_t in the direction towards the optimal portfolio $A_t^* \geq 0$. So whenever the parameter setup fulfill $A_t^* \geq 0$, the optimal allocation rate α^* keeps the amount of assets A_t positive.

On one hand, all our parameter setups in the previous studies fulfill this assumption, as we see a posteriori. On the other hand, we can not predict the optimal portfolio A^* a priori. Therefore, we simulate Problem 3 also for initial amounts of assets $a < 0$. Note that the short-selling prohibition purely comes for the real world application of production (physical) assets and is not a technical constraint of the model. So our numeric scheme does not need any adjustment for $a < 0$.

7.6.2 Control constraints

The control constraints of Problem 3 are:

$$\alpha_t \in [\alpha_{min}, \alpha_{max}], \quad \chi_t \in [0, \chi_{max}], \tag{7.24}$$

In contrast to the short-selling constraint, the control constraints are not given through the problem setup. They are rather technical constraints and guarantee the boundedness of the controls. In all previous studies, the optimal controls $\alpha_t(a, b, c)$ and $\chi_t(a, b, c)$ are continuous (smooth) functions w.r.t. a, b, c and t. Therefore, α_t and χ_t are bounded whenever the initial values a, b, c and the time horizon T are finite. Moreover, whenever we kept parameter values reasonable in the previous studies, the optimal consumption χ_t stayed positive.

In summary, we can neglect the state (short-selling) and the control constraints in our numerical scheme. They are automatically fulfilled whenever we choose reasonable parameters.

7.6.3 Bankruptcy constraints

We discussed the bankruptcy constraints for the time-discrete model in Section 4.6 and for the continuous-time model in Section 5.2.7. The result is the following exit strategy in order to prevent bankruptcy with uncertain losses:

$$(\alpha_t, \chi_t) = \begin{cases} (-\alpha_{bry}, 0), & \text{if } (A_t, B_t) \notin \mathcal{L}_t(\alpha_{bry}) \wedge A_t > 0, \\ (0, 0), & \text{if } (A_t, B_t) \notin \mathcal{L}_t(\alpha_{bry}) \wedge A_t = 0, \end{cases} \tag{7.25}$$

where:

$$\mathcal{L}(\alpha_{bry}) = \{(x, y) \in \mathbb{R}_+ \times \mathbb{R} \mid y + (1 - c_2\,\alpha_{bry})\,x \geq 0\}. \tag{7.26}$$

Figure 32 shows the optimal consumption $\chi_t^h(a, b)$ and allocation $\alpha_t^h(a, b)$ at time $t = 0$ for the constrained case with $c_{bry} = 0.2$. One can clearly see the region where the exit strategy is forced to be exercised. The other region seems to be similar to the unconstrained case, see Figure 11.

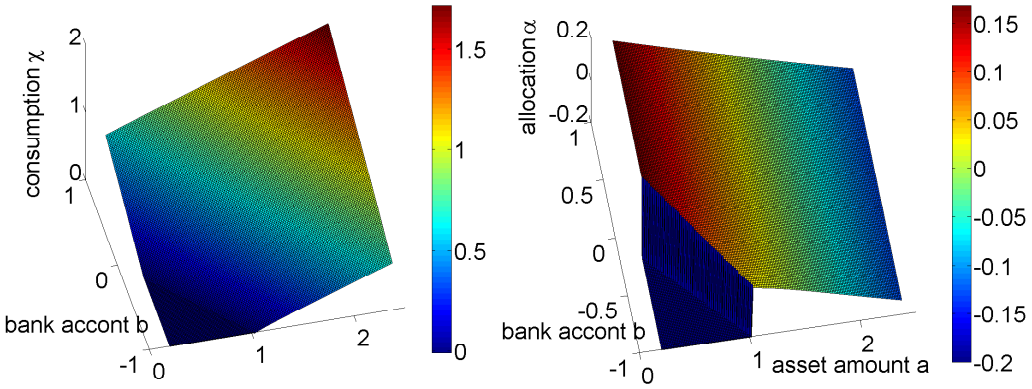

Figure 32: Optimal consumption $\chi_t^h(a, b)$ and allocation $\alpha_t^h(a, b)$ at time $t = 0$, for $h = 0.07$.

Figure 33 shows the differences of the optimal consumption $\chi_t^h(a, b)$ and allocation $\alpha_t^h(a, b)$ between the unconstrained and the constrained case at time $t = 0$. We observe that consumption and allocation is reduced close to the exit region in order to preserve an entering into it.

7.6.4 Conclusions

Considering constrains for the case of quadratic transaction costs the consumption and allocation is reduced close to the exit region in order to preserve an entering into it.

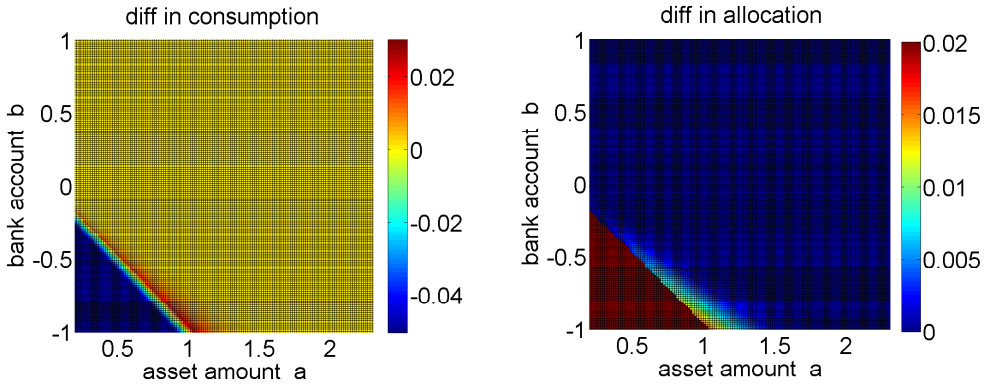

Figure 33: Differences of optimal consumption $\chi_t^h(a,b)$ and allocation $\alpha_t^h(a,b)$ between the unconstrained and the constrained case at time $t = 0$.

7.7 Considering correlation

Since the return rate R_t is not controlled (i.e., it does not depend on α nor χ), the optimal controls:

$$\begin{aligned} \alpha_t^* &= \frac{1}{2c_2}\frac{\partial_a v - \partial_b v}{\partial_b v}, \\ \chi_t^* &= \dot{u}^{-1}\left(e^{\delta t}\partial_b v\right). \end{aligned}$$

do not depend on the correlation ρ, see our discussion in Section 6.4.1. Therefore we do not analyze the effect of correlations here.

7.8 Summary

We introduced the quadratic transaction costs as second order Taylor expansion of temporary illiquidity costs in Section 3.4.2. In contrast to the bang-bang control and the hold region under proportional transaction costs, the allocation strategy under quadratic transaction costs is smooth and the optimal portfolio A_t^* is a sharp curve in the (a,b)-space for all $t \in [0,T]$. On the other hand, the model's behavior is very similar to the original Merton's model without transaction costs.

The optimal allocation rate $\alpha_t^*(a,b)$ is linear in a, independent of b and sensitive to the transaction cost factor c_2. The optimal portfolio A_t^* is smaller than in Merton's original case due to the asymmetry of the problem. The optimal consumption is reduced proportional to the transaction costs $c_2{\alpha_t^*}^2$.

The introduction of stochastic returns shift the solutions v, χ^* and α^* upwards as long as the mean return $\bar{R}$ is positive. On the other hand, a higher return volatility σ_R decreases the values due to the risk aversion of the investor (concave utility function). In contrast, when the stochastic returns are optional $\max(R_t, -M)$, a higher volatility is favorable as long as the asset's fixed costs M are small.

Enforcing an exit strategy to prevent bankruptcy, the allocation rate and the consumption rate are reduced close to the exit region in order to preserve the entering into it.

7.8.1 Economical applications

The two unique aspects of our model are the consideration of temporary illiquidity (quadratic transaction) costs and the optimized returns. Possible real world applications are huge transactions and peak power plants.

The temporary illiquidity effect is very important for transactions at fragmented (commodity) markets (see Section 3.4.2). In reality, bigger transactions are placed by computer based algorithms. These algorithms use slow (smooth) allocation rates to avoid high costs for illiquidity that are only triggered due to the allocation process itself. Disregarding illiquidity costs could easily lead into a new Flash Crash [Wallstreet 2010].

A peak-power plant (see Section 2) is a perfect example of an production asset with optimized returns and low fixed costs. Those plants have usually high production costs and makes their profits only through the high volatility of power markets. The investment strategy in those assets should be considered according to the volatility.

8 Numerical results for the case of two generation assets

In this section we present a multi-dimensional case of our continuous-time model for optimal portfolio allocation of commodity related assets, Problem 2 (5.13)-(5.16). In particular, we consider a portfolio of generation assets. Portfolio theory is applicable to power generation because: 1) energy sources and power are homogeneous goods[30] and 2) power generation is smoothly scalable because generation assets of the same technology have similar properties.

The purpose of this section is to show the applicability of our continuous-time model for optimal portfolio allocation of commodity related assets to multiple asset portfolios. Our controlled forward-backward (CFB) algorithm from Part III provides us the ability (in terms of performance) to study this multi-dimensional convex stochastic optimal control problem numerically. A detailed study of a certain business case is left to further research.

In Section 8.1 we give a short overview of recent literature on generation portfolios. In Section 8.2 we introduce different types of generation assets and their properties. Then we state the problem for two assets in Section 8.3. The results are presented in Section 8.4.

8.1 Literature review

[Awerbuch, Burger 2003] firstly applied a one time step Mean-Variance Portfolio (MVP) model[31] to generation assets. The assets returns are the inverses of the expected levelised generation costs. The variances are measured through the standard deviation of historic annual outlays for fuel, operation & maintenance and construction period costs. So these outlays are assumed to be normally distributed. [Awerbuch, Stirling, Jansen, Beurskens 2005] combine the MVP approach with a multi-criteria diversity index (developed by [Stirling 1994]) to study a so called *full-spectrum risk measure*. They point out the importance of renewable energy for risk reduction of a nation wide generation portfolio.

[30] See preliminaries, Section 2 and 1.1.1.
[31] See preliminaries, Section 3.2.

[Roques, Newbery, Nutttall 2007] propose the following two-step approach to find the optimal generation portfolio. In the first step they calculate the Net Present Values (NPV) of several generation assets by using a Discounted Cash Flow (DCF) method[32] and Monte Carlo simulations. The DCF method is based on probability distributions for fuel prices, carbon taxes and power prices, and a constant average operation level that is derived through the peak load pricing concept[33]. In the second step they measure the mean and the variance from these simulations in order to apply a one time step MVP analysis on $\frac{\text{NPV}}{\text{capacity}}$. They found out that "high degrees of correlation between gas and electricity prices - as observed in most European markets - reduce the risk of gas power plants and make portfolios dominated by gas power plants more attractive". The high correlation is an effect of the peak load pricing concept, where the power price is derived from the highest marginal costs among all power plants. Typically gas power plants have the highest marginal costs.

[Huang, Wu, 2007] are searching the optimal generation mix for Thailand with an one time step MVP model. The goal is to minimize total generation costs for the country, instead of maximizing profits for a company. In particular, their model integrates portfolio theory into an electricity planning framework: each power plant's yearly output is derived through a the load duration curve[34] (fixed step function) that satisfy the countries's demand. The merit order[35] is fixed. Fuel prices, technological changes and capital cost are assumed to be normally distributed. Additional considerations are transmission loss, average utilization rate, average out of service times for maintenance and gas import constraints. Similar to [Awerbuch, Burger 2003], they found that "replacing fossil fuel with renewable energy helps reduce generating cost risk".

[Weber, Sunderkoetter 2009] also apply a one time step MVP analysis to generation assets in an electricity planning (peak-load pricing) framework. Their approach combine conceptual elements of peak-load pricing and MVP, where the operation level of each technology is derived from a biding process on the load duration curve. The load duration curve is fixed and the merit order curve excludes reversals. The source of uncertainty are the fuel prices. Their main contribution is an analytical solution of the problem but it provides less new insights for practice.

[32] See preliminaries, Section 1.4.1.
[33] See preliminaries, Section 2.3.1.
[34] See preliminaries, Section 2.3.3.
[35] See preliminaries, Section 2.3.1.

[Delarue 2010] considers a one time step MPV model that minimizes total operation costs. Both, the installed capacity and the generation levels of each technology are optimized in this so-called *integrated portfolio investment model*. In particular, the mean and the variance of the MVP model are sums of means and variances, respectively, to account for different time periods of the year. In order to do so, fuel costs, hourly demand, wind output and the values for certain risk categories are deterministic from historical data for each period. The approach explicitly accounts for dispatch constraints and thus emphasis the variability of wind power. However, he found out that the risk reduction through the installation of wind power is smaller then proposed in earlier papers. The reason is the requirement of sufficient fast rampable technologies to compensate the wind fluctuation.

According to our literature research, all found portfolio analyses for generation assets are one time step approaches. In contrast, our discrete-time model in Section 4 and our continues-time model in Section 5 are multi step approaches that start with the current portfolio and consider the allocation process in the future.

8.2 Types of assets

In order to keep the number of dimensions low, we need to cluster generation assets into categories. According to [Burger, Graeber, Schindlmayr 2007] page 183, a clustering of generation assets into base load, peak load and renewables is advisable when considering fundamental market models for power prices, see Section 2.3. Therefore, let $\mathbb{J} := \{\mathcal{J}^{\text{base}}, \mathcal{J}^{\text{peak}}, \mathcal{J}^{\text{renew}}\}$ denote the set of generation assets available in the market, where:

$$\begin{aligned} \mathcal{J}^{\text{base}} &:= \{\text{nuclear, brown-coal, black-coal, ...}\}, \\ \mathcal{J}^{\text{peak}} &:= \{\text{gas, GCCT}^{36}\text{, biomass, ...}\}, \\ \mathcal{J}^{\text{renew}} &:= \{\text{wind, solar, hydro, ...}\}. \end{aligned} \tag{8.1}$$

The assets in each category have different technologies but the generation properties of theses technologies are similar:

1. **Base load** power plants, by definition, provide the basic all-time demand. In particular, their operation does not depend on the daily demand periods. Base load plants have stable return rates. One reasons is that their fuel's market prices (Uranium, Lignite and Anthracite) are relatively stable compared to oil or gas. Moreover, in most markets base load plans provide the biggest share of power and the average

power price is likely to follow their operation costs. The fixed costs for operation & maintenance and construction are relatively low due to long lifetimes and permanent operation.

One the other side, base load power plants face high political and economical risks. Coal plants produce a high amount of emissions. Nuclear plants produce nuclear waste and face a risk of reactor melt-down. In recent years, authorities introduced taxes for the emission of pollutants and fees on nuclear waste. In some countries (like the U.K.) authorities even limit the amount of nuclear or coal power produced. Moreover, operation failures or safety issues could lead to a forced phase out of a power plant. In consequence, considering the high construction and de-construction costs, the monetary amount of base load assets are volatile even they have stable returns.

2. **Peak load** power plants are usually highly flexible natural gas plants or biomass plants. On one hand, their returns are more uncertain than base load returns because of higher fuel (gas) price volatility and uncertain peak demand. On the other hand, the plants produce less emissions and face less political risks.

3. **Renewables** like wind, solar, and riverside hydro power plants have no fossil fuel input and negligible operation costs. The operation of renewables depend on weather effects like natural water flow from melting snow, wind penetration or sunshine. In consequence, renewables operates whenever they can because it would be a wastage of energy to shut off renewable power generation. In some countries (like Germany) it is even prohibited by law to intervene into renewable power generation. Nevertheless, the average annual insolation, wind intensity and amount of snow fall are quite stable. Together with government guaranteed feed-in prices, the returns of renewables are quite stable.

   On the other hand, renewables are less efficient compared to their construction costs and their support in society are not guaranteed: water-pump-reservoirs destroy landscape, solar panels take farming land away and wind-turbines are ugly in the landscape, loud near cities and influencing the wild life.

The above properties are justified in a heuristic way. Nevertheless, they are carefully chosen for different sources to approximate the reality properly and are mostly consistent with [Burger, Graeber, Schindlmayr 2007]. Table 7 summarizes the properties with relation to the parameter of our model, .

	total retrun $\bar{R} - M$	retrun volatility σ_R	asset volatility σ
base load	high	low	high
peak load	medium	high	low
renewables	low	low	medium

Table 7: Properties of generation assets in relation to each other.

8.3 Problem statement for base and peak load plants

In this thesis we consider the case of two generation assets, namely a general base load and a general peak load asset. The analysis for more assets are open to further research. So let A_t^B and A_t^P denote the monetary amount of base load and peak load assets in the portfolio, respectively, and let us recall Problem 2 (5.13)-(5.16) with negative exponential utility for this two assets case.

Problem 4 (Portfolio problem in case of two generation assets)
Let the objective function be given by:

$$\begin{aligned} &J(t, a^B, a^P, b\,, c\,, \alpha^B, \alpha^P, \chi) = \\ &E_t^{\alpha^B, \alpha^P, \chi}\left[\int_t^T e^{-\delta s} \tfrac{1-e^{-\gamma \chi_s}}{\gamma}\, ds \;+\; e^{-\delta T} \tfrac{1-e^{\gamma(A_T^B + A_T^P + B_T)}}{\gamma} \;\middle|\; a^B, a^P, b, c \right], \end{aligned} \tag{8.2}$$

and let the states evolve according to:

$$\begin{aligned} dA_t^B &= \left(\mu^B A_t^B + \alpha_t^B\right) dt \;+\; \sigma^B A_t^B\, dW_t^B, && A_0^B = a^B, \\ dA_t^P &= \left(\mu^P A_t^P + \alpha_t^P\right) dt, && A_0^P = a^P, \\ dR_t &= \left(\kappa(\bar{R}^P - R_t)\right) dt \;+\; \sigma_R\, dW_t^R && R_0 = c, \\ dB_t &= \Big(r B_t + (\bar{R}^B - M^B) A_t^B + \max(R_t - M^P, -M^P) A_t^P \\ &\quad -\chi_t - \alpha_t^B - \alpha_t^P - c_2^B \alpha^{B\,2} - c_2^P \alpha^{P\,2}\Big)\, dt, && B_0 = b, \end{aligned} \tag{8.3}$$

where:

$$dW_t^B\, dW_t^R \;= 0. \tag{8.4}$$

We search for the optimal allocation policies α^{B*}, α^{P*} and the optimal consumption policy χ^*:

$$\begin{aligned}&(\alpha^{B*}(t,a^B,a^P,b,c),\alpha^{P*}(t,a^B,a^P,b,c),\chi^*(t,a^B,a^P,b,c)) = \\ &\arg\max_{(\alpha^B,\alpha^P,\chi)} J(t,a^B,a^P,b\,,c\,,\alpha^B,\alpha^P,\chi),\end{aligned} \tag{8.5}$$

and the value function v:

$$\begin{aligned}v(t,a^B,a^P,b,c) = &\max_{(\alpha^B,\alpha^P,\chi)} J(t,a^B,a^P,b\,,c\,,\alpha^B,\alpha^P,\chi),\\ v(T,a^B,a^P,b,c) = &\; g(a^B,a^P,b,c),\end{aligned} \tag{8.6}$$

for any given initial time $t \in [0,T]$.

In Problem 4 all parameters are given as positive constants: $T, r, \delta, \gamma \geq 0$, $\mu^B, \mu^P, \sigma^B, c_2^B, c_2^P, \bar{R}^B, \bar{R}^P, M^B, M^P \geq 0$ and $\kappa, \sigma_R \geq 0$. Note that we neglect the state and control constraints in this study. Note also that the peak load asset's volatility as well as the correlations are assumed to be zero. The base load's return rate is constant. This setup is consistent with our heuristic analysis of Table 7.

8.3.1 Basic parameter set up

Our region of interest is:

$$\begin{array}{ll} a^B, a^P \in [0,2], & \text{initial monetary amount of assets in €,}\\ b \in [-1,1], & \text{initial monetary amount in bank account in €,}\\ c \in [-0.2,0.2], & \text{initial return of asset in 100\%.}\end{array} \tag{8.7}$$

The model parameters are defined in Table 8. Note that all parameters are constant over time $[0,T]$.

T	γ	δ	r	μ^B	σ^B	c_2^B	$\bar{R}^B$	M^B
2	1	0.01	0.02	0.02	0.2	0.1	0.05	0.0

μ^P	c_2^P	$\bar{R}^P$	M^P	κ	σ_R	ρ		
0.02	0.1	0.06	0.02	0.5	0.05	0.0		

Table 8: Parameters of the two generation assets case.

The market in Table 8 has a low interest rate r, a low discounting rate δ and low asset's drift rates μ^B, μ^P. On one hand, the total return of base load

assets $\bar{R}^B - M^B = 0.05$ is constant and higher than the total mean return of peak load assets $\bar{R}^P - M^P = 0.04$. On the other side the base load asset is volatile $\sigma^B = 0.2$ while the peak load asset has volatile retuns $\sigma_R = 0.05$. Gamma is chosen such that the utility has a significant curvature in the region of interest $(a^B, a^P, b) \in [0, 2] \times [-1, 1]$.

8.3.2 CFB algorithm and approximation parameters

Our controlled forward-backward (CFB) algorithm from Part III provides us the ability (in terms of performance) to study this convex problem numerically. All following results are generate with this method. The CFB algorithm is developed in Part III and all details about the implementation are given in Part IV of this thesis.

Let h denote our approximation parameter. For our calculations, we use the time step size $\Delta t = h$. All other grid step sizes A^B, A^P, B, R and the number of simulations of the Brownian increments $\Delta W^B, \Delta W^R$ are derived from h. In particular, they are optimized according to equation (13.12) in Section 13.2 in Part III, in order to provide linear convergence of our CFB algorithm when $h \to 0$. In the following, we use the superscript h to indicate a numeric approximation χ^h of the opitmal solution χ^*.

8.4 Numerical results

Figure 34 shows the value function V^h and the optimal consumption χ^h at time $t = 0$ when starting with a bank account $B_0 = b = 0.0$ and a (peak load) return rate $R_0 = c = 0.0$. The axes are the initial amount of base load assets $A_0^B = a^B$ and peak load assets $A_0^P = a^P$. We observe that the value function $V_0^h(a^B, a^P, 0, 0)$ is increasing and concave w.r.t. a^B and a^P. The optimal $\chi_0^h(a^B, a^P, 0, 0)$ consumption is increasing and linear w.r.t. a^B and a^P. This is consistent with the one asset case.

Figure 35 shows the optimal allocation rates $\alpha^{B,h}$ and $\alpha^{P,h}$ at time $t = 0$ when starting with a bank account $B_0 b = 0.0$ and a (peak load) return rate $R_0 = c = 0.0$. The axes are the initial amount of base load assets $A_0^B = a^B$ and peak load assets $A_0^P = a^P$. On one hand, we observe that the optimal allocation rate $\alpha_0^{B,h}(a^B, a^P, 0, 0)$ depends (linearly) on a^B but it is independent of a^P. On the other hand, the optimal allocation rate $\alpha_0^{P,h}(a^B, a^P, 0, 0)$ depends (linearly) on a^P and also on a^B.

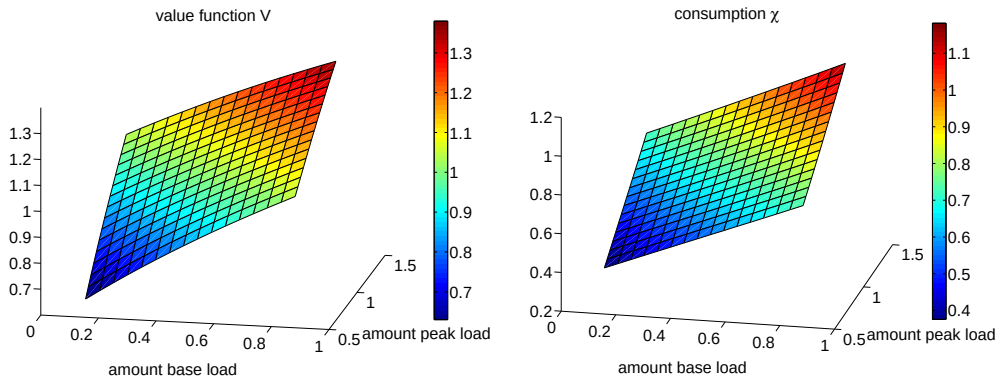

Figure 34: Value function $V_0^h(a^B, a^P, 0, 0)$ and optimal consumption $\chi_0^h(a^B, a^P, 0, 0)$ at time $t = 0$, for $h = 0.1$.

Extrapolating the allocation rates, the optimal portfolio (A_0^{B*}, A_0^{P*}) is approximately at $(1.6, 4.5)$. So the optimal portfolio holds three times more peak load assets A^{P*} than base load assets A^{B*} at time $t = 0$. However, the allocation towards the optimal amount of peak load A^{P*} is much slower than the allocation towards the optimal amount of base load A^{B*}, given the current (peak load) return rate $R_0 = 0$.

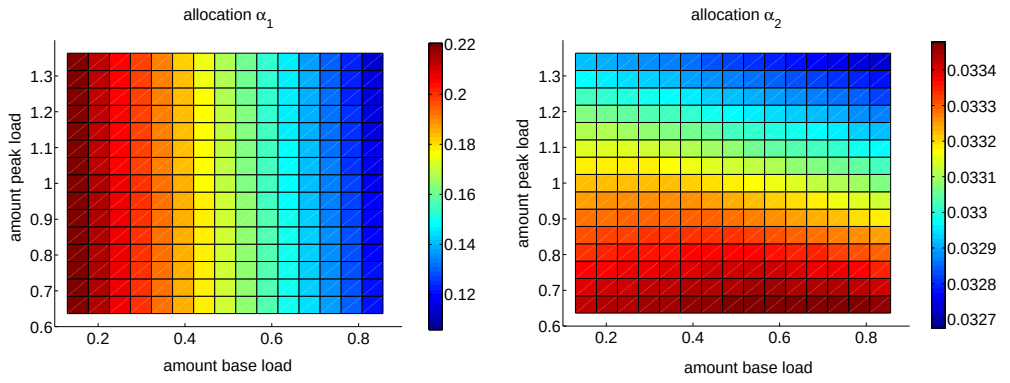

Figure 35: Optimal allocation rates $\alpha_0^{B,h}(a^B, a^P, 0, 0)$ and $\alpha_0^{P,h}(a^B, a^P, 0, 0)$ at time $t = 0$, for $h = 0.1$.

Figure 36 shows the optimal allocation rates $\alpha^{B,h}$ at time $t = 0$ when starting with an amount of peak load assets $A_0^P = a^P = 1.0$ and a bank account $B_0 b = 0.0$. The axes are the initial amount of base load assets $A_0^B = a^B$ and

the (peak load) return rate $R_0 = c$. We observe that the optimal allocation $\alpha_0^{B,h}(a^B, a^P, 0, 0)$ decrease linearly for increasing a^B but it is independent of the (peak load) return rate c. This is consistent with the one asset case.

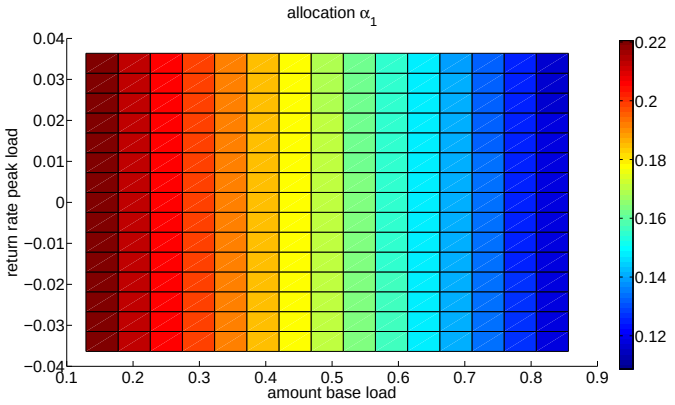

Figure 36: Optimal allocation rates $\alpha_t^{B,h}(a^B, 1, 0, c)$ and $\alpha_t^{p,h}(a^B, 1, 0, c)$ at time $t = 0$, for $h = 0.1$.

Figure 37 shows the optimal allocation rates $\alpha^{P,h}$ at time $t = 0$ when starting with an amount of base load assets $A_0^B = a^B = 1.0$ and a bank account $B_0 b = 0.0$. The axes are the initial amount of peak load assets $A_0^P = a^B$ and its return rate $R_0 = c$. We observe that the optimal allocation $\alpha_0^{B,h}(a^B, a^P, 0, 0)$ depends (linearly) on the return rate c but the dependency of the initial amount of assets a^P is comparatively low. This is different to the one asset case.

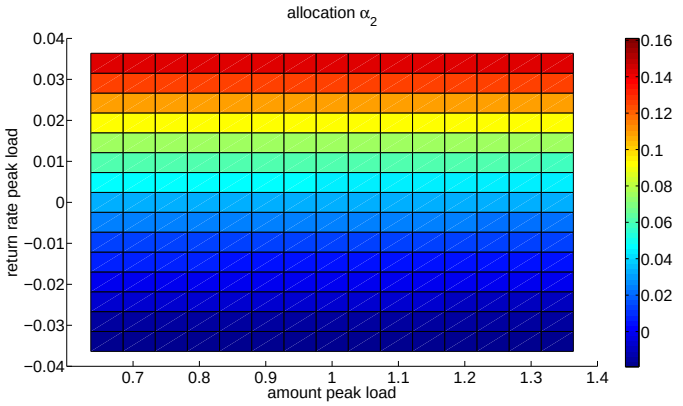

Figure 37: Optimal allocation rates $\alpha_t^{b,h}(1, a^P, 0, c)$ and $\alpha_t^{p,h}(1, a^P, 0, c)$ at time $t = 0$, for $h = 0.1$.

In particular, the allocation rate α_t^P - the speed of portfolio changes towards the optimal portfolio A_t^{P*} - highly depends on the (optimized) stochastic returns rates R_t and less on the current amount of assets A_t^P which are non-stochastic. Further investigation is needed in order to understand the observed effects and the dynamics of the problem. For now, we leave this to future research.

8.5 Numerical complexity of the two asset case

The presented two asset case, Problem 4 (8.2) - (8.6), is a stochastic optimal control problem in four state dimensions and three control dimensions. The numerical treatment of this multi-dimensional problem is a complex task. Common methods utilize the Hamilton-Jacobi-Bellman (HJB) equation (a second order non-linear parabolic PDE) or use a stochastic method to solve the related dynamic programming (DP) equation. The crux of the problem is the intrinsic optimization. For every step towards an optimal control in the control space, the HJB equation (or the DP equation) must be solved. The optimization multiplies the numerical complexity and amplify the curse of dimensionality, especially when the control spaced is also multi-dimensional.

In the next Part III of this thesis, we present a fairly new stochastic approach that we used to obtain the above results. The approach is based on the Stochastic Maximum Principle (SMP) and considers the problem as a forward-backward SDE coupled through an optimality condition. Therefore we call it a controlled forward-backward (CFB) algorithm. Besides introducing the theoretical background of the SMP and developing the CFB algorithm, we show in Section 15 that the CFB algorithm is less exposed the multiplication of complexity caused by the optimization. We also perform test calculations to compare this effect between the CFB and a DP algorithm in Section (14).

Part III
A controlled forward-backward algorithm for stochastic optimal control problems

9 Overview

9.1 Motivation

Our Problem 2 in Part II Section 5 of this thesis (reduced continuous-time model for optimal portfolio allocation of commodity related assets) is a stochastic optimal control problem. The two characteristic properties of the problem are:

- the stochastic state variable X_t is controlled, i.e., the SDE that describes the state dynamics dX_t depends on the control π_t,
- the gain function $f(t, x)$ is concave w.r.t. the state variable x.

These types of problems typically appear in mathematical finance and economics but also in other areas. A standard example is portfolio theory where the investor's allocation affects the stochastic evolution of his wealth. The concavity is provided whenever the investor is risk-averse, i.e., his utility function is concave.

The numerical treatment of such stochastic control problems is little explored. Standard methods are widely explored for the case of non-controlled SDEs but they become numerically expensive in the case of controlled SDEs, even in low dimensions. The sticking point is that the stepwise approach to an optimal control point $\pi_t^k \to \pi_t^*, k = 1, 2, ...$, in the control space is influencing the stochastic state variable $X_t(\pi_t^k)$. Therefore, the state simulations for $X_t(\pi_t^k)$ must be adjusted or repeated after each optimization step k until a fixed point is found. This effect is also called feedback control.

9.2 Current state of the art

Currently, the Dynamic Programming Principle (DPP) is the most common method to solve stochastic optimal control problems.

9.2.1 Deterministic approach

In continuous time, the DPP theory leads to the well-known Hamilton-Jacobi-Bellman (HJB) equation, a second order parabolic, partial differential equation (PDE) which describes the dynamics of the value function $v(t, x)$. In order to solve the PDE, various numerical schemes use a finite approximation of the value function and its derivatives (e.g. finite differences) and plug these approximations into the HJB equation. Applied to our problem, the crux of the matter is the optimization operator $\max_{\pi_t}(\cdot)$ in the HJB equation which makes it highly non-linear.

9.2.2 Stochastic approach

Other schemes take advantage of the DDP's discrete version, the Dynamic Programming (DP) equation. Theses schemes are based on an iteration backwards in time in order to calculate the value function. At each time step, one simulates the future state variable X_{t+1} (e.g. using Monte Carlo methods) and then exploits the knowledge of the value function $v(t+1, X_{t+1})$ on the ahead time point for all outcomes X_{t+1}. Applied to our problem where the state variable is controlled, the crux of the matter is the above mentioned feedback control $X_{t+1}(\pi_t)$.

Note that the HJB-based schemes are deterministic approaches because the considered PDEs are hiding the stochastic nature of the problem. In particular, the approach does not include the calculation of any expectation value. In contrast, the DP-based schemes are stochastic approaches, and they do include the calculation of expectation values. A comprehensive reference to numerical methods for stochastic control problems and Dynamic Programming is [Kushner, Dupuis 2001].

9.3 A new stochastic approach

In this part of the thesis, we propose a general stochastic algorithm that is based on the Stochastic Maximum Principle (SMP). The algorithm solves stochastic optimal control problems of the above type efficiently by utilizing the adjoint variables and exploiting the concavity of the problem.

It is well known that the SMP theory leads to a system of forward-backward stochastic differential equations (FBSDE) coupled through an external optimality condition. This was first studied by [Peng 1990]. The coupling arises

through the dependency of the state process X_t on the control π.

Recently, only a few authors proposed numerical schemes for pure FBSDE systems. By pure we mean that the FBSDE system does not include an optimality condition. They basically studied two types of approaches: the deterministic four-step scheme in [Yong, Zhou 1999] and the stochastic forward-backward algorithm from [Delarue, Menozzi 2005]. The latter is originally designed to solve quasi-linear PDEs.

Our attempt is to **synthesize the SMP together with a numerical scheme for FBSDE systems** in order to obtain an implementable stochastic algorithm for stochastic optimal control problems. In particular, we extend [Delarue, Menozzi 2005]'s forward-backward stochastic algorithm by a Newton-Raphson method in order to handle the coupled optimality condition.

As far as we are aware, this is the first stochastic approach for the solution of stochastic optimal control problems through the solution of the corresponding extended FBSDE system. Therefore, we call the scheme **controlled forward-backward (CFB) algorithm**.

9.4 Outline

First, we define the general problem statement for the special class of stochastic optimal control problems in Section 10. Second, we briefly review the DPP and the SMP in Section 11. In particular, we show the connection of both principles by deducing the FBSDE representation directly from the HJB equation. We also state a verification theorem for the SMP's sufficient conditions. In Section 12, we discretize the time-continuous FBSDE which arises in the SMP. Then, we develop our numerical scheme - the CFB algorithm - using a Newton-Raphson method for the optimization. Afterwards, we comment on convergence and stability issues in Section 13.

In Section 14, we show the structural advantage of our CFB algorithm for the specific problem class by comparing the complexity (computational costs) of our CFB scheme and a plain DP scheme. Finally, we present a performance study for the valuation of Swing options in order to support the theoretical advantages of our CFB scheme in Section 15. Alongside we show that our CFB algorithm is still considerable, but less efficient, when the state variable is non-controlled.

9.5 Notation

In this Part III of the thesis we use a different notation than in Part I and Part II in order to formulate the CFB algorithm in a general way. The reason is that our CFB algorithm is not limited to a specific problem. Instead, it can be easily applied to other examples of the same problem class. Therefore, we consider a general continuous-time stochastic optimal control problem where the state variable X is a controlled process of Markovian type, and the objective function J depends on the state X and on the control π.

10 Problem statement

Throughout the following, let $(\Omega, \mathcal{F}, P)$ denote a probability space, endowed with a d-dimensional Brownian motion $\{W_t,\, t \geq 0\}$, whose natural filtration is denoted by $\mathbb{F} = \{F_t,\, t \geq 0\}$, see Appendix, Definition A.8 and A.9. Now, let us consider the following general problem.

10.1 General problem and basic conditions

Problem 5 (General problem)
The dynamics of a controlled Markov[37] *process X_t, which represents the state of our system, are given by:*

$$dX_t = b(t, X_t, \pi_t)dt + \sigma(t, X_t, \pi_t)dW_t, \quad X_0 = x \in \mathbb{R}^n, \tag{10.1}$$

where $b : [0,T] \times \mathbb{R}^n \times A \to \mathbb{R}^n$ and $\sigma : [0,T] \times \mathbb{R}^n \times A \to \mathbb{R}^{n\times d}$ are real-valued functions and $A \subset \mathbb{R}^r$ is a convex and compact set. The goal is to maximize a given objective function J, which is defined on a finite time interval $[0,T]$, over the set of all admissible[38] *control policies:*

$$\mathbb{A}_{t,T} := \left\{ \{\pi_s \in \mathbb{R}^r,\, t \leq s \leq T\} \mid \pi_s \in A,\ \pi \text{ is } \mathbb{F}\text{-adapted} \right\}. \tag{10.2}$$

The objective function is defined for all $(t,x) \in [0,T] \times \mathbb{R}^n$ and $\pi \in \mathbb{A}_{t,T}$ as:

$$J(t,x,\pi) := E_t^{\pi}\left[\int_t^T f(s, X_s, \pi_s)ds + g(X_T) \middle| X_t = x \right], \tag{10.3}$$

where $f : [0,T] \times \mathbb{R}^n \times A \to \mathbb{R}$ and $g : \mathbb{R}^n \to \mathbb{R}$ are real-valued functions.

Here, the operator $E_t^{\pi}[\cdot]$ denotes the conditional expectation w.r.t. F_t, while using the fixed control policy π. Note that we assume the control-value set $A \subset \mathbb{R}^r$ to be convex and compact. This is an important property which we utilize in further theory.

If one exists, we define an optimal control policy by:

$$\pi^*(t,x) := \arg\max_{\pi \in \mathbb{A}_{t,T}} J(t,x,\pi), \quad \forall (t,x) \in [0,T] \times \mathbb{R}^n, \tag{10.4}$$

and the value function by:

$$\begin{aligned} v(t,x) &:= \sup_{\pi \in \mathbb{A}_{t,T}} J(t,x,\pi), && \forall\, (t,x) \in [0,T] \times \mathbb{R}^n, \\ v(T,x) &:= g(x), && \forall\, x \in \mathbb{R}^n. \end{aligned} \tag{10.5}$$

[37] See Appendix, Definition A.10
[38] See Appendix, Definition A.7

In summary, the goal is to maximize (10.3) over (10.2) subject to (10.1).

Problem 5 is a general setup for a stochastic optimal control problem. In particular, it does not state any assumption on the involved coefficient b, σ, f and g. Therefore, we can neither guarantee a solution of the SDE (10.1) nor the existence of the objective function (10.3). In order to do so, we introduce the following basic conditions.

Conditions 10.1 (Basic conditions)

1. *Let the maps $b : [0,T] \times \mathbb{R}^n \times A \to \mathbb{R}^n$ and $\sigma : [0,T] \times \mathbb{R}^n \times A \to \mathbb{R}^{n\times d}$ be measurable in $t \in [0,T]$, uniformly Lipschitz continuous w.r.t. x, continuous w.r.t. π, and bounded in $L^2([0,T],\mathbb{R})$. So there exists a fixed constant $L > 0$, such that:*

$$\begin{aligned} &|b(t,x,\pi_t) - b(t,\hat{x},\pi_t)| + |\sigma(t,x,\pi_t) - \sigma(t,\hat{x},\pi_t)| \le L|x-\hat{x}|, \\ &\qquad \forall t \in [0,T],\ \forall\, x,\hat{x} \in \mathbb{R}^n,\ \forall \pi \in \mathbb{A}_{0,T}, \\ &|b(\cdot,0,\pi_{(\cdot)})| + |\sigma(\cdot,0,\pi_{(\cdot)})| \in L^2([0,T],\mathbb{R}) \quad \forall \pi \in \mathbb{A}_{0,T}. \end{aligned} \tag{10.6}$$

2. *The maps $f : [0,T] \times \mathbb{R}^n \times A \to \mathbb{R}$ and $g : \mathbb{R}^n \to \mathbb{R}$ are measurable in $t \in [0,T]$, continuous w.r.t. π and satisfy a quadratic growth condition w.r.t. x, i.e., there exists a fixed constant $C \ge 0$, such that:*

$$|g(x)| + |f(t,x,\pi)| \le C(1+|x|^2) \quad \forall (t,x,\pi) \in [0,T] \times \mathbb{R}^n \times A. \tag{10.7}$$

Here, the notation $|\cdot|$ denotes the operator norm. From standard theory of stochastic differential equations, the SDE (10.1) has a unique $\mathbb{F}$-adapted solution X_t with continuous sample path and bounded moments. In consequence, the quadratic growth Condition 2 for the gain functions f and g ensures the finiteness of the objective function J in (10.3).

Theorem 10.2 (Existence of a unique solution of the SDE)
Let Condition 1 of the Basic conditions 10.1 hold. Then, for all $\pi \in \mathbb{A}_{0,T}$, there exists a unique $\mathbb{F}$-adapted continuous solution X_t of the controlled forward SDE (10.1):

$$\begin{aligned} X_0 &= x \quad a.s., \\ X_t &= X_0 + \int_0^t b(s,X_s,\pi_s)ds + \int_0^t \sigma(s,X_s,\pi_s)dW_s, \quad \forall t \in [0,T],\ a.s., \end{aligned} \tag{10.8}$$

with:

$$\int_0^t |b(s, X_s, \pi_s)|^2 + |\sigma(s, X_s, \pi_s)|^2 \, ds < \infty, \quad a.s. \tag{10.9}$$

The solution has bounded moments:

$$E[\sup_{t\in[0,T]} |X_t|^m] < \infty, \quad \forall m \geq 1. \tag{10.10}$$

A proof can be found in [Yong, Zhou 1999] chapter 1, theorem 6.3. and corollary 6.4. The key of the proof is a fix-point theorem for the contraction mapping $\xi \to X_t = x + \int_0^t b(s, \xi, \pi_s)ds + \int_0^t \sigma(s, \xi, \pi_s)dW_s$ on a Banach space.[39] The bounded moments are obtained through the Burkholder-Davis-Gundy inequality, [Burkholder, Davis, Gundy 1972].

Theorem 10.3 (Boundedness of the objective function)
Let the Basic conditions 10.1 hold. Then, for all $(t, x) \in [0, T] \times \mathbb{R}^n$ *and* $\pi \in \mathbb{A}_{t,T}$*:*

$$E_t^\pi \left[\int_t^T |f(s, X_s, \pi_s)| \, ds + |g(X_T)| \,\middle|\, X_t = x \right] < \infty.$$

A proof can be found in [Pham 2009] chapter 3.2, remark 3.2.1. The key of the proof is the boundedness of $E[\sup_t |X_t|^2]$ together with the quadratic growth condition on f.

In certain cases, the existence of an optimal solution pair $(\hat{X}, \pi^*)$ of our stochastic optimal control problem can be guaranteed when the Basic conditions 10.1 hold. For example, the following theorem provides existence of solutions for a special linear-concave problem case.

Theorem 10.4 (Existence of an optimal control for a special case)
Assume that the Basic conditions 10.1 hold. Additionally, assume that b, σ, f *are autonomous, i.e. they do not depend on time, that* f, g *are concave w.r.t.* (x, π) *and that* b, σ *are linear w.r.t.* (x, π)*, i.e.,* $\varphi = b, \sigma$ *are of the form:*

$$\varphi(x_t, \pi_t) = Ax_t + B\pi_t,$$

where A *and* B *are matrices of suitable size. Then the General problem 5 admits an optimal control* $\pi^* \in \mathbb{A}_T$ *and an according solution* $\hat{X}_t$*.*

[39] In our case, we have to extend the proof by a time dependent constant K_T for the contraction mapping in [Yong, Zhou 1999] equation 6.12. Since our time horizon T is finite, we can use their proof on a one-to-one basis replacing K with K_T. We even could extent the existence and uniqueness of a solution to the limit $T \to \infty$.

Proof 10.5
The proof for this linear-concave[40] *case can be found in [Yong, Zhou 1999] chapter 2, theorem 5.2. It utilizes the fact that the control domain A is convex and compact. Then, using Mazur's theorem, a minimizing sequence $\pi^j \in \mathbb{A}_{0,T}$ has a subsequence that converges strongly in $L^2_{\mathbb{F}}([0,T],\mathbb{R}^d)$. Due to the linearity of b and σ, the uniform convergence of the according X^j can be shown via the convergence of the subsequence π^j. The optimality is given due to the concavity of f and g.*

Comments

- In Condition 1, the Lipschitz continuity together with the boundedness of b, σ imply a linear growth condition w.r.t x.
- In Condition 2, the first integral in (10.8) is a usual Lebesgue integral and the second is an Itô integral. (10.9) ensures that both integrals are well-defined.
- The fact that the set of control values A is compact and b, σ, f are continuous w.r.t. π implies that b, σ, f attain their maximum w.r.t. π according to the extreme value theorem. Therefore, no π-terms are needed on the right hand side of (10.6) and (10.7).
- Since A is compact, all $\mathbb{F}$-adapted control policies $\pi \in \mathbb{A}_{0,T}$ are admissible, see Appendix, Definition A.7.

10.2 HJB conditions and existence of solutions

Let us consider a set of conditions in order to guaranty the existence of solutions for the general stochastic optimal control Problem 5.

In the next Section 11.1, we will state a necessary and a sufficient condition of optimality from the standard HJB theory. Roughly speaking, the necessary condition states that if the value function v in (10.5) is sufficiently smooth, $v \in C^{1,2}([0,T]\times\mathbb{R}^n)$, then it satisfies the well-known HJB equation, a second-order PDE:

$$\begin{aligned} &\partial_t v(t,x) + \sup_{\pi\in A} G\left(t,x,\pi,\nabla_x v(t,x),\nabla^2_{xx} v(t,x)\right) = 0, && \forall\,(t,x)\in[0,T]\times\mathbb{R}^n,\\ &v(T,x) = g(x), && \forall\, x\in\mathbb{R}^n. \end{aligned} \tag{10.11}$$

[40] Linear-convex for minimization problems.

where the generator of the diffusion is defined by:

$$G(t,x,\pi,p,M) := b(x,\pi)p + \frac{1}{2}\mathrm{tr}\left[\sigma\sigma'(x,\pi)M\right] + f(t,x,\pi). \tag{10.12}$$

The sufficient condition states that $\pi^* \in \mathbb{A}_{t,T}$ is an optimal control if:

$$\pi^*(t,x) = \arg\max_{\pi\in A} G\left(t,x,\pi,\nabla_x v(t,x),\nabla^2_{xx} v(t,x)\right). \tag{10.13}$$

In order to guaranty the existence of a solution of Problem 5, we state conditions under which the HJB equation (10.11) attains a solution $v \in C^{1,2}([0,T]\times\mathbb{R}^n)$ and that satisfies (10.13).

Conditions 10.6 (HJB conditions)

3. *Let b,σ,f be continuously differentiable w.r.t. t, twice continuously differentiable w.r.t. x, continuous w.r.t. π, and let all partial derivatives be bounded:*
$$b,\sigma,f \in C_b^{1,2,0}([0,T]\times\mathbb{R}^n\times A). \tag{10.14}$$

4. *Let g be three-times continuously differentiable w.r.t. x and let all partial derivatives be bounded:*
$$g \in C_b^3(\mathbb{R}^n). \tag{10.15}$$

5. *Let $\sigma(t,x,\pi)$ be an $(n\times n)$-matrix such that for all $(t,x,\pi)\in[0,T]\times\mathbb{R}^n\times A$ and $c>0$:*
$$\sum_{i,j=1}^{n}\sigma_{i,j}(t,x,\pi)\xi_i\xi_j \;\geq\; c|\,\xi|^2, \quad \forall\xi\in\mathbb{R}^n, \tag{10.16}$$
or,

   let $\sigma(t,x)$ be a nonsingular $(n\times n)$-matrix such that for all $(t,x)\in[0,T]\times\mathbb{R}^n$ and $c>0$:
$$|\,\sigma(t,x)^{-1}| \;\leq\; c. \tag{10.17}$$

Theorem 10.7 (Existence of solutions)
Let the HJB conditions 10.6 hold. Then, the HJB equation (10.11) admits a unique solution $v\in C_b^{1,2}([0,T]\times\mathbb{R}^n)$.

The proof can be found in [Fleming, Soner 2005] in chapter IV, theorem 4.2. It is based on standard theory for second order nonlinear parabolic PDEs as well as on results of [Krylov 1980] about the existence of solutions to SDEs with measurable coefficients.

Theorem 10.8 (Value function and optimal control)
Let the HJB conditions 10.6 hold. Then the unique classical solution $v(t,x) \in C_b^{1,2}([0,T] \times \mathbb{R}^n)$ *of the HJB equation (10.11) is the value function (10.5) of Problem 5 for all* $(t,x) \in [0,T] \times \mathbb{R}^n$.

Moreover, the control $\pi^*(t,x) \in \mathbb{A}_{0,T}$ *defined by (10.13) exists and is of Markovian*[41] *type. Therefore,* $\pi^*(t,x)$ *is an optimal control (10.4) of Problem 5 for all* $(t,x) \in [0,T] \times \mathbb{R}^n$.

The proof can be found in [Fleming, Soner 2005] in chapter IV, theorem 4.4.

Comments

- Condition 5 implies that $a := [\sigma\sigma^\top]$ is uniformly parabolic for all $(t,x,\pi) \in [0,T] \times \mathbb{R}^n \times A$:

$$\sum_{i,j=1}^{n} a_{i,j}(t,x,\pi)\xi_i\xi_j \;\geq\; c|\,\xi|^2, \quad \forall \xi \in \mathbb{R}^n. \tag{10.18}$$

  Therefore, the HJB equation becomes a second order nonlinear parabolic PDE. If the coefficients do not depend on time, the HJB equation is uniformly elliptic.

- If $\sigma(t,x)$ does not depend on the control, the HJB equation turns into a second order quasi-linear parabolic PDE.

- Note also that the HJB theory deals with a whole family of optimal control problems, one problem for every pair of initial values $(t,x) \in [0,T] \times \mathbb{R}^n$.

10.3 SMP conditions

Our CFB algorithm, which we introduce in section 12, is based on the Stochastic Maximum Principle (SMP), which we introduce in Section 11. However, the SMP framework utilizes the theory of stochastic calculus whereas the HJB framework basically utilizes PDE theory. Therefore, the SMP needs different conditions than the HJB framework. We call them SMP conditions.

[41] See Appendix, Definition A.11

Conditions 10.9 (SMP conditions)

6. *Let b, σ, f and g be continuously differentiable w.r.t. x and π:*

$$b, \sigma, f, g \in C^{0,1,1}([0,T] \times \mathbb{R}^n \times A). \tag{10.19}$$

7. *Let f and g be uniformly Lipschitz continuous w.r.t. x and bounded, i.e., there exists a fixed constant $L > 0$, such that:*

$$\begin{aligned} &|f(t,x,\pi) - f(t,\hat{x},\pi)| + |g(x) - g(\hat{x})| \;\le\; L|x - \hat{x}|, \\ &\qquad \forall t \in [0,T],\ \forall x, \hat{x} \in \mathbb{R}^n,\ \forall \pi \in A, \\ &|f(t,0,\pi)| + |g(0)| \;\le\; L \quad \forall t \in [0,T],\ \forall \pi \in A, \end{aligned} \tag{10.20}$$

and let also $\varphi = b_x, \sigma_x, f_x, g_x$ be uniformly Lipschitz continuous w.r.t. x, i.e., there exists a fixed constant $L > 0$, such that:

$$\begin{aligned} &|\varphi(t,x,\pi) - \varphi(t,\hat{x},\pi)| \;\le\; L|x - \hat{x}|, \\ &\forall t \in [0,T],\ \forall x, \hat{x} \in \mathbb{R}^n,\ \forall \pi \in A. \end{aligned} \tag{10.21}$$

8. *Let g be concave w.r.t x and for any fixed $y \in \mathbb{R}^n$ and $z \in \mathbb{R}^{n \times d}$ let the function:*

$$H(t,x,\pi) = b(t,x,\pi)y + tr[\sigma(t,x,\pi)^\top z] + f(t,x,\pi),$$

be uniformly concave w.r.t. x and π for all $t \in [0,T]$.

Comments

- Comparing the HJB conditions 3 and 4 with the SMP conditions 6 and 7, we observe that the SMP theory needs less smoothness and no boundedness of the coefficients. Instead, it requires global Lipschitz continuity of the coefficients and their derivatives.

- The HJB condition 5 states a parabolic PDE property in order to guarantee smooth solutions. Instead, the SMP imposes the concavity Condition 8 on the Hamilton function H and the terminal value function g in order to be sufficient for optimality, see Section 11.3.

10.4 A class of concave problems

The concavity Condition 8 characterizes the specific class of stochastic optimal control problems on which we focus in this thesis. We refer to it as the class of concave problems. Our CFB algorithm is designed especially for this problem class which we define as follows.

Problem 6 (Concave problem)
The goal is to maximize (10.3) over (10.2) subject to (10.1) while the Basic conditions 10.1 and the SMP conditions 10.9 are satisfied.

A rigorous proof of the existence of solutions for the general class of Concave problems 6 is similar to providing the existence of a solution for a general second order nonlinear PDE. This for sure goes far beyond the scope of this thesis. Nevertheless we state the following hypothesis and a promising outline of a hypothetical proof.

Hypothesis 10.10 (Existence of solutions for Concave problems)
The Concave problem 6 admits an optimal state-control pair:

$$(\hat{X}, \pi^*) \in (L^2_{\mathbb{F}}([0,T], \mathbb{R}^n), \mathbb{A}_{0,T}).$$

A proof could have the following outline. As in the proof of Theorem 10.4, we utilize the fact that the control domain A is convex and compact plus Mazur's theorem to provide a converging sequence $\pi^j \to \pi^* \in \mathbb{A}_T$. Moreover, for every $\pi^j \in \mathbb{A}_T$, we can provide an unique strong solution $X^j \in L^2_{\mathbb{F}}([0,T], \mathbb{R}^n)$ by Theorem 10.2. We denote the solution under π^* by $\hat{X}$. The optimality is given via the SMP due to the concavity of H and g, see Section 11.

Comments

- If we additionally assume that the HJB conditions 10.6 hold, in particular the parabolic property, then the existence of solutions is given due to Theorem 10.8 and the connection between the HJB theory and SMP theory, which we present in the following Section 11.

11 The Stochastic Maximum Principle

In this section, we review the Stochastic Maximum Principle (SMP) in order to find solutions to stochastic optimal control problems. The SMP formulates a necessary condition for optimality. Together with a concavity[42] assumption, the condition is sufficient and lays the foundation of our CFB algorithm in Section 12.

We also compare the SMP with the Dynamic Programming Principle (DPP). The DPP is most common to numerically solve stochastic optimal control problems and serves as a landmark for us to measure the efficiency of our CFB algorithm. The DPP also states necessary and sufficient conditions of optimality. We put an emphasis on the comparison of the SMP and the DPP in order to clarify the connections between both and to show the transferability of theoretical results. One of our main contributions in this thesis is the conceivable overview of SMP and DDP theory in Table 9 and their connection through Theorem 11.14.

Before we start, we briefly review the Dynamic Programming Principle (DPP) and Pontryagin's maximum principle (PMP), which is the deterministic foundation of the SMP, in Section 11.1 and 11.2, respectively. In both reviews, we basically follow the book of [Yong, Zhou 1999].[43]

11.1 Brief review of dynamic programming

The well known Bellman principle of optimality [Bellman 1957] states that the value function $v(t,x)$ in (10.5) satisfies the following dynamic programming (DP) equation for any $(t,x) \in [0,T] \times \mathbb{R}^n$:

$$v(t,x) = \sup_{\pi \in \mathbb{A}_{t,\tau}} E_t\left[\int_t^\tau f(s, X_s^\pi, \pi_s) ds + v(\tau, X_\tau^\pi) \middle| X_t = x \right], \tag{11.1}$$

where $\mathbb{A}_{t,\tau}$ is the set of all admissible control policies over the time interval $[t,\tau]$ and X_t^π is the solution of the SDE (10.1) while using the control π. Using Ito's formula [Ito 1951] and the Feynman-Kac formula, one can derive the well-known HJB equation out of the DP equation and state the following necessary and sufficient conditions.

[42] Convexity when minimizing the objective function.

[43] This means that informal statements without proofs are cited from this book or are otherwise referenced appropriately.

Theorem 11.1 (Necessary condition)
Let the Basic Conditions 10.1 hold. Also, let f and g be uniformly Lipschitz continuous w.r.t. x and bounded, i.e., there exists a fixed constant $L > 0$, such that:

$$\begin{aligned} |f(t,x,\pi) - f(t,\hat{x},\pi)| + |g(x) - g(\hat{x})| \;&\leq\; L|x-\hat{x}|, \\ \forall t \in [0,T],\ \forall x,\hat{x} \in \mathbb{R}^n,\ \forall \pi \in A, \\ |f(t,0,\pi)| + |f(t,0,\pi)| \;&\leq\; L \quad \forall t \in [0,T],\ \forall \pi \in A, \end{aligned} \tag{11.2}$$

If the value function $v(t,x)$ of Problem 5 is in $C^{1,2}([0,T]\times\mathbb{R}^n)$, then it satisfies the HJB equation:

$$\begin{aligned} &\partial_t v(t,x) + \sup_{\pi\in A} G\left(t,x,\pi,\nabla_x v(t,x),\nabla^2_{xx} v(t,x)\right) = 0, && \forall\,(t,x) \in [0,T]\times\mathbb{R}^n, \\ &v(T,x) = g(x), && \forall\, x \in \mathbb{R}^n. \end{aligned} \tag{11.3}$$

where the generator of the diffusion is defined by:

$$G(t,x,\pi,p,M) := b(x,\pi)p + \frac{1}{2} tr[\sigma\sigma'(x,\pi)M] + f(t,x,\pi). \tag{11.4}$$

Theorem 11.2 (Sufficient condition)
Let the Basic Conditions 10.1 hold and let f and g be uniformly Lipschitz continuous w.r.t. x and bounded. Assume that the value function $v(t,x) \in C^{1,2}([0,T]\times\mathbb{R}^n)$. An admissible control policy $\pi^* \in \mathbb{A}_T$ is an optimal control of Problem 5, if it satisfies the following optimality condition for all $t \in [0,T]$:

$$G(t,\hat{X}_t,\pi^*_t,v_x(t,\hat{X}_t),v_{xx}(t,\hat{X}_t)) = \max_{\pi\in A} G(t,\hat{X}_t,\pi_t,v_x(t,\hat{X}_t),v_{xx}(t,\hat{X}_t)), \tag{11.5}$$

where $\hat{X}_t$ is the solution of the SDE (10.1) while using the optimal control π^* and starting at $\hat{X}_0 = x$.

The proofs for the above conditions to be necessary and sufficient for optimality can be found for example in [Yong, Zhou 1999] chapter 4, proposition 3.5. and chapter 5, theorem 5.1, respectively. The latter optimality condition is often called verification. It is a sufficient condition such that a solution $v(t,x)$ of the HJB equation (11.3) provides the optimal controls through (11.5). Remember, in Theorem 10.7 we showed that the HJB equation has a unique solution $v \in C^{1,2}([0,T]\times\mathbb{R}^n)$ whenever the HJB conditions 10.6 are satisfied.

11.1.1 DPP based schemes to find solutions

Now we take a look on how to find (numerical) solutions of our stochastic optimal control Problem 5 by using the DPP (HJB equation).

The three-step scheme

In order to solve a stochastic optimal control problem via the HJB equation, one needs to perform the following three steps.

1. Step: Find the solution $v(t,x) \in C^{1,2}$ of the HJB equation (11.3) for all $(t,x) \in [0,T] \times \mathbb{R}^n$.
2. Step: Find the optimal control $\pi^*(t,x)$ via the optimality condition (11.5) for all $(t,x) \in [0,T] \times \mathbb{R}^n$.
3. Step: Solve the state equation (10.1) for X_t, using the optimal control $\pi^*(t,X_t)$.

Especially the first step in the tree-step scheme could be a complex task since the HJB equation is a non-linear second-order parabolic PDE. In the case that no analytical solution can be provided, numerical methods need to be considered.

Numerical schemes

As outlined in the introduction (Section 9) there are two basic approaches to solve a stochastic optimal control problem numerically via the DPP.

- **Deterministic approach**: By using finite approximations of the value function and its derivatives, e.g. finite differences or finite elements, one can calculate solutions $v(t_i,x_j)$ of the **HJB equation** (11.3) for a finite number of points $(t_i,x_j) \in [0,T] \times \mathbb{R}^n$, $i = 1,...,N$, $j = 1,...,M$. Once knowing the value function $v(t_i,x_j)$, step 2 and 3 of the tree-step scheme can be performed to obtain an approximative optimal control policy $\pi^*(t_i,x_j)$ and the according state process $\hat{X}_t(t_i,x_j)$ for any initial pair (t_i,x_j).
- **Stochastic approach**: Starting with $v(T,x) = g(x)$ at time T, one can calculate an approximative solution of the value function on a discrete time grid $t_i \in [0,T]$, $i = 1,...,N$, using the **DP equation** (11.1). This is often called a backward iteration. Common methods to calculate approximations of the expectation operator in the DP equation are Monte Carlo methods, Quantization methods and Markov-chain approximations. In this approach, the optimal control policy π^* and the state $\hat{X}_t$ are calculated on the way backwards in time.

Both approaches require an optimization over the control space A. There are two ways to solve the optimization.

- In the **direct approach** one discretizes the control space $A \subset \mathbb{R}^r$ into a finite number of controls $\pi^j \in A$, $j = 1, ..., K$, and compares the solutions for every point π^j. The crux of the matter is, that this approach leads to a discretization of $1 \times n \times r$ dimensions in total, which is numerically complex.

- On the other hand, an **indirect approach** discretizes only the state space only and uses a numerical optimizer to search the optimal control values π^* in A. In view of our Problem 5, this is also numerically complex because of the feedback control $X_t(\pi)$. For example, using the stochastic approach and a line-search algorithm, one has to calculate the expectation value $E^{\pi^k}[\cdot]$ in the DP equation (11.1) multiple times at every optimization step π^k in order to find the next appropriate search direction $k+1$ in the control space A. In the case of a feedback control, one has to simulate $X_t(\pi^k)$ repeatedly for each optimization step π^k.

In best practice, supplementary methods are used to reduce the computational costs. Common tools are global grid-refinement iterations (which work good in the deterministic case, but worse in the stochastic case), importance sampling and variance reduction (which may not be appropriate for feedback controls), predictor-corrector methods or least-square Monte Carlo simulations. We do not want to go into any further detail here. All methods are subject to plenty of research papers.

11.1.2 Summary

The intention of the above review of dynamic programming is to remember the reader that:

- the DPP/HJB theory states a necessary and a sufficient condition of optimality.

- one can derive deterministic and stochastic approaches to find numerical solutions through the HJB equation and the DP equation, respectively.

- the resulting numerical schemes are complex and expensive, especially in the case of feedback controls.

Before we consider an alternative approach that utilizes the Stochastic Maximum Principle (SMP), we review its deterministic version, the well-known Pontryagin's maximum principle (PMP).

11.2 Brief review of Pontryagin's maximum principle

Let us briefly consider the following deterministic optimal control problem.

Definition 11.3 (Deterministic problem)
Let the functions f, g and b are given by the General problem 5 (10.1) - (10.3). The goal is to maximize the objective function:

$$J(t,x,\pi) = \max_{\pi} \int_t^T f(s, X_s, \pi_s)ds + g(X_T), \tag{11.6}$$

over a given admissible control set $\mathbb{A}_{t,T}$ and under the condition:

$$\dot{X}_t = b(t, X_t, \pi_t), \quad X_0 = x. \tag{11.7}$$

For these type of problems, the well-known Pontyagin's maximum principle (PMP) gives a set of first-order necessary conditions on optimal solution pairs $(\hat{X}, \pi^*)$.

Theorem 11.4 (Necessary condition)
Let the Basic Conditions 10.1 and the SMP Conditions 10.9, 6 - 7 hold. Assuming that $(\hat{X}, \pi^*)$ is an optimal state-control pair of Problem 11.3, there exits a process $Y : [0,T] \to \mathbb{R}^n$ satisfying the following backward SDE:

$$\dot{Y}_t = -\nabla_x H(t, \hat{X}_t, \pi_t^*, Y_t), \quad Y_T = \nabla_x g(\hat{X}_T), \tag{11.8}$$

and the optimality condition:

$$H(t, \hat{X}_t, \pi_t^*, \hat{Y}_t) = \max_{\pi \in A} H(t, \hat{X}_t, \pi, \hat{Y}_t), \tag{11.9}$$

where the Hamilton H is defined as:

$$H(t,x,\pi,y) := b(t,x,\pi)y + f(t,x,\pi). \tag{11.10}$$

Y_t is called adjoint variable and the backward ODE (11.8) is called first-order adjoint equation. Equation (11.9) is an equivalent formulation of the variational inequality.

Theorem 11.5 (Sufficient condition)
Let the Basic Conditions 10.1 and the SMP Conditions 10.9, 6 - 8 hold (i.e., g and H are uniformly[44] concave w.r.t. x and π). Let $(\hat{X}, \pi^*)$ be an admissible state-control pair that satisfies equation (11.7) and (11.9) while Y_t is the corresponding adjoint variable satisfying (11.8). Then the state-control pair $(\hat{X}, \pi^*)$ is an optimal solution of Problem 11.3.

The proofs for the above conditions to be necessary and sufficient for optimality can be found in [Yong, Zhou 1999], chapter 3.2.

[44] That means concave for all $t \in [0,T]$.

Further remarks

- Pontyagin's maximum principle (PMP) states a necessary and a sufficient condition similar to the DPP.
- The sufficient condition of the PMP needs the concavity conditions for the terminal function g and for the Hamilton function H. In contrast, the sufficient condition of the DPP needs the assumption that $v(t,x) \in C^{1,2}$. This assumption can be guaranteed through Theorem 10.7 when $[\sigma\sigma']$ is uniformly parabolic.
- The main steps in the proof for the necessary and the sufficient conditions of the PMP are a first-order Taylor expansion and a duality relation, respectively.
- In economic theory, the adjoint variable Y_t corresponds to the so-called shadow price for the resource X_t.
- In Lagrange theory, the adjoint variable Y_t corresponds to the Lagrange multiplier, commonly denoted by λ.

11.2.1 A two-point boundary problem

By using the PMP, we can find solutions of Problem (11.6)-(11.7), Definition 11.3, through the solutions of the forward ODE (11.7) and the backward adjoint equation (11.8):

$$\begin{aligned} \dot{X}_t &= b(t, X_t, \pi_t^*), & X_0 &= x, \\ \dot{Y}_t &= -\nabla_x H(t, X_t, \pi_t^*, Y_t), & Y_T &= \nabla_x g(X_T), \end{aligned} \tag{11.11}$$

where the optimal control $\pi^* \in \mathbb{A}_{t,T}$ must satisfy the optimality condition:

$$H(t, \hat{X}_t, \pi_t^*, \hat{Y}_t) = \max_{\pi \in A} H(t, \hat{X}_t, \pi, \hat{Y}_t). \tag{11.12}$$

System (11.11) is a two-point boundary problem coupled through an optimality condition (11.12). Solving this two-point boundary problem, one can obtain an optimal state-control pair $(\hat{X}, \pi^*)$ without considering the value function $v(t,x)$. Moreover, knowing the optimal $(\hat{X}, \pi^*)$, one can easily calculate $v(t,x)$ via the simplified Dynamic Programming equation:

$$v(t,x) = \int_t^\tau f(s, X_s^{\pi^*}, \pi_s^*) ds + v(\tau, X_\tau^{\pi^*}), \quad X_t = x. \tag{11.13}$$

11.2.2 Summary

The intention of the above subsection is to remember the reader that:

- the PMP states a necessary and a sufficient condition of optimality,
- the PMP needs concavity conditions for g and H, while the DPP can be used when $[\sigma\sigma']$ is uniformly parabolic.
- the PMP leads to a two-point boundary problem.

After our review of the DPP and the PMP, we are now prepared to take a closer look at the Stochastic Maximum Principle (SMP).

11.3 The Stochastic Maximum Principle

Now, let us consider the stochastic Problem 5 (10.1) - (10.3) where the state variable is stochastic:

$$dX_t = b(t, X_t, \pi_t)dt + \sigma(t, X_t, \pi_t)dW_t, \quad X_0 = x \in \mathbb{R}^n. \tag{11.14}$$

Let S^n denote the set of all symmetric $(n \times n)$-matrices. We define the Hamilton function of the stochastic problem that will appear in the SMP.

Definition 11.6 (Hamiltonian)
The Hamiltonian $H : [0,T] \times \mathbb{R}^n \times A \times \mathbb{R}^n \times \mathbb{R}^{n \times d} \to \mathbb{R}$ of the stochastic optimal control Problem 5 is defined by:

$$H(t, x, \pi, y, z) := b(t, x, \pi)y + tr[\sigma(t, x, \pi)^\top z] + f(t, x, \pi). \tag{11.15}$$

Now we state the general Stochastic Maximum Principle (SMP). We will see later that a reduced version of the SMP is sufficient for many applications. The SMP states that a state-control pair $(\hat{X}, \pi^*)$ must satisfy the following necessary conditions to be optimal.

Theorem 11.7 (Necessary condition)
Let the Basic Conditions 10.1 and the SMP Conditions 10.9, 6 - 7 hold. Additionally, let b, σ, f and g be twice continuously differentiable w.r.t. x and let there exist a modulus of continuity $\bar{w} : [0, \infty) \to [0, \infty)$ such that for $\varphi = b, \sigma, f, g$, we have:

$$\begin{aligned} &|\varphi_{xx}(t, x, \pi) - \varphi_{xx}(t, \hat{x}, \hat{\pi})| \leq \bar{w}(|x - \hat{x}|), \\ &\forall t \in [0, T], \quad \forall\, x, \hat{x} \in \mathbb{R}^n, \quad \forall \pi, \hat{\pi} \in A. \end{aligned} \tag{11.16}$$

If $(\hat{X}, \pi^*)$ is an optimal state-control pair of Problem 5, then there exist pairs of $\mathbb{F}$-adapted processes[45]*:*

$$\begin{cases} (Y,Z) \in L^2_{\mathbb{F}}([0,T],\mathbb{R}^n) \times \left(L^2_{\mathbb{F}}([0,T],\mathbb{R}^n)\right)^d, \\ (P,Q) \in L^2_{\mathbb{F}}([0,T],S^n) \times \left(L^2_{\mathbb{F}}([0,T],S^n)\right)^d, \end{cases} \tag{11.17}$$

which satisfy the first-order adjoint equation:

$$\begin{aligned} dY_t = &\quad -\nabla_x H(t,\hat{X}_t,\pi_t^*,Y_t,Z_t)dt + Z_t dW_t, \\ Y_T = &\quad \nabla_x g(\hat{X}_T), \end{aligned} \tag{11.18}$$

and the second-order adjoint equation:

$$\begin{aligned} dP_t = &\quad -\Big[\nabla_x b(\hat{X}_t,\pi_t^*)^\top P_t + P_t \nabla_x b(\hat{X}_t,\pi_t^*) \\ &+ \textstyle\sum_{j=1}^d \nabla_x \sigma^j(\hat{X}_t,\pi_t^*)^\top P_t \nabla_x \sigma^j(\hat{X}_t,\pi_t^*) \\ &+ \textstyle\sum_{j=1}^d \nabla_x \sigma^j(\hat{X}_t,\pi_t^*)^\top Q_t^j + Q_t^j \nabla_x \sigma^j(\hat{X}_t,\pi_t^*) \\ &+ \nabla^2_{xx} H(t,\hat{X}_t,\pi_t^*,Y_t,Z_t)\Big]dt + \textstyle\sum_{j=1}^d Q_t^j dW_t^j, \\ P_T = &\quad \nabla^2_{xx} g(\hat{X}_T), \end{aligned} \tag{11.19}$$

for all $t \in [0,T]$, respectively, such that the following optimality condition hold:

$$\tilde{H}(t,\hat{X}_t,\pi_t^*,Y_t,Z_t,P_t) = \max_{\pi \in A} \tilde{H}(t,\hat{X}_t,\pi,Y_t,Z_t,P_t), \quad \forall t \in [0,T]. \tag{11.20}$$

Here, the function $\tilde{H}$ is defined as:

$$\begin{aligned} \tilde{H}(t,x,\pi,Y_t,Z_t,P_t) := H(t,x,\pi,Y_t,Z_t) - \tfrac{1}{2} tr\Big[\sigma(\hat{X}_t,\pi_t^*)\sigma(\hat{X}_t,\pi_t^*)^\top P_t\Big] \\ + \tfrac{1}{2} tr\Big[[\sigma(x,\pi) - \sigma(\hat{X}_t,\pi_t^*)][\sigma(x,\pi) - \sigma(\hat{X}_t,\pi_t^*)]^\top P_t\Big]. \end{aligned} \tag{11.21}$$

Note that the adjoint variables Y_t and P_t must be adapted to the filtration $\mathbb{F}$ through (11.17). This is an important condition in the above theorem. The condition 'determine' the processes Z_t and Q_t. Let us now state the sufficient conditions of optimality.

[45]The definition of $L^2_{\mathbb{F}}$ is given in Section 5.2.1.

Theorem 11.8 (Sufficient condition)
Let the Basic Conditions 10.1 and the SMP Conditions 10.9, 6 - 8 hold (i.e., g and H are uniformly[46] *concave w.r.t. x and π). Let* $(\hat{X}, \pi^*)$ *be an admissible state-control pair that satisfies (11.14) and (11.20) while* Y_t, Z_t, P_t, Q_t *are the corresponding adjoint variables satisfying (11.17), (11.18) and (11.19). Then the state-control pair* $(\hat{X}, \pi^*)$ *is an optimal solution of Problem 5.*

The proofs for the above conditions to be necessary and sufficient for optimality can be found in [Yong, Zhou 1999] chapter 3, section 4 and 5, respectively.

A stochastic backward SDE has a solution pair

Comparing the deterministic and the stochastic case, the reader may wonder why there are four adjoint variables in (11.17) but only two equations (11.18) and (11.19) for them. The reason is that Y_t and P_t must satisfy an additional condition, namely they must be adapted to the given filtration $\mathbb{F}$ according to (11.17).

For an ordinary differential equation, under the usual Lipschitz condition, both an initial value problem (11.7) and a terminal value problem (11.8) are well-posed. In particular, they become equal through a time reversion. This is fundamental different in the stochastic case (11.18) and (11.19) since the solutions Y_t and P_t must be adapted to the given filtration $\mathbb{F}$. So the time can not be easily reverted. Therefore, the variables Z_t and Q_t are needed to ensure that Y_t and P_t are $\mathbb{F}$-adapted, respectively.

[Yong, Zhou 1999] states that "it is the second component (Z_t and Q_t) that corrects the possible "non-adaptiveness" caused by the backward nature of the equations, including the given terminal value of the first component. The solution pair has an interesting interpretation in functional analysis and duality analysis. It is nothing but the (unique) *Riesz Representation* of a certain functional on the Hilbert space $L^2_{\mathbb{F}}([0,T],\mathbb{R}^n) \times L^2_{\mathbb{F}}([0,T],\mathbb{R}^{n\times d})$ defined via a forward SDE that is dual to the backward SDE."

In consequence, under the SMP Conditions 10.9, both backward SDEs (11.18) and (11.19) admit a unique adapted solution pair (Y_t, Z_t) and (P_t, Q_t), respectively. A proof can be found in [Yong, Zhou 1999] chapter 7, Theorem 3.2.

[46] That means concave for all $t \in [0,T]$.

Further remarks

- The proof of the SMP needs a second-order Taylor expansion since dW is only of the order $\sqrt{(\epsilon)}$ when using variations. The PMP needs only a first-order Taylor expansion.

- We can not reverse time in the stochastic case, as it would destroy the non-anticipativness of the solutions. Therefore, each adjoint equation consists of a tuple of adjoint variables.

- In economic theory, Y_t is the random shadow price process and Z_t is the instantaneous standard deviation of the depreciation rate of the random shadow price process.

- Nevertheless, the shadow price deviation pair (Y_t, Z_t) is not able to fully characterize the trade between the costs and control gain in an uncertain, non-concave environment. The optimality condition must be extended by a risk adjustment, and thus the second-order adjoint variable (P_t, Q_t) becomes necessary, see [Yong, Zhou 1999], page 116-118 for details on this.

11.3.1 Simplifications of the SMP

Lemma 11.9 (Simplification for non-controlled diffusions)
If the diffusion $\sigma(x, \pi) = \sigma(x)$ does not depend on the control variable, then the optimality conditions (11.20) over $\tilde{H}$ simplify to an optimality condition over H:

$$H(t, \hat{X}_t, \pi_t^*, Y_t, Z_t) = \max_{\pi \in A} H(t, \hat{X}_t, \pi, Y_t, Z_t), \quad \forall t \in [0, T]. \tag{11.22}$$

This follows from the definition of $\tilde{H}$ in (11.21). In consequence, the second-order adjoint variables (P_t, Q_t) are not needed when solving the stochastic control problem.

Lemma 11.10 (Simplification for concave problems)
Consider the Concave problem 6 of Section 10.4 (i.g. the Basic Conditions 10.1 and the SMP Conditions 10.9 hold). Then the optimality condition (11.20) in the necessary conditions simplifies to:

$$H(t, \hat{X}_t, \pi_t^*, Y_t, Z_t) = \max_{\pi \in A} H(t, \hat{X}_t, \pi, Y_t, Z_t), \quad \forall t \in [0, T]. \tag{11.23}$$

Proof 11.11
[Yong, Zhou 1999], page 120, show that, when the control domain A is convex and all coefficients b, σ, f and g are continuously differentiable w.r.t. π, then the optimality condition (variational inequality) (11.20) implies a local form:

$$\left\langle \partial_\pi H(t, \hat{X}_t, \pi_t^*, Y_t, Z_t), \pi - \pi_t^* \right\rangle \leq 0, \quad \forall t \in [0,T],\ \forall \pi \in A.$$

Moreover, when the Hamiltonian H is concave w.r.t. π, (11.11) is equal to:

$$H(t, \hat{X}_t, \pi_t^*, Y_t, Z_t) = \max_{\pi \in A} H(t, \hat{X}_t, \pi, Y_t, Z_t).$$

qed.

Note that in both cases, the simplified optimality conditions (11.22) and (11.23) do not involve the second-order adjoint processes P_t and Q_t. When the second order adjoint variables (P_t, Q_t) can be neglected, the proof of the SMP does not need the twice continuous differentiability of b, σ, f and g w.r.t. x in (11.16). Instead, the C^1-property is enough to apply the SMP.

[Yong, Zhou 1999] states that "when the Hamiltonian $H(t, x, \pi, y, z)$ is already concave in π, the second-order adjoint process P_t plays no role at all. This is because the concavity of H reflects the risk adjustment already." We summarize this result in the following section.

11.3.2 A stochastic two-point boundary problem

Using SMP's necessary and sufficient condition, the forward state SDE (10.1) together with the backward adjoint equations (11.18) and (11.19) build a system of forward-backward stochastic differential equations (FBSDE) coupled through the optimality condition (11.20). This is a stochastic version of the two-point boundary problem that we faced in Pontryagin's maximum principle in Section 11.2.1.

Theorem 11.12 (Coupled FBSDE system)
Our Concave Problem 6 (i.e. the Basic Conditions 10.1 and the SMP Conditions 10.9 hold) is equivalent to searching for the $\mathbb{F}$-adapted quadruple $(\hat{X}_t, \pi_t^*, Y_t, Z_t)$, which satisfies the following coupled FBSDE system:

$$\begin{aligned} dX_t &= b(t, X_t, \pi_t^*)dt + \sigma(t, X_t, \pi_t^*)dW_t, && X_0 = x \in \mathbb{R}^n, \\ dY_t &= -\nabla_x H(t, \hat{X}_t, \pi_t^*, Y_t, Z_t)dt + Z_t dW_t, && Y_T = \nabla_x g(X_T), \\ \pi_t^* &= \arg\max_{\pi_t \in A} H(t, X_t, \pi_t, Y_t, Z_t). \end{aligned} \tag{11.24}$$

This follows from the the Necessary Condition 11.7 and the Sufficient Condition 11.8 of the SMP together with Lemma 11.10.

By solving the stochastic two-point boundary problem, the optimal state-control pair $(\hat{X}, \pi^*)$ can be found without considering the value function $v(t,x)$. However, knowing the optimal solution-pair $(\hat{X}, \pi^*)$, the value function $v(t,x)$ can be calculated via the simplified Dynamic Programming equation:

$$\begin{aligned} v(t,x) &= E_t^{\pi^*}\left[\int_t^\tau f(s,\hat{X}_s,\pi_s)ds + v(\tau,\hat{X}_\tau)\Big| X_t = x\right], \\ v(T,x) &= g(x). \end{aligned} \tag{11.25}$$

11.3.3 Summary

The intention of the above subsection is to remember the reader that:

- the SMP states a necessary and a sufficient condition of optimality,
- the SMP needs concavity conditions for g and H, while the DPP can be used when $[\sigma\sigma']$ is uniformly parabolic.
- the SMP leads to a stochastic two-point boundary problem, namely a FBSDE system coupled through an optimality condition.

11.4 Solving a coupled FBSDE system

Now we take a closer look on how to solve a coupled FBSDE system. Yet, there does not exist a numerical scheme for stochastic optimal control problems that operates via the SMP. However, methods are available that find numerical solutions of FBSDE systems for different degrees of coupling. To apply these methods to our problem, let us assume for a moment that the optimal control process π_t^* can be expressed as a measurable function $\pi^* : [0,T] \times \mathbb{R}^n \times \mathbb{R}^n \times S^n \to A$, $\pi_t^* = \pi^*(t, \hat{X}_t, Y_t, Z_t)$. Then we search for an $\mathbb{F}$-adapted solution triple (X_t, Y_t, Z_t) for the following coupled FBSDE system for $t \in [0,T]$:

$$\begin{aligned} dX_t &= \tilde{b}(t,X_t,Y_t,Z_t)dt + \tilde{\sigma}(t,X_t,Y_t,Z_t)dW_t \qquad & X_0 &= x, \\ dY_t &= \tilde{f}(t,X_t,Y_t,Z_t)dt + Z_t dW_t, & Y_T &= \tilde{g}(X_T), \end{aligned} \tag{11.26}$$

where:

$$\begin{aligned} &\tilde{b}(t,x,y,z) := b(x,\pi^*(t,x,y,z)), \quad \tilde{\sigma}(t,x,y,z) := \sigma(x,\pi^*(t,x,y,z)), \\ &\tilde{f}(t,x,y,z) := \nabla_x H(t,x,\pi^*(t,x,y,z),y,z), \quad \tilde{g}(x) := \nabla_x g(x). \end{aligned}$$

Using these reformulation, we can find solutions of the stochastic optimal control problem through appropriate solution schemes for FBSDE systems. Therefore, let us review the analytical four-step scheme for FBSDE systems and briefly discuss a few numerical schemes.

11.4.1 The four-step scheme

The most common method is called four-step scheme and converts a FBSDE back to a parabolic PDE. It was proposed by [Ma, Protter, Young 1994]. They assumed that the relationship between the state variable X_t and the adjoint variable Y_t can be expressed by a function $\Phi \in C^{1,2}([0,T] \times \mathbb{R}^n, \mathbb{R}^n)$ via $Y_t = \Phi_t(X_t)$. Using this relationship and assuming that there exists a solution of the coupled FBSDE system (11.26), the solution may be found through the following four steps:

1. Step: Find a function $z(t,x,y,p)$ which satisfies the following relation:

$$\begin{aligned} &z(t,x,y,p) = p\tilde{\sigma}\left(t,x,y,z(t,x,y,p)\right), \\ &\forall (t,x,y,p) \in [0,T] \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^{n\times n}. \end{aligned} \tag{11.27}$$

2. Step: Use the function $z_t = z(t,x,\Phi_t,\nabla_x\Phi_t)$ obtained above to solve the following parabolic PDE for $\Phi_t(x)$:

$$\begin{aligned} &\partial_t\Phi_t^l + \tfrac{1}{2}\mathrm{tr}[\tilde{\sigma}\tilde{\sigma}^\top(t,x,\Phi_t,z_t)\nabla_{xx}^2\Phi_t^l] + \tilde{b}(t,x,\Phi_t,z_t)\nabla_x\Phi_t^l \\ &-\tilde{f}^l(t,x,\Phi_t,z_t) = 0, \quad \forall (t,x) \in [0,T] \times \mathbb{R}^n, \quad l = 1,...,n, \\ &\Phi_T(x) = \tilde{g}(x), \quad \forall x \in \mathbb{R}^n. \end{aligned}$$

3. Step: Replace Y_t and Z_t by Φ_t and z_t obtained in step 1-2, respectively, and solve the forward SDE in (11.26) for X_t:

$$X_t = x + \int_0^t \bar{b}(s,X_s)ds + \int_0^t \bar{\sigma}(s,X_s)dW_s,$$

where:

$$\begin{aligned} \bar{b}(t,x) :=& \quad \tilde{b}(t,x,\Phi_t(x),z(t,x,\Phi_t(x),\nabla_x\Phi_t(x))), \\ \bar{\sigma}(t,x) :=& \quad \tilde{\sigma}(t,x,\Phi_t(x),z(t,x,\Phi_t(x),\nabla_x\Phi_t(x))). \end{aligned}$$

4. Step: Set:

$$\begin{aligned} Y_t &= \Phi(t,X_t), \\ Z_t &= z(t,X_t,\Phi_t(X_t),\nabla_x\Phi_t(X_t)). \end{aligned}$$

The crucial parts of the four-step scheme are 1) to find a function z in step one, if it exists, and 2) to find a solution of the quasi-linear, second order parabolic PDE in step two.

11.4.2 Numerical schemes

As for Dynamic Programming (Section 11.1.1), there exists two basic approaches to solve a FBSDE system numerically.

- **Deterministic approach**: One follows the four-step scheme above and uses finite approximations of the function Φ_t (e.g. finite differences or finite elements) in order to solve the quasi-linear, second order parabolic PDE in step two. Examples for this method can be found in [Douglas, Ma, Protter 1996] or [Milstein, Tretyakov 2006].

- **Stochastic approach**: Only recently, a few authors proposed stochastic schemes to find numerical solutions of coupled FBSDE systems. Most schemes are extensions of schemes for uncoupled FBSDE systems. However, with an extensive literature research, we found only two fully traceable algorithms for coupled FBSDEs:

  1. [Delarue, Menozzi 2005] proposed a backward iteration on a discretized time interval $[0, T]$ and state space $\mathbb{R}^n$. Their explicit backward iteration starts at time T with $Y_T(\xi^j) = \tilde{g}(\xi^j)$ and $Z_T(\xi^j) = \nabla_x \tilde{g}(\xi^j)$, for all state space points ξ^j, $j = 1...M^n$. Then it calculates Y_t and Z_t for all time points $t = t_i$, $i = N-1, ..., 1$, backward in time, using the previous values. We will explain this method in much more detail in Section 12.
  2. [Bender, Zhang 2008] proposed a numerical algorithm to simulate high-dimensional coupled FBSDEs under weak coupling or monotonicity conditions. Their method is based on a global Picard iteration, commonly used for BSDEs. In contrast to Delarue and Menozzi's scheme, the method does not discretize the state space $\mathbb{R}^n$.

  Note that, a part of any stochastic scheme is the calculation of expectation values. Common methods to calculate these expectation values are Monte Carlo simulations, Quantization methods, Markov-chain approximations or Least-Square Monte-Carlo techniques. The choice of the method usually depends on the specific problem and its parameters but it does not affect the general stochastic scheme.

We want to emphasize that both stochastic schemes are developed recently and the proofs for convergence are known only for special, weak coupling conditions so far. Nevertheless, we will apply the first scheme to problems with stronger coupling conditions and discuss the results and convergence in Part IV. We believe that algorithms for stronger coupling conditions will be provided and proofed in the near future.

11.5 Connection between the DPP and the SMP

In the previous paragraphs we already saw a few similarities between the DPP and the SMP. In this section we show a direct connection. First, we formally show analogies between the optimal control case and the non-controlled[47] case. In the latter, the FBSDE system is a Hamilton system and the HJB equation is the Hamilton-Jacobi equation. These analogies provide the reader a more intuitive understanding of the connection between the DPP and the SMP theory. We then formaly show this connection by deriving the SMP's necessary condition of optimality out of the HJB equation and by proving the SMP's sufficient condition via the HJB verification.

11.5.1 Analogies

- **The deterministic, non-controlled case**: The classical Hamilton-Jacobi theory tells us, that a *linear, parabolic first-order PDE* can be represented as a finite *system of ODEs*. The PDE is called Hamilton-Jacobi equation and the system of backward ODEs (BODE) is called Hamilton system. The method of characteristics is used to solve the PDE via the system of ODEs. Vice versa, one can use complete integrals and the implicit function theorem to tansform the ODE system into a PDE. See [Evans 2002] chapter 3 for details.

- **The deterministic, optimal control case**: The classical Hamilton-Jacobi-Bellmann theory tells us, that an optimal control problem can be expressed by the HJB equation, a *non-linear, parabolic first-order PDE*. The non-linearity arises through the intrinsic optimization. On the other hand, Pontryagin's maximum principle (see Section 11.2) provides the theory to solve an optimal control problem via a *system of forward-backward ODEs* (FBODE), which is a two-point boundary problem. See [Evans 2002] chapter 10 for details.

[47] There is no control variable and no optimization.

- **The stochastic, non-controlled case**: It is the well-known Feynman-Kac theorem that tells us that a special type of *linear, parabolic second-order PDEs* can be represented as a *system of backward SDEs* (BSDEs[48]) and vice versa. See [Klebaner 2005] chapter 6 for details.

- **The stochastic, optimal control case**: An extension of the Feynman-Kac formula tells us that also a quasi-linear, parabolic second-order PDE can be represented as a system for coupled FBSDEs including a (first-order) adjoint equation. However, this FBSDE representation cannot be directly applied to stochastic optimal control problems. Here the HJB equation (11.3) is a fully *non-linear, parabolic second-order PDE*. Instead, by using the SMP, we obtain a *system of second-order FBSDEs coupled through an optimality condition* (11.24), a stochastic two-point boundary problem. See [Pham 2009].

	Non-controlled case	Optimal control case
Deterministic case	Hamilton system: system of first-order BODEs	Pontryagin's maximum principle: system of first-order FBODEs + external optimality condition
	HJ equation: linear, first-order PDE	HJB equation: non-linear, first-order PDE with inherent optimization
Stochastic case	Hamilton system: system of second-order BSDEs	Stochastic Maximum Principle: system of second-order FBSDEs + external optimality condition
	HJ equation: linear, second-order PDE	HJB equation: non-linear, second-order PDE with inherent optimization

Table 9: Analogies of the ODE-PDE representation between the non-controlled and optimal control case for the deterministic and stochastic case.

[48] BSDEs are uncoupled FBSDEs. By uncoupled we mean that the forward SDE for X_t does not depend on the adjoint variables (Y_t, Z_t) of the backward SDEs.

We obtain that going form the non-controlled to the optimal control case, we switch from linear to non-linear equations. Moreover, going from the deterministic to the stochastic case, we switch from first-order to second-order equations, see Table 9.

Now we show a direct connection between the DPP and the SMP by deriving the SMP's necessary condition of optimality from the HJB equation.

11.5.2 Derivation of the SMP's necessary condition from the DPP

Following [Pham 2009], let us suppose there exists a unique solution $v \in C^{1,3}([0,T)\times\mathbb{R}^n)\cap C^0([0,T]\times\mathbb{R}^n)$ of the value function (10.5) and an optimal control $\pi^* \in \mathbb{A}_T$ described in (10.4) with an associated controlled diffusion $\hat{X}_t$ satisfying (10.1). Then we can derive the adjoint equations of the SMP out of the HJB equation through the following three steps.

First, we know from the HJB theory that:

$$\partial_t v(t,\hat{X}_t) + G(t,\hat{X}_t,\pi_t^*,\nabla_x v(t,\hat{X}_t),\nabla_x^2 v(t,\hat{X}_t)) = 0. \tag{11.28}$$

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

$$G(t,x,\pi,p,M) := b(t,x,\pi)p + \frac{1}{2}\mathrm{tr}\left[\sigma\sigma'(t,x,\pi)M\right] + f(t,x,\pi). \tag{11.29}$$

In detail, to obtain (11.28) we represent:

$$\begin{aligned} v(t,\hat{X}_t) &= E_t^{\tilde{\pi^*}}\left[\int_t^T f(s,\hat{X}_s,\pi_s^*)ds + g(\hat{X}_T)\Big| F_t\right] \\ &= -\int_0^t f(s,\hat{X}_s,\pi_s^*)ds + E_t^{\tilde{\pi^*}}\left[\int_0^T f(s,\hat{X}_s,\pi_s^*)ds + g(\hat{X}_T)\Big| F_t\right]. \end{aligned} \tag{11.30}$$

The second term of the last equation is a martingale. Therefore, we apply Ito's formula to $v(t,\hat{X}_t)$ and identify the terms in dt with (11.30). This yields (11.28).

Second, we use (11.28) together with the HJB equation to derive $\forall(t,x)\in[0,T]\times\mathbb{R}^n$:

$$\begin{aligned} &\partial_t v(t,\hat{X}_t) + G(t,\hat{X}_t,\pi_t^*,\nabla_x v(t,\hat{X}_t),\nabla_x^2 v(t,\hat{X}_t)) = 0 \\ &= \partial_t v(t,x) + \sup_{\pi\in A} G(t,x,\pi,\nabla_x v(t,x),\nabla_x^2 v(t,x)) \\ &\geq \partial_t v(t,x) + G(t,x,\pi_t^*,\nabla_x v(t,x),\nabla_x^2 v(t,x)). \end{aligned}$$

Since $v \in C^{1,3}$ this implies:

$$\partial_x \left(\partial_t v(t,x) + G(t,x,\pi_t^*, \nabla_x v(t,x), \nabla_x^2 v(t,x)) \right) \Big|_{x=\hat{X}_t} = 0. \tag{11.31}$$

Third, we apply Ito's formula to $\nabla_x v(t, \hat{X}_t)$ and plug the result into (11.31). After a few calculations and using the fact that $v(T,x) = g(x)$, we obtain the (adjoint) equations:

$$\begin{aligned} -d\nabla_x v(t,\hat{X}_t) = \; & \nabla_x H(t,\hat{X}_t,\pi_t^*,\nabla_x v(t,\hat{X}_t),\nabla_x^2 v(t,\hat{X}_t)\sigma(\hat{X}_t,\pi_t^*))dt \\ & -\nabla_x^2 v(t,\hat{X}_t)\sigma(\hat{X}_t,\pi_t^*)dW_t, \\ \nabla_x v(T,\hat{X}_T) = \; & \nabla_x g(\hat{X}_t), \end{aligned} \tag{11.32}$$

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

$$H(t,x,\pi,y,z) := b(t,x,\pi)y + \operatorname{tr}[\sigma'(t,x,\pi)z] + f(t,x,\pi). \tag{11.33}$$

Furthermore, we assume that b, σ, f, g and thus G are continuously differentiable with respect to π. Since matrix multiplication is associative and the trace operator is invariant under cyclic permutations, we obtain that:

$$\begin{aligned} &\partial_\pi G(t,x,\pi,\nabla_x v(t,x),\nabla_x^2 v(t,x)) = \partial_\pi H(t,x,\pi,\nabla_x v(t,x),\nabla_x^2 v(t,x)\sigma(x,\pi)), \\ &\forall (t,x) \in [0,T] \times \mathbb{R}^n. \end{aligned}$$

Since this is true for all $\pi \in A$, H attains a maximum whenever G attains one. Since the optimal control π^* fulfills the optimality condition of the DPP (11.5), we obtain the optimality condition of the SMP (11.20). Let us summarize this result.

Theorem 11.13 (Connection between DPP and SMP)
Let the Basic Conditions 10.1 and the SMP Conditions 10.9, 6 - 7 hold. If there exists a unique solution $v \in C^{1,3}([0,T) \times \mathbb{R}^n) \cap C^0([0,T] \times \mathbb{R}^n)$ of the value function (10.5) and an optimal state-control pair $(\hat{X}_t, \pi^*)$ satisfying (10.1) and (10.4), respectively, then for all $t \in [0,T]$ the triple:

$$\begin{aligned} (\hat{X}_t, Y_t, Z_t) := \; & \left(\hat{X}_t, \nabla_x v(t,\hat{X}_t), \nabla_x^2 v(t,\hat{X}_t)\sigma(\hat{X}_t,\pi_t^*) \right) \\ \in \; & L^2_{\mathbb{F}}([0,T],\mathbb{R}^n) \times L^2_{\mathbb{F}}([0,T],\mathbb{R}^n) \times \left(L^2_{\mathbb{F}}([0,T],\mathbb{R}^n)\right)^d, \end{aligned}$$

is the unique solution to the coupled FBSDE system:

$$\begin{aligned} \hat{X}_t = &\quad x + \int_0^t b(s, \hat{X}_s, \pi_s^*)ds + \int_0^t \sigma(s, \hat{X}_s, \pi_s^*)dW_s, \\ Y_t = &\quad \nabla_x g(X_T) + \int_t^T \nabla_x H(s, \hat{X}_s, \pi_s^*, Y_s, Z_s)ds \\ &- \int_t^T Z_s dW_s, \end{aligned} \tag{11.34}$$

such that the following optimality condition holds:

$$\pi_t^* = \arg\max_{\pi \in A} H(t, \hat{X}_t, \pi, Y_t, Z_t). \tag{11.35}$$

11.5.3 Derivation of the SMP's sufficient condition from DPP

Now we show the connection between SMP and DPP by proving the SMP's sufficient condition via the HJB verification (sufficient condition).

Theorem 11.14 (Connection between DPP and SMP 2)
Let the Basic Conditions 10.1 and the SMP Conditions 10.9, 6 - 8 hold (i.e., g and H are uniformly[49] concave w.r.t x and π). Let $(\hat{X}, Y_t, Z_t)$ and the control policy $\pi^* = \{\pi_t^*, t \in [0,T]\}$ be associated solutions to the FBSDE system (11.34) such that the optimality condition (11.35) holds. Furthermore, suppose that there exists a unique solution $v \in C^{1,3}([0,T] \times \mathbb{R}^n, \mathbb{R})$ of the value function (10.5). Then:

1. *$(\hat{X}_t, \pi^*)$ is the optimal state-control pair of the Concave Problem 6 (10.1) - (10.3).*

2. *The solutions (Y_t, Z_t) of the first-order adjoint equation are the first derivative and an adjusted second derivative of the value function v (10.5), respectively:*

$$(Y_t, Z_t) = \left(\nabla_x v(t, \hat{X}_t), \nabla_x^2 v(t, \hat{X}_t)\sigma(\hat{X}_t, \pi_t^*)\right). \tag{11.36}$$

Proof 11.15
The first statement is the SMP's sufficient condition. An alternative proof, which uses the DPP, can be found in [Pham 2009], theorem 6.5.4. In order to prove the second statement, we define a function $u(t,x) \in C^{1,3}([0,T] \times \mathbb{R}^n, \mathbb{R})$ through its first derivatives $\nabla_x u(t, \hat{X}_t^x) := Y_t^x$ and its terminal value $u(T,x) :=$

[49] That means concave for all $t \in [0,T]$.

$g(x)$. The superscript x at $\hat{X}_t^x$ or Y_t^x indicates the solution of (11.34) when starting at time t with $\hat{X}_t = x$[50]*. Then we apply Ito's formula to $\nabla_x u$:*

$$\begin{aligned} dY_t &= d\,\nabla_x u = \partial_t \nabla_x u\, dt + \nabla^2_{xx} u\, dX + \tfrac{1}{2} tr\big(\sigma\sigma' \nabla^2_{xx}\big)\nabla_x u\, dt \\ &= \Big(\partial_t \nabla_x u + \nabla^2_{xx} u\, b + \tfrac{1}{2} tr\big(\sigma\sigma' \nabla^2_{xx}\big)\nabla_x u\Big)\, dt + \nabla^2_{xx} u\, \sigma\, dW_t. \end{aligned}$$

Comparing the diffusion term of the above equation with the diffusion term of the BSDE for Y_t^x in (11.34), we obtain that $Z_t^x = \nabla^2_{xx} u(t, \hat{X}_t^x)\sigma(\hat{X}_t^x, \pi_t^{*,x})$. Comparing the drift terms of these equations, we get a third-order PDE:[51]

$$\begin{aligned} 0 &= \nabla_x H(t, \hat{X}_t^x, \pi_t^{*,x}, Y_t^x, Z_t^x) + \partial_t \nabla_x u + \nabla^2_{xx} u\, b + \tfrac{1}{2} tr\big[\sigma\sigma' \nabla^2_{xx}\big]\nabla_x u \\ &= b_x Y_t^x + \nabla tr\big[\sigma_x' Z_t^x\big] + f_x + \partial_t \nabla_x u + \nabla^2_{xx} u\, b + \tfrac{1}{2} tr\big[\sigma\sigma' \nabla^2_{xx}\big]\nabla_x u \\ &= b_x \nabla_x u + \nabla tr\big[\sigma_x' \nabla^2_{xx} u\, \sigma\big] + f_x + \partial_t \nabla_x u + \nabla^2_{xx} u\, b + \tfrac{1}{2} tr\big[\sigma\sigma' \nabla^2_{xx}\big]\nabla_x u \\ &= \partial_t \nabla_x u + b_x \nabla_x u + \nabla^2_{xx} u\, b + 2\tfrac{1}{2}\nabla tr\big[\sigma\sigma_x' \nabla^2_{xx} u\big] + \tfrac{1}{2} tr\big[\sigma\sigma' \nabla^2_{xx}\big]\nabla_x u + f_x \\ &= \nabla_x \left(\partial_t u + \nabla_x u\, b + \tfrac{1}{2} tr\big[\sigma\sigma' \nabla^2_{xx} u\big] + f\right). \end{aligned}$$

The last term is the gradient of the HJB equation w.r.t x. Since the solution of this PDE is unique, and equation (11.31) holds for the value function v, it follows that $u(t,x) = v(t,x)$. qed.

In the above theorem, we started with solutions $\hat{X}_t$, Y_t, and Z_t of the FBSDE system (11.34) and constructed a function that is equal to the solution of the HJB equation, if it exists. A further reasearch topic would be to investigate the following Hypothesis:

Hypothesis 11.16
If the Basic Conditions 10.1 and the SMP Conditions 10.9 hold (i.e., g and H are uniformly concave w.r.t x and π), then the HJB equation has a unique solution $v(t,x) \in C^{1,2}([0,T] \times \mathbb{R}^n)$.

11.5.4 Summary

The analogies in Table 9 show the connection between the SMP and the DPP. Moreover, Theorems 11.14 and 11.13 reveal the transferability of necessary and sufficient conditions in the smooth case. We obtain that:

$$(\hat{X}_t, Y_t, Z_t) = \big(\hat{X}_t, \nabla_x v(t, \hat{X}_t), \nabla^2_x v(t, \hat{X}_t)\sigma(\hat{X}_t, \pi_t^*)\big)$$

[50] By definition, $\hat{X}_t^x = x$.

[51] Note that b_x is a $(n \times n)$-matrix, f_x is a vector, and $\nabla \mathrm{tr}[\sigma_x' \cdot]$ indicates the derivative of the trace operator component-by-component.

In consequence, if the according conditions are satisfied, one can use either way to search for solutions of the stochastic optimal control Problem 5 (10.1) - (10.3).

The concavity of H and g is important for the SMP's sufficient conditions to hold. This concavity condition mainly specifies the problem class for which we can apply the SMP theory, namely the class of Concave Problems 6.

12 A controlled forward-backward algorithm

In this section, we propose a numerical scheme for concave stochastic optimal control problems (Problem 6 in Section 10) by utilizing the Stochastic Maximum Principle (Section 11). We call our scheme Controlled forward-backward (CFB) algorithm because it is based on the forward-backward stochastic algorithm proposed by [Delarue, Menozzi 2005].

We already showed in Theorem 11.12 that the Concave Problem 6 is equivalent to the problem of searching for the $\mathbb{F}$-adapted quadruple $(\hat{X}_t, \pi_t^*, Y_t, Z_t)$, which satisfies the following coupled FBSDE system:

$$\begin{aligned} dX_t &= b(t, X_t, \pi_t^*)dt + \sigma(t, X_t, \pi_t^*)dW_t, && X_0 = x \in \mathbb{R}^n, \\ dY_t &= -\nabla_x H(t, \hat{X}_t, \pi_t^*, Y_t, Z_t)dt + Z_t dW_t, && Y_T = \nabla_x g(X_T), \\ \pi_t^* &= \arg\max_{\pi_t \in A} H(t, X_t, \pi_t, Y_t, Z_t), \end{aligned} \tag{12.1}$$

where $H(t,x,\pi,y,z) := b(t,x,\pi)y + \mathrm{tr}[\sigma'(t,x,\pi)z] + f(t,x,\pi)$ is concave w.r.t. x and π. To give the reader an upfront idea of the CFB algorithm, we summarize the main steps in the following informal Statement 12.1.

Statement 12.1 (Controlled forward-backward algorithm)
Let $\{t_k, k = 0, ..., N\}$ be a discretization of the time interval $[0,T]$ and let $\left\{\mathrm{C}_k^h, k = 0, ..., N\right\}$ be an according sequence of spartial grids in $\mathbb{R}^n$, such that $\mathrm{C}_k^h \subset \mathrm{C}_l^h$, for all $k < l$. Then, for every point (t_k, ξ^j) on the time-space grid, we can approximate an optimal control π_k^j for the Concave problem 6, together with approximations Y_k^h and Z_k^h, by using the following controlled forward-backward algorithm:

Starting at time T and going backwards in time for $k = N, ..., 0$, the control π_k^j on each discrete space-point ξ^j can be approximated through the routine:

$$\begin{aligned} \hat{\xi}_k^j &= \xi^j + \sigma(t_k, \xi^j, \pi_{k+1}^j)\,\Delta W_k, \\ \hat{Z}_k^j &= E_k\left[Z_{k+1}^h(\hat{\xi}_k^j)\right], \\ \pi_k^j &= \arg\max_{\pi \in A} H(t_k, \xi^j, \pi, Y_{k+1}^j, \hat{Z}_k^j), \\ \xi_{k+1}^j &= \xi^j + b(t_k, \xi^j, \pi_k^j)\Delta t_k + \sigma(t_k, \xi^j, \pi_k^j)\Delta W_k, \\ Z_k^j &= \tfrac{1}{\Delta t_k} E_k\left[Y_{k+1}^h(\xi_{k+1}^j)\,\Delta W_k^\top\right], \\ Y_k^j &= \nabla_x H(t_k, \xi^j, \pi_k^j, Y_{k+1}^j, Z_k^j)\,\Delta t_k + E_k\left[Y_{k+1}^h(\xi_{k+1}^j)\right], \end{aligned} \tag{12.2}$$

The reader may immediately see that this algorithim is explicit because every variable on the right hand side is known. We suggest a Newton-Raphson method to perform the optimization and a Quantization method to simulate the expectation values.

Now let us develop the CFB algorithm step-by-step. However, we do not state every detail of [Delarue, Menozzi 2005]'s paper since this would be extensive without revealing new insights. We rather provide the reader all main features of the algorithm in an understandable way, while keeping mathematical correctness, and such that the CFB algorithm can be implemented. For the interested reader, we recommend to read through [Delarue, Menozzi 2005] to get all details.

12.1 FBSDE approximation

Following [Kushner, Dupuis 2001], let us define a fixed, scalar approximation parameter $h > 0$. Throughout the following, the superscript h denotes the dependency on this approximation parameter.

12.1.1 Time discretization

As in every numerical scheme, we discretize the time horizon $[0, T]$ into $N-1$ intervals $[t_k, t_{k+1}] = \Delta t_k^h$, such that for $k = 0, ..., N-1$:

$$t_0 := 0, \quad t_1 := \Delta t_0^h, \quad t_k := \sum_{i=0}^{k-1} \Delta t_i^h, \quad t_N := T, \tag{12.3}$$

and:

$$\Delta t_k^h \to 0 \text{ as } h \to 0.$$

Now we introduce the set of piecewise constant, admissible control policies by:

$$\mathbb{A}_T^h = \{\pi \in \mathbb{A}_T \,|\, \pi_t \text{ is constant over } [t_k, t_{k+1}),\ \forall k \leq N-1\}. \tag{12.4}$$

Note that we do not discretize the control space here but rather keep the control constant over each time interval Δt_k^h.

12.1.2 Localization

Let $\pi^* \in \mathbb{A}_T$ denote an optimal control policy. The local evolution of X_t and Y_t in (12.1) over each time interval $[t_k, t_{k+1})$ is given by:

$$\begin{aligned} X_{k+1} &= X_k + \int_{t_k}^{t_{k+1}} b(s, X_s, \pi_s^*) ds + \int_{t_k}^{t_{k+1}} \sigma(s, X_s, \pi_s^*) dW_s, \\ Y_k &= Y_{k+1} + \int_{t_k}^{t_{k+1}} \nabla_x H(s, X_s, \pi_s^*, Y_s, Z_s) ds - \int_{t_k}^{t_{k+1}} Z_s dW_s. \end{aligned} \tag{12.5}$$

So, starting from a known point X_k at t_k, the state X_{k+1} can be simulated forward through the first equation in (12.5). In order to calculate the adjoint variable Y_k backwards from Y_{k+1}, which is a random variable at time t_k, we take the expectation of the second equation in (12.5) conditional to F_k:

$$\begin{aligned} Y_k &= E_k^{\pi^*}[Y_k] \\ &= E_k^{\pi^*}[Y_{k+1}] + E_k^{\pi^*}\left[\int_{t_k}^{t_{k+1}} \nabla_x H(s, X_s, \pi_s^*, Y_s, Z_s) ds\right]. \end{aligned} \tag{12.6}$$

In order to calculate the adjoint variable Z_k backwards in a similar way, we first integrate the second equation in (12.5) by $\int_{t_k}^{t_{k+1}} \cdot \, dW_l^\top$ and then take the expectation conditional to F_k:

$$\begin{aligned} \mathbf{0} = \; & Y_k E_k^{\pi^*}\left[(W_{k+1} - W_k)^\top\right] = Y_k E_k^{\pi^*}\left[\int_{t_k}^{t_{k+1}} dW_l^\top\right] = E_k^{\pi^*}\left[\int_{t_k}^{t_{k+1}} Y_k \, dW_l^\top\right] \\ = \; & E_k^{\pi^*}\left[\int_{t_k}^{t_{k+1}} Y_{k+1} \, dW_l^\top + \int_{t_k}^{t_{k+1}} \int_{t_k}^{t_{k+1}} \nabla_x H(s, X_s, \pi_s^*, Y_s, Z_s) \, ds \, dW_l^\top\right] \\ & - E_k^{\pi^*}\left[\int_{t_k}^{t_{k+1}} \int_{t_k}^{t_{k+1}} Z_s \, dW_s \, dW_l^\top\right] \\ = \; & E_k^{\pi^*}\left[Y_{k+1} \, (W_{k+1} - W_k)^\top\right] - E_k^{\pi^*}\left[\int_{t_k}^{t_{k+1}} Z_s \, ds\right] + O([t_{k+1} - t_k]^{\frac{3}{2}}). \end{aligned}$$

The remaining term $O([t_{k+1} - t_k]^{\frac{3}{2}})$ is a consequence of the boundedness of $\nabla_x H(s, X_s, \pi_s^*, Y_s, Z_s)$. This follows from the definition of H and the SMP Conditions 10.9. Thus we obtain:

$$E_k^{\pi^*}\left[\int_{t_k}^{t_{k+1}} Z_s \, ds\right] = E_k^{\pi^*}\left[Y_{k+1} \, (W_{k+1} - W_k)^\top\right] + O([t_{k+1} - t_k]^{\frac{3}{2}}). \tag{12.7}$$

12.1.3 Approximations

Now, let $\pi^{h*} \in \mathbb{A}_T^h$ denote a piecewise constant control policy and let X^h, Y^h and Z^h denote the corresponding approximations of the state and adjoint processes that are piecewise constant over $[t_k, t_{k+1})$ for all $k = 0, ..., N-1$. Let $\Delta W_k = W_{k+1} - W_k$ denote the Brownian increments. Then, the standard Euler approximation of the state process (12.5) is given by:

$$X_{k+1}^h = X_k^h + b(t_k, X_k^h, \pi_k^{h*}) \Delta t_k^h + \sigma(t_k, X_k^h, \pi_k^{h*}) \Delta W_k, \tag{12.8}$$

and the corresponding piecewise constant approximations of the adjoint processes can be derived from (12.6) and (12.7), for all $0 \leq k < N$, as:

$$\begin{aligned} Z_k^h &= \tfrac{1}{\Delta t_k^h} E_k^{\pi^{h*}}\left[Y_{k+1}^h \Delta W_k^\top\right], \\ Y_k^h &= \nabla_x H(t_k, X_k^h, \pi_k^{h*}, \bar{Y}_k^h, Z_k^h)\,\Delta t_k^h + E_k^{\pi^{h*}}\left[Y_{k+1}^h\right], \end{aligned} \tag{12.9}$$

where $\bar{Y}_k^h$ denotes a constant predictor of $\{Y_s,\, s \in [t_k, t_{k+1}]\}$ that needs to be defined. We also have:

$$\begin{aligned} Y_N^h &= \nabla_x g(X_N^h), \\ Z_N^h &= \nabla_{xx} g(X_N^h)\sigma(X_N^h, \pi_N^{h*}). \end{aligned} \tag{12.10}$$

Note that we switched the order of calculation such that we can use Z_k^h as constant predictor on the right hand side. Note also that Z_k^h is not just any approximation of Z_{t_k} but the best constant approximation of $\{Z_s,\, s \in [t_k, t_{k+1}]\}$ in the $L^2_{\mathbb{F}}([t_k, t_{k+1}])$ sense when we neglect the term $O([t_{k+1} - t_k]^{\frac{3}{2}})$, see (12.7).

Furthermore, we say that π^{h*} is an approximation of the optimal control policy if it satisfies the following optimality condition:

$$\pi_k^{h*} = \arg\max_{\pi \in A} H(t_k, X_k^h, \pi, \bar{Y}_k^h, \bar{Z}_k^h), \quad \forall 0 \leq k \leq N. \tag{12.11}$$

Here, $\bar{Z}_k^h$ denotes a constant predictor of $\{Z_s,\, s \in [t_k, t_{k+1}]\}$ that needs to be defined. Z_k^h would definitely be the best choice but Z_k^h again depends on a π_k^{h*}, see (12.9).

Note that in contrast to the DDP, it is not necessary to calculate an approximation of the value function $V_k := v(t_k, X_k)$ when using the SMP. Nevertheless, in some applications the value function is of major interest. In compliance with DPP, we can calculate an approximation of the value function V^h by dynamic programming for $k = N-1, ..., 0$ via:

$$V_k^h = f(t_k, X_k^h, \pi_k^{h*})\Delta t_k^h + E_k^{\pi^{h*}}\left[V_{k+1}^h\right], \quad V_N^h = g(X_N^h). \tag{12.12}$$

12.2 State space discretization

A key point in Delarue and Menozzi's method for pure FBSDEs is the introduction of spatial grids C_k^h for each time point k, together with an iteration backward in time that calculates the approximations $Y_k^j = Y_k^h(\xi^j)$, $Z_k^j = Z_k^h(\xi^j)$ on all grid points $\xi^j \in \mathrm{C}_k^h$, for all $k = N-1, ..., 0$. When the backward iteration is finished, the process X^h can be easily simulated forward in

time by using the approximation values Y_k^j and Z_k^j.

We extend this method by means of an optimization method that calculates the optimal control $\pi_k^j = \pi_k^{h*}(\xi^j)$ additionally on every grid point, during the backward iteration. In order to simulate the process X^h afterwards, we must only store the optimal control values π_k^j and can discard Y_k^j and Z_k^j.

In detail, for each $0 \leq k \leq N$, let $I_k^n \subset \mathbb{N}^n$, $|I_k^n| = M_k$, denote a n-dimensional, finite index set and let us define the spatial grids by:

$$\mathrm{C}_k^h = \left\{ \xi^j \in \mathbb{R}^n \,|\, j \in I_k^n \right\} \subset \mathbb{R}^n, \tag{12.13}$$

such that $\mathrm{C}_j^h \subset \mathrm{C}_i^h$, for all $j < i$.

Note that due to the finiteness of the index set, the state space $\mathbb{R}^n$ is truncated at some point. Several truncation procedures may be considered, but all should consider the specific drift and geometry of the forward diffusion X_t.

When it comes to the choice of a grid, a Cartesian grid is the most natural one. In particular, because of the strong coupling of the FBSDE system, little is known on the behavior of the path for the forward process. Hence, we cannot compute a kind of optimal grid for X, see [Delarue, Menozzi 2005] section 3.4.

12.3 State transition

Given a spatial grid, there exists several methods to calculate the transition of the state process (12.8) from a point $\xi_k \in \mathrm{C}_k^h$ to a point $\xi_{k+1} \in \mathrm{C}_{k+1}^h$ and thereby the expectation values in (12.9) and (12.12). Delarue and Menozzi's choice is a Quantization method. They claim that standard Monte Carlo methods are inefficient because of their slow convergence. A contra argument would be that there exist efficient Monte Carlo methods like importance sampling. On the other hand, these methods are commonly used only for uncontrolled processes and, roughly speaking, the strong coupling of the forward and backward equations prohibits an optimization of Monte Carlo methods.

Thus, we will use the Quantization method here. In order to present an alternative to Quantization anyway, we outline the method of Markov-chain approximations in the next subsection, too.

12.3.1 Quantization

Quantization consists in approximating a random variable by a suitable discrete law. In particular, it replaces the Brownian increments $\Delta W_k \sim N(0, 1^d)$ appearing in (12.8) by its projection ΔW_k^h onto a finite grid:

$$\Lambda_k^h = \left\{ y^l \in \mathbb{R}^d \,|\, l \in L_k^d \right\} \subset \mathbb{R}^d, \tag{12.14}$$

where $L_k^d \subset \mathbb{N}^d$, $|L_k^d| = L_k$, denotes a d-dimensional, finite index set for each $0 \le k \le N$. The associated probability weights are:

$$P_k^h = \left\{ p^l \,,\, l \in L_k^d \right\}. \tag{12.15}$$

The set P_k^h is equal to a probability measure on $\mathbb{R}^d$ with finite support Λ_k^h. In order to measure the error associated to the grid Λ_k^h, we refer to the so-called p-distortion:

$$D_{\Delta W_k, p}(\Delta W_k^h) := \|\Delta W_k - \Delta W_k^h\|_{L^p(P)}, \quad p \ge 1, \tag{12.16}$$

see the monograph of [Graf, Luschgy 2000] for details. We will discuss the measurement of the error and the choice of optimal discrete grids Λ_k^h in Section 13.

If we start at a grid point $\xi^j \in \mathrm{C}_k$ and simulate the state process (12.8) one time step forward with the above Quantization method, then the simulated points ξ_{k+1}^j does not necessarily lie on the grid C_{k+1}. Therefore, let $\Pi_k : \mathbb{R}^n \to \mathrm{C}_k$ denote a projection mapping onto the grid C_k for all $0 \le k \le N$. We use the orthogonal projection in the hypercube here.

Then, for all $\xi^j \in \mathrm{C}_k$, let us define the propagation of the state process up to time t_{k+1} conditional to the Brownian increment y^l by:

$$\xi_{k+1}^{k,j}(\pi_k^j \,|\, y^l) = \Pi_{k+1}\left[\xi^j + b(\xi^j, \pi_k^j)\Delta t_k^h + \sigma(\xi^j, \pi_k^j)\, y^l \right] \in \mathrm{C}_{k+1}, \quad \forall l \in \Lambda_k^h. \tag{12.17}$$

Using these transitions and the associated probability weights $p^l \in P_k^h$, the expectations appearing in the induction schemes (12.9) and (12.12) become computable finite sums. To make this clearer, let us suppose our backward algorithm has reached time point t_k and grid point $\xi^j \in \mathrm{C}_k^h$ and we already know the values Y_{k+1}^i, Z_{k+1}^i and V_{k+1}^i for all $i \in I_{k+1}^n$. Then the induction

schemes (12.9) and (12.12) read:

$$\begin{aligned} Z_k^j &= \frac{1}{\Delta t_k^h} \sum_{l \in L_k^d} p^l \, Y_{k+1}^h \left(\xi_{k+1}^{k,j}(\pi_k^j \,|\, y^l) \right) y^{l\top}, \\ Y_k^j &= \nabla_x H(t_k, \xi^j, \pi_k^j, \bar{Y}_k^j, Z_k^j) \, \Delta t_k^h + \sum_{l \in L_k^d} p^l \, Y_{k+1}^h \left(\xi_{k+1}^{k,j}(\pi_k^j \,|\, y^l) \right), \\ V_k^j &= f(t_k, \xi^j, \pi_k^j) \Delta t_k^h + \sum_{l \in L_k^d} p^l \, V_{k+1}^h \left(\xi_{k+1}^{k,j}(\pi_k^j \,|\, y^l) \right). \end{aligned} \tag{12.18}$$

and the optimality condition (12.11) becomes:

$$\pi_k^j = \arg\max_{\pi \in A} H(t_k, \xi^j, \pi, \bar{Y}_k^j, \bar{Z}_k^j). \tag{12.19}$$

(12.17) - (12.19) formulate an implicit forward-backward iteration that calculates the values π_k^j on all grid points $\xi^j \in \mathrm{C}_k$ for $k = N-1, ..., 0$.

12.3.2 Markov-chain approximations as an alternative to Quantization

Markov-chain approximation is an alternative method to Quantization which we do not implement. We want to present it anyway in order to show how an alternative method work. Therefore, let $p_k^{lj}(\pi_k^j) = P(\xi^l \,|\, \xi^j, \pi_k^j)$ denote fixed, discrete transition probabilities of the state process from point $\xi^j \in \mathrm{C}_k^h$ at time t_k to point $\xi^l \in \mathrm{C}_{k+1}^h$ at time t_{k+1} while the fixed control π_k^j is used. If $\sum_{l \in I_k^n} p_k^{lj}(\pi_k^j) = 1$, then we can say for the process ξ that:

$$\begin{aligned} &\forall k = 0, ..., N-1, \ \forall \xi^j \in \mathrm{C}_k^h, \ \forall \xi^l \in \mathrm{C}_{k+1}^h, \\ &\xi_k = \xi^j \to \xi_{k+1} = \xi^l \text{ with probability } p_k^{lj}(\pi_k^j). \end{aligned} \tag{12.20}$$

Moreover, if $\pi_k^j = \pi^h(t_k, \xi^j, Y_k^j, Z_k^j)$ is a Markov control[52], then the transition probabilities $p_k^{lj}(\pi_k^j)$ depend only on information of the current time t_k too and (12.20) defines a so-called Markov-chain approximation of process (10.1) with increments $\Delta\xi_k = \xi_{k+1} - \xi_k$.

The discrete Markov chain approximation ξ converges to the real state process X in (10.1) as $h \to 0$ if the following local consistency conditions hold:

$$\begin{aligned} &E_k\left[\Delta\xi_k\right] = b(\xi_k, \pi_k)\Delta t_k^h + o(\Delta t_k^h), \\ &\mathrm{Var}_k\left[\Delta\xi_k^h\right] = [\sigma\sigma^\top](\xi_k, \pi_k)\Delta t_k^h + o(\Delta t_k^h), \\ &\sup_{k<N} \sup_{\omega \in \Omega} |\Delta\xi_k| \to 0, \text{ as } h \to 0, \end{aligned} \tag{12.21}$$

[52] It depends only on information of the current time k.

where $E_k[\,\cdot\,]$ is the conditional expectation given all information up to time k and $\mathrm{Var}_k[\,\cdot\,]$ is the conditional variance accordingly. The proof and methods to derive proper transition probabilities can be found in [Kushner, Dupuis 2001].

One limitation of using Markov chain approximations in our algorithm is that σ needs to be invertible. Only then we can calculate the discrete Brownian increments y^l, when starting at ξ^j, via:

$$y^l = \sigma^{-1}(\xi^j, \pi_k^j)\,\left(\xi^l - \xi^j - b(\xi^j, \pi_k^j)\right), \quad \forall l \in \Lambda_k^h. \tag{12.22}$$

The Brownian increments are needed to calculate Z in (12.18). This also means that $d = n$.

As before, let us now suppose our backward algorithm reached time point t_k and grid point $\xi^j \in \mathrm{C}_k^h$ and we already know the values Y_{k+1}^i, Z_{k+1}^i and V_{k+1}^i for all $i \in I_{k+1}^n$. Then the inductions (12.9) and (12.12) read:

$$\begin{aligned} Z_k^j &= \frac{1}{\Delta t_k^h} \sum_{l \in I_k^n} p_k^{lj}(\pi_k^j)\; Y_{k+1}^h\left(\xi^l\right)\; {y^l}^\top, \\ Y_k^j &= \nabla_x H(t_k, \xi^j, \pi_k^j, \bar{Y}_k^j, \bar{Z}_k^j)\,\Delta t_k^h + \sum_{l \in I_k^n} p_k^{lj}(\pi_k^j)\; Y_{k+1}^h\left(\xi^l\right), \\ V_k^j &= f(t_k, \xi^j, \pi_k^j)\Delta t_k^h + \sum_{l \in I_k^n} p_k^{lj}(\pi_k^j)\; V_{k+1}^h\left(\xi^l\right), \end{aligned} \tag{12.23}$$

where y^l is calculated via (12.22) and the optimality condition (12.11) becomes:

$$\pi_k^j = \arg\max_{\pi \in A} H(t_k, \xi^j, \pi, \bar{Y}_k^j, \bar{Z}_k^j). \tag{12.24}$$

With (12.23) and (12.24) we formulate an alternative implicit iteration to calculate the value π_k^j on all grid points $\xi^j \in \mathrm{C}_k$ backwards in time for $k = N-1, ..., 0$.

12.4 Choice of predictors

We need to find adequate predictors $\bar{Y}_k^h$ and $\bar{Z}_k^h$ in order to implemented the scheme (12.17) - (12.19). The second key of Delarue and Menozzi's method is the special choice of predictors in order to make the scheme explicit.

In detail, let us suppose again that the algorithm has already calculated the approximations $Y_{k+1}^h(\cdot)$, $Z_{k+1}^h(\cdot)$ for all spatial grid points $\xi^j \in \mathrm{C}_{k+1}^h$ at time t_{k+1}. At time t_k and grid-point $\xi^j \in \mathrm{C}_k^h$, we replace $\bar{Y}_k^j$ and $\bar{Z}_k^j$ on the right hand side of (12.18) and (12.19) with the values Y_{k+1}^j and Z_{k+1}^j as natural

predictors, respectively. Remember that $\mathrm{C}_k^h \subset \mathrm{C}_{k+1}^h$, so this replacement is applicable.

Now we transfer this technique to our controlled FBSDE system and obtain the following explicit scheme: Starting at terminal time T and going backwards in time, at time t_k and for every grid point $\xi^j \in \mathrm{C}_k$, we first pre-calculate the optimal control value by:

$$\pi_k^j = \arg\max_{\pi \in A} H(t_k, \xi^j, \pi, Y_{k+1}^j, Z_{k+1}^j). \tag{12.25}$$

Then, we use this pre-calculated π_k^j together with the natural predictors Y_{k+1}^j and Z_{k+1}^j on the right hand side of (12.17) - (12.19) which become:

$$\begin{aligned}
\xi_{k+1}^{k,j}(\pi_k^j \,|\, y^l) &= \Pi_{k+1}\left[\xi^j + b(\xi^j, \pi_k^j)\Delta t_k^h + \sigma(\xi^j, \pi_k^j)\, y^l\right], \qquad \forall l \in \Lambda_k^h, \\
Z_k^j &= \frac{1}{\Delta t_k^h} \sum_{l \in L_k^d} p^l\, Y_{k+1}^h\left(\xi_{k+1}^{k,j}(\pi_k^j \,|\, y^l)\right) {y^l}^\top, \\
Y_k^j &= \nabla_x H(t_k, \xi^j, \pi_k^j, Y_{k+1}^j, Z_k^j)\, \Delta t_k^h + \sum_{l \in L_k^d} p^l\, Y_{k+1}^h\left(\xi_{k+1}^{k,j}(\pi_k^j \,|\, y^l)\right), \\
V_k^j &= f(t_k, \xi^j, \pi_k^j)\Delta t_k^h + \sum_{l \in L_k^d} p^l\, V_{k+1}^h\left(\xi_{k+1}^{k,j}(\pi_k^j \,|\, y^l)\right).
\end{aligned} \tag{12.26}$$

The scheme (12.25) - (12.26) is fully explicit and thus it can be implemented now.

12.4.1 An intermediate predictor to improve convergence

In order to improve convergence [Delarue, Menozzi 2005] introduced an intermediate "regularized" predictor $\hat{Z}_k^j$ that replaces Z_{k+1}^j as predictor. Applied to our Quantization method, the regularized predictor is calculated via:

$$\begin{aligned}
\hat{\xi}_k^j(\pi_k^j \,|\, y^l) &= \Pi_{k+1}\left[\xi^j + \sigma(\xi^j, \pi_k^j)\, y^l\right], \qquad \forall l \in \Lambda_k^h, \\
\hat{Z}_k^j &= E_k^{\pi_k^j}\left[Z_{k+1}^h(\hat{\xi}_k^j)\right] = \sum_{l \in L_k^d} p^l\, Z_{k+1}^h\left(\hat{\xi}_k^j(\pi_k^j \,|\, y^l)\right),
\end{aligned} \tag{12.27}$$

Therefore, in our CFB-algorithm, we replace Z_{k+1}^j by $\hat{Z}_k^j$ in (12.25).

12.5 Separate optimization

Regarding the optimization, the difference between the CFB-algorithm and the standard dynamic programming algorithm is that in the former method the optimization does not have to be performed over an expectation operator. Instead, for a fixed time and space tuple (t_k, ξ^j), the optimization (12.25) is performed over a known function H, which is given by:

$$H(\pi) = b(\xi^j, \pi)Y_{k+1}^j + \mathrm{tr}[\sigma'(\xi^j, \pi)\hat{Z}_k^j] + f(t_k, \xi^j, \pi). \tag{12.28}$$

In the following section, we want to make this optimization concrete. Therefore, we consider a Newton-Raphson method together with a line search algorithm in order solve the optimization.

12.5.1 Line search with Newton-Raphson

Let us briefly recapitulate the line search method. Suppose we have the unconstraint optimization problem:

$$\min_{\pi \in A} f(\pi), \tag{12.29}$$

for a scalar-valued function $f : \mathbb{R}^r \to \mathbb{R}$. Note that we can use $\min -f$ to solve a $\max f$ problem. Now let $\pi^0 \in A$ be a starting point for the search of an optimal control value π^*. The following iteration is called line search method. For $i = 1, ..., K$ do:

$$\pi^{i+1} = \pi^i + a^i p^i. \tag{12.30}$$

The $a^i > 0$ are called step lengths and the $p^i \in \mathbb{R}^r$ are called search directions. The key of the line-search algorithm is to choose the p^i's and a^i's such that the decent $f(\pi^i) \to f(\pi^{i+1})$ is sufficiently small for each i. A common criteria for a sufficient decent is the fulfilling of the Wolf conditions, which are defined by:

$$\begin{aligned} f(\pi^{i+1}) &\leq f(\pi^i) + c_1 a^i \nabla_\pi f(\pi^i)^\top p^i, \\ \nabla_\pi f(\pi^{i+1}) p^i &\geq c_2 \nabla_\pi f(\pi^i)^\top p^i, \end{aligned} \tag{12.31}$$

where $0 < c_1 < c_2 < 1$ are constants.

Theorem 12.2 (Wolf conditions)
If f is smooth enough, $\nabla_\pi f$ is Lipschitz continuous in π and the Wolf conditions are fulfilled, then the Newton-Raphson line-search converges to a local minimum.

The proof can be found for example in [Nocedal, Wright 2006], Theorem 3.2. The most popular choice for a search direction p^i is such that p^i satisfies the following equation:

$$B^i p^i = -\nabla_\pi f(\pi^i),$$

where B^i is a symmetric and nonsingular matrix. Examples are the so-called steepest descent method, where B^i is simply the identity matrix I, the Newton method, where B^i is the exact Hessian matrix $\nabla^2_\pi f(\pi^i)$, and the quasi-Newton methods, where B^i is an approximation to the Hessian which is updated in an iterative manner by means of a low-rank formula. For illustrative purpose, we use the exact Newton method with step length $a^i = 1$ here. Then (12.30) becomes:

$$\begin{aligned} \pi^{i+1} &= \pi^i + p^i \\ \nabla^2_\pi f(\pi^i)\, p^i &= -\nabla_\pi f(\pi^i), \qquad \text{for } i = 1, ..., K. \end{aligned} \tag{12.32}$$

Theorem 12.3 (Convergence of the optimization)
If the Hessian $\nabla^2_\pi f(\pi^i)$ is positive definite in a region close to the optimal solution and also Lipschitz continuous, then the following is true:

- *If the starting point π_0 is sufficiently close to π^*, the sequence of iterates $\{\pi^i\}$ converges to π^*;*
- *The rate of convergence of the sequence $\{\pi^i\}$ is quadratic;*
- *The sequence of gradient norms $\{|\nabla_\pi f(\pi^i)|\}$ converges quadratically to zero.*

The proof can be found for example in [Nocedal, Wright 2006], Theorem 3.5. The book also includes a good introduction into optimal control theory and further details.

12.5.2 Back to our optimization problem

Remember that for the Concave Problem 6, the SMP Conditions 10.9 hold, i.e. $H(t, \cdot, \cdot, y, z)$ and $g(\cdot)$ are concave w.r.t. x and π. Now, suppose our backward algorithm reached time point t_k and grid point $\xi^j \in \mathrm{C}^h_k$. Then, we select $\pi^{j\,0}_k = \pi^{h\,*}_{k+1}(\xi^j) \in A$ as starting point and perform the following iteration:

$$\pi^{j\,i+1}_k = \pi^{j\,i}_k + p^i, \qquad \text{for } i = 1, .., K^j_k. \tag{12.33}$$

We choose the search direction p^i as the solution of the linear system:

$$\nabla^2_{\pi\pi} \tilde{H}^j_k(\pi^{j\,i}_k)\, p^i = -\nabla_\pi \tilde{H}^j_k(\pi^{j\,i}_k). \tag{12.34}$$

where:

$$\begin{aligned} \nabla_\pi \tilde{H}_k^j(\pi) :=& \quad b_\pi(\xi^j,\pi) Y_{k+1}^j + \sum_{a,b} \sigma_\pi^{ab}(\xi^j,\pi)\, \hat{Z}_k^{ab} + f_\pi(t_k,\xi^j,\pi), \\ \nabla^2_{\pi\pi} \tilde{H}_k^j(\pi) :=& \quad b_{\pi\pi}(\xi^j,\pi) Y_{k+1}^j + \sum_{a,b} \sigma_{\pi\pi}^{ab}(\xi^j,\pi)\, \hat{Z}_k^{ab} + f_{\pi\pi}(t_k,\xi^j,\pi), \end{aligned} \tag{12.35}$$

Here we denote the gradient $\nabla_\pi \phi$ by ϕ_π for all functions $\phi = b, \sigma, f, g$ and the Hessians by $\phi_{\pi\pi}$ correspondingly. We repeat the iteration in (12.33) until $|\pi_k^{j\,i+1} - \pi_k^{j\,i}| = |p^i| \leq \epsilon^h$. The parameter K denotes the number of iterations.

Theorem 12.4 (Convergence for our concave maximization)
The above iteration converges to the unique global maximum π^* and the rate of convergence is quadratic.

Proof 12.5
According to the SMP Conditions 10.9, b_π, σ_π, f_π, $b_{\pi\pi}$, $\sigma_{\pi\pi}$, $f_{\pi\pi}$ are smooth functions and locally Lipschitz continuous w.r.t. π. Also, H, g are concave w.r.t. π. Thus H is smooth in π, $\nabla_\pi \tilde{H}_k^j(\cdot)$, $\nabla^2_{\pi\pi} \tilde{H}_k^j(\cdot)$ are locally Lipschitz continuous and $\nabla^2_{\pi\pi} \tilde{H}_k^j(\cdot)$ is positive definite. If we choose h small enough, such that the starting value $\pi_k^{j\,0} = \pi_{k+1}^{h\,*}(\xi^j)$ is close enough to $\pi_k^{h\,*}(\xi^j)$, we can apply Theorem 12.3 and receive quadratic convergence.

If b, σ, f are too complex to derive the first or second derivatives by hand, one can use methods of *automatic differentiation* to calculate them. An introduction into these methods can be found for example in [Nocedal, Wright 2006] chapter 8.

12.6 Full CFB-algorithm

Now, let us consider our Concave Problem 6, i.e., equation (10.1) - (10.5) while the Basic Conditions 10.1 and the SMP Conditions 10.9 hold. Then, we propose the following controlled forward-backward algorithm in order to find an approximation π_k^j of the optimal control on every time-space grid point (k, j) via the following iteration:

$\forall \xi^j \in \mathrm{C}_N^h$ calculate:

$$\begin{aligned} \pi_N^j =& \quad \arg\max_{\pi \in A} G(t_N, \xi^j, \pi, \nabla_x g(\xi^j), \nabla^2_{xx} g(\xi^j)), \\ V_N^j(\xi^j) =& \quad g(\xi^j), \\ Y_N^j =& \quad \nabla_x g(\xi^j), \\ Z_N^j =& \quad \nabla^2_{xx} g(\xi^j)\, \sigma(\xi^j, \pi_N^j). \end{aligned} \tag{12.36}$$

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

$\forall \xi^j \in \mathrm{C}_k^h$,

$\forall l \in \Lambda_k^h$, calculate:

$$\begin{aligned} \hat{\xi}_k^j(\pi_k^j \,|\, y^l) &= \Pi_{k+1}\left[\xi^j + \sigma(\xi^j, \pi_k^j)\, y^l\right], \\ \hat{Z}_k^j &= \sum_{l \in L_k^d} p^l \, Z_{k+1}^h\left(\hat{\xi}_k^j(\pi_k^j \,|\, y^l)\right). \end{aligned} \tag{12.37}$$

Then set $\pi_k^{j\,0} = \pi_{k-1}^j$, and

for $i = 1, ..., K_k^j$,

solve until $|p^i| \le \epsilon^h$:

$$\begin{aligned} \nabla^2_{\pi\pi} \tilde{H}_k^j(\pi_k^{j\,i})\, p^i &= -\nabla_\pi \tilde{H}_k^j(\pi_k^{j\,i}), \\ \pi_k^{j,i+1} &= \pi_k^{ji} + p^i. \end{aligned} \tag{12.38}$$

where $\nabla_\pi \tilde{H}_k^j$ and $\nabla^2_{\pi\pi} \tilde{H}_k^j$ are defined by:

$$\begin{aligned} \nabla_\pi \tilde{H}_k^j(\pi) &= b_\pi(\xi^j, \pi) Y_{k+1}^j + \sum_{a,b} \sigma_\pi^{ab}(\xi^j, \pi)\, \hat{Z}_k^{ab} + f_\pi(t_k, \xi^j, \pi), \\ \nabla^2_{\pi\pi} \tilde{H}_k^j(\pi) &= b_{\pi\pi}(\xi^j, \pi) Y_{k+1}^j + \sum_{a,b} \sigma_{\pi\pi}^{ab}(\xi^j, \pi)\, \hat{Z}_k^{ab} + f_{\pi\pi}(t_k, \xi^j, \pi). \end{aligned} \tag{12.39}$$

Then set $\pi_k^j = \pi_k^{j,i+1}$.

Now, $\forall l \in \Lambda_k^h$, calculate:

$$\xi_{k+1}^{k,j}(\pi_k^j \,|\, y^l) = \Pi_{k+1}\left[\xi^j + b(\xi^j, \pi_k^j) + \sigma(\xi^j, \pi_k^j)\, y^l\right] \in \mathrm{C}_{k+1}, \tag{12.40}$$

and then set:

$$\begin{aligned} V_k^j &= f(t_k, \xi^j, \pi_k^j)\Delta t_k^h + \sum_{l \in L_k^d} p^l \, V_{k+1}^h\left(\xi_{k+1}^{k,j}(\pi_k^j \,|\, y^l)\right); \\ Z_k^j &= \frac{1}{\Delta t_k^h} \sum_{l \in L_k^d} p^l \, Y_{k+1}^h\left(\xi_{k+1}^{k,j}(\pi_k^j \,|\, y^l)\right) y^{l\top}, \\ Y_k^j &= \nabla_x H(t_k, \xi^j, \pi_k^j, Y_{k+1}^j, Z_k^j)\, \Delta t_k^h + \sum_{l \in L_k^d} p^l \, Y_{k+1}^h\left(\xi_{k+1}^{k,j}(\pi_k^j \,|\, y^l)\right). \end{aligned} \tag{12.41}$$

After this procedure we can simulate a sample path X_t: We start with $X_0 = x \in \mathrm{C}_0^h$ at time t_0 and simulate the process X_{t_k} forward using the Euler approximaton (12.8) and an appropriate interpolation of the stored optimal controls $\left\{\pi_k^j \in A \,|\, j \in I_k^n\right\}$ for each $k = 1, ..., N$.

Note that, in the case $\sigma(x, \pi)$ depends on the control π, we use the generator G (11.29) to solve the optimization problem at terminal time T. If $\sigma(x)$ is independent of π, we first set up Y_T and Z_T and then optimize over H.

12.6.1 Comments

A key point of our CFB algorithm is that we approximate V, Y, and Z backwards in time via (12.41). In contrary, the dynamic programming scheme approximates only V backward in time, based on the DP equation (11.1). By doing so, we are able to exploit the additional information provided in Y and Z - the first and the adjusted second derivative of V - for the optimization.

In particular, by utilizing the SMP we obtain an optimality condition as an autonomous equation in (12.1) that depends only on Y and Z. Therefore, the optimization for π_k^* can be performed in a separate step (12.38). Because of the special choice of (explicit) predictors Y_{k+1} and $\hat{Z}_k$, this optimization step is done previously to the backward calculations of Z_k, Y_k and V_k in (12.41).

As a consequence, the autonnomous optimization procedure does not involve the calculation of expectation values, which is the crux of the matter in dynamic programming (DP). In particular, this feature reveals an structural advantage in the complexity (operation count) of the CFB algorithm, compared to DP algorithms. We will analyze this performance advantage in Section 14.

12.7 Summary

In the last section, we proposed a complete stochastic algorithm (12.36) - (12.41) to solve optimal control problems through the associated FBSDE system. By complete we mean that the algorithm explicitly includes the optimization step. Therefore we call it controlled forward-backward (CFB) algorithm.

Our main contribution is that we applied an Euler scheme for pure FBSDE systems, that was pre-studied by [Delarue, Menozzi 2005], and adapted it to

the case of controlled FBSDE systems, which arise through the Stochastic Maximum Principle. Therfore we are able to solve a stochastic optimal control problem through a FBSDE representation.

The significance of our algorithm is that we transfer approximations of Y_t and Z_t backwards in time instead of approximations of V_t. We then exploit the (additional) information provided in Y_t and Z_t - the first and the adjusted second derivative of V_t - for the optimization. This is an alternative to dynamic programming.

As a consequence of the special choice of (explicit) predictors, we can precalculate the optimal control $\pi_k^{h*}(\xi^j)$ in a separate step (12.38), previous to the backward calculations of Z_k, Y_k and V_k in (12.41). This is different to dynamic programming and the performance advantages will be outlined more precisely in Section 14.

12.8 Extensions

The following extensions can be considered:

- [Zhang, Elliott, Siu 2012] state a SMP for a Markov regime-switching jump-diffusion model, recently. The CFB algorithm may be extended to this type of processes. The convergence analysis would be much harder.

- To allow for constraints on the control variables (and to some extend on the state variables) in the CFB algorithm, we can use a method for constrainted optimization instead of the exact Newton line-search. An example would be the active set method. However, in a discrete time setting, we can not guarantee to satisfy the state constraints almost surely, even when we use a constraint optimization. This fact is due to the stochastic nature of the states.

  In general, little is known about problems with state constraints in stochastic control theory and even less about the SMP; see for example the discussion in [Yong, Zhou 1999] section 6. In spite of that, using a fixed finite discretization of the Brownian increments (like a Quantization), one could check the state constraints for every realization of the state variable in the optimization procedure. If the state hits the boundary for any realization, the control constaints become active. This drastically increases the complexity of the problem.

13 Convergence

In this section we discuss the convergence of our CFB algorithm. Using a particular choice of grids for the time intervall, the state space and the Quantization, we can achieve linear convergence in some cases. To obtain this result, we start with the convergence results of [Delarue, Menozzi 2005]'s algorithm for pure FBSDE systems and transfer them to our controlled FBSDE system. The outline is as follows.

In Section 13.1, we choose the discretization for time, state space and Quantization, which is crutial for convergence analysis. In Section 13.2, we state [Delarue, Menozzi 2005]'s convergence results for a pure FBSDE system, which basically involve the transfer of PDE regularity estimates to the stochastic case. In Section 13.3, we show how the estimates can be transfered from the pure FBSDE case to our Concave (stochastic optimal control) Problem 6. In particular, we introduce additional conditions to the coefficients b, σ, f and g in order to apply the results from the pure FBSDE case. In section 13.4, we comment on possible relaxations of these additional conditions such that the convergence analysis still holds. We close with a summary and state other possible extensions of our CFB algorithm which can improve convergence.

13.1 Choice of the grids

In general, the convergence of an algorithm crucially depends on the choice of the discretization. Therefore, we specify fixed time, space and Quantization grids before we start with the convergence analysis.

13.1.1 Time grid

Let us fix the scalar approximation parameter $h > 0$ and consider an equidistant time grid $\Delta t_k^h = h$, for all $0 \leq k \leq N$:

$$t_0 := 0, \quad t_1 := h, \quad t_k := kh, \quad t_N := T. \tag{13.1}$$

Note that we used our approximation parameter h as time step size. Therefore, all further discretization parameters are orientated to the time step size. However, sometimes we drop the index h in order to keep notations clear and simple.

13.1.2 Spatial grid

For a fixed parameter $\delta^h > 0$, we choose $R^h > 0$ and $\rho^h > 0$ as multitudes of δ^h and define the spatial grids on every time point t_k by:

$$\begin{aligned} \mathrm{C}_k^h &= \delta\mathbb{Z}^n \cap \Delta^\rho, \qquad \forall 0 < k \le N, \\ \mathrm{C}_0^h &= \delta\mathbb{Z}^n \cap \Delta^0, \qquad k = 0, \end{aligned} \tag{13.2}$$

where:

$$\Delta^\rho := \{\xi \in \mathbb{R}^n \,|\, \forall 1 \le i \le n,\ -(R+\rho) \le \xi_i \le R+\rho\}. \tag{13.3}$$

So C^h is basically a Cartesian grid of a n-dimensional cube in $\mathbb{R}^n$. The radius of C_0^h in each coordinate is R and the radius of C_k^h is $R+\rho$ for all $k > 0$. The minimal radius R bounds the area on which we look for solutions. The extended radius $R+\rho$ is needed to take the influence of outside points into account. Thus, ρ regulates the truncation error. Using this Cartesian grid, the i's coordinate of the projection mapping $\Pi_k[\cdot]$ in (12.17) is given, for every ξ inside the grid, by:

$$\left(\Pi_k[\xi]\right)_i = \delta\left\lfloor \delta^{-1}\xi_i + 1/2 \right\rfloor. \tag{13.4}$$

In literature, $\Pi_k[\cdot]$ is sometimes called *nearest neighbor* method. Note that the number of grid points for C_k^h is $M_k = |I_k^n| = (2\frac{R+\rho}{\delta} + 1)^n$. Since M_k is constant for all $k > 0$, we denote it simply by M.

13.1.3 Quantization grid

Remember that the Quantization grid $\Lambda_k^h = \left\{y_k^l \in \mathbb{R}^d \,|\, l \in \mathrm{L}_k^d\right\}$ (12.14) is the finite support of the discrete Brownian increment ΔW_k^h. In our discretization, we use a constant Quantization grid Λ^h for all $0 \le k \le N$. Therefore, let $\tilde{L}^h \in \mathbb{N}$ denote a fixed number and let us define a partial grid in one dimension by:

$$\tilde{\Lambda}^+ = \left\{\frac{1}{\tilde{L}}, \frac{2}{\tilde{L}}, ..., \frac{\tilde{L}-2}{\tilde{L}}, \frac{\tilde{L}-1}{\tilde{L}}, 1, \frac{\tilde{L}}{\tilde{L}-1}, \frac{\tilde{L}}{\tilde{L}-2}, ..., \frac{\tilde{L}}{2}, \tilde{L}\right\}.$$

Then, our d-dimensional Quantization grid Λ^h is given by:

$$\Lambda^h = \left[-\tilde{\Lambda}^+ \cup \tilde{\Lambda}^+\right]^d =: \left\{y^l \in \mathbb{R}^d \,|\, l \in \mathrm{L}^d\right\}, \tag{13.5}$$

where L^d is a d-dimensional index set. The number of Quantization points is $L^d = |\,\mathrm{L}^d\,| = 2^d(2\tilde{L}^h - 1)^d$. So $L = 4\tilde{L}^h - 2$.

For the fixed Quantization grid Λ^h, we can calculate the associated probability weights:

$$P^h = \left\{ p^l \,,\, l \in \mathrm{L}^d \right\}, \tag{13.6}$$

once for all k, by using a Monte Carlo method and the nearest neighbor method as projection mapping onto the grid Λ^h. For this choice of Quantization, [Delarue, Menozzi 2005] showed that for every $p \geq 1$:

$$E\left[|y^l(\Delta W^h) - \Delta W^h|^p \right]^{1/p} \leq C_{p,d}\, h^{1/2}\, L^{-1/d}, \tag{13.7}$$

where $C_{p,d}$ is a constant only depending on the dimension of the Brownian motion d and the p-distortion of the projection ΔW^h onto the finite grid Λ^h in (12.16). Note that other *optimal grids* can be used for the Quantization, as long as they fulfill the estimate (13.7). Now we can analyze the convergence of the CFB algorithm for this grid choice.

13.2 Convergence of the pure FBSDE case

Let us consider the following pure FBSDE system:

$$\begin{aligned} dX_t &= b(X_t, Y_t, Z_t)dt + \sigma(X_t, Y_t)dW_t, \qquad && X_0 = x \in \mathbb{R}^n, \\ dY_t &= \tilde{f}(\hat{X}_t, Y_t, Z_t)dt + Z_t dW_t, && Y_T = \tilde{g}(X_T), \end{aligned} \tag{13.8}$$

where, $b, \tilde{f}, \tilde{g} \in \mathbb{R}^n$, W_t is a d-dimensional Brownian motion and $\sigma, Z_t \in \mathbb{R}^{n\times d}$ are matrices. Note that σ is independent of Z_t here. The consequences of this simplification in view of our optimal control problem are pointed out in the next paragraph.

The following theorem holds when we apply our CFB algorithm (12.36) - (12.41) in Section 12.6 to the pure FBSDE system (13.8) with the above discretization (13.1) - (13.6). In the pure FBSDE case, the optimization step (12.38) - (12.39) of the CFB algorithm can be neglected.

Theorem 13.1 (Convergence FBSDE)
Assume $b, \sigma, \tilde{f}$ and $\tilde{g}$ are bounded w.r.t x, have at most linear growth w.r.t. y, z, and are uniformly Lipschitz continuous w.r.t x, y, z. Also assume that $\sigma\sigma^\top$ is positive definite and $\tilde{g}$ is bounded in $C^{2+\alpha}(\mathbb{R}^n)$ for $\alpha > 0$. Then, for $p \geq 2$, there exist constants c, C_1 and C_2, only depending on p, T and other known parameters, such that, for $h < c$, $\delta^2 < h$, $L^{-2/d} < h$ and $\rho \geq 1$:

$$\sup_{\xi \in C_0^h} \left| Y_{t_0} - Y_0^h \right|^2 \leq C_1\, \mathcal{E}^2(\mathit{global}), \tag{13.9}$$

and, where τ is the first exit time of X_t^h leaving the radius $R+\rho$ of the spatial grid C_τ^h:

$$\begin{aligned} E\left[\sup_{k\in\{1,\dots,N\}}\left|X_{t_k}-X_k^h\right|^2\mathbf{1}_{\{t_k=\tau\}}\right]+E\left[\sup_{k\in\{1,\dots,N\}}\left|Y_{t_k}-Y_k^h\right|^2\mathbf{1}_{\{t_k=\tau\}}\right] \\ +h\sum_{k=0}^{N-1}E\left[\left|Z_{t_k}-Z_k^h\right|^2\mathbf{1}_{\{t_k<\tau\}}\right]\le C_2\,\mathcal{E}^2(global), \end{aligned} \tag{13.10}$$

where $\mathcal{E}^2(global)$

$$\begin{array}{llllll} = & \mathcal{E}^2(time) & +\mathcal{E}^2(space) & +\mathcal{E}^2(trunc) & +\mathcal{E}^2(quantiz) & +\mathcal{E}^2(gradient) \\ = & h & +h^{-2}\delta^2 & +\frac{R^2}{(R+\rho)^2} & +h^{-1}L^{-\frac{2}{d}} & +h^{p+\frac{d}{2}-1}L^{-\frac{2p}{d}}\delta^{-2p-d}. \end{array}$$

The proof can be found in [Delarue, Menozzi 2005] theorem 4.1, theorem 4.2 and theorem 4.3. for the case that $\tilde{f},\tilde{g},Y_t\in\mathbb{R}^1$, $Z\in\mathbb{R}^{1\times d}$ and $d=n$.

13.2.1 Multi-dimensional case

In order to transfer the results to the multi-dimensional case $n>1$, note that the fundamental theorem used in their convergence analysis is the Hölder regularity for parabolic PDEs, theorem 2.1. It states that for $u(t,x)=Y_t^{t,x}$:

$$\begin{aligned} |u(t,x)|+|\nabla_x u(t,x)|+|\nabla_{xx}^2 u(t,x)|+|\partial_t u(t,x)| \\ +\sup_{t'\in[0,T],t\neq t'}\left[|t-t'|^{-1/2}\,|\nabla_x u(t,x)-\nabla_x u(t',x)|\right]\le C, \end{aligned} \tag{13.11}$$

where $C>0$ is a constant. Since this Hölder regularity and Gronwall's lemma also holds in n-dimensions, [Delarue, Menozzi 2005]'s convergence analysis can be transfered to the case when $Y\in\mathbb{R}^n$.

Moreover a change from $d=n$ to $d\le n$ affect only their lemma 6.3. and imply no difficulties. In contrast, the Quantization error $\mathcal{E}^2$(quantiz) reduces to $h^{-1}L^{-\frac{2}{d}}\le h^{-1}L^{-\frac{2}{n}}$ for $d\le n$.

13.2.2 Choice of discretization parameter

The referenced paper provides more details about the nature of the different errors $\mathcal{E}^2$(time), $\mathcal{E}^2$(space), $\mathcal{E}^2$(trunc), $\mathcal{E}^2$(quantiz) and $\mathcal{E}^2$(gradient). We do not repeat them here. Instead, let us set:

$$\rho=R\,h^{-\frac{1}{2}},\qquad \delta=h^{\frac{3}{2}},\qquad L^{-\frac{2}{d}}=h^{2+\frac{(d+2)}{p}}. \tag{13.12}$$

Then we obtain:

$$\mathcal{E}^2(global) = \; h + h + h\frac{1}{(1+\sqrt{h})^2} + h\,h^{\frac{(d+2)}{p}} + h \;\leq\; 5h. \tag{13.13}$$

Therefore, we can provide the following corollary.

Corollary 13.2 (Rate of convergence)
Using the parameterization (13.12), the rate of convergence for $X_k^h \to X_{t_k}$ and $Y_k^h \to Y_{t_k}$ is h, in the sense of (13.9) and (13.10) inside the boundaries of the spartial grid C_k^h.

13.3 Convergence of an uncontrolled-diffusion case

Now, let us transfer the convergence results from the previous pure FBSDE case to the stochastic optimal control case when the diffusion is not controlled, i.e., σ is independent from π. Therefore, let us consider our Concave Problem 6 (Section 10.4) and its Hamilton function:

$$\begin{aligned} H(t,x,\pi,y,z) &= b(t,x,\pi)y + \mathrm{tr}\big[\sigma(t,x)^\top z\big] + f(t,x,\pi), \\ \nabla_\pi H(t,x,\pi,y) &= \nabla_\pi\left(b(t,x,\pi)y\right) + \nabla_\pi f(t,x,\pi), \\ \nabla^2_{\pi\pi} H(t,x,\pi,y) &= \nabla^2_{\pi\pi}\left(b(t,x,\pi)y\right) + \nabla^2_{\pi\pi} f(t,x,\pi). \end{aligned}$$

It is clear that $\nabla_\pi H(t,x,\pi,y)$ and $\nabla^2_{\pi\pi} H(t,x,\pi,y)$ are independent of z when $\sigma(t,x)$ is independent of π. In consequence, the optimization step (12.38) - (12.39) of the CFB algorithm does not involve the adjoint variable Z_t and we can define the optimal control without loss of generality at $z = 0$ by:

$$\pi^*(t,x,y) \coloneqq \arg\max_{\pi\in A} H(t,x,\pi,y,0). \tag{13.14}$$

As a result, by using our CFB-algorithm from Section 12.6 with an exact optimization and the discretization from Section 13.1, the following theorem holds.

Theorem 13.3 (Convergence uncontrolled diffusion)
Consider the Concave Problem 6 (Section 10.4) and let $p \geq 2$. Additionally assume that:

1. *b, σ, f do not explicitly depend on time,*
2. *b is linear w.r.t (x,π), i.e., b is of the form $b(x,\pi) = Ax + B\pi + C$,*

3. *σ is independent of π and linear w.r.t. x,*

4. *$\nabla_x f$ has at most linear growth w.r.t. π,*

5. *π^* depends linearly on y.*

Then, the convergence results of Theorem 13.1 hold for X_t, Y_t and Z_t. In particular, if we choose the parameters δ, ρ, M according to (13.12), then the rate of convergence for $X_t^h \to \hat{X}_t$ and $\pi_t^h \to \pi_t^*$ is h inside the spatial grid C_t^h.

Proof 13.4
First, we note that the problem admits a unique solution due to Theorem 10.4 as a special linear case. Second, the problem's parameters have the following properties:

- *g: due to the Basic Conditions 10.1, g has at most quadratic growth and is continuous differentiable. Thus $\nabla_x g$ has at most linear growth w.r.t. π and is bounded in $C^2(\mathbb{R}^n)$.*

- *σ: due to the SMP Conditions 10.9, $\sigma\sigma^\top$ is positive definite. By assumption, σ is independent of π and linear w.r.t. x.*

- *b: By assumption, b is linear w.r.t (x, π). Together with the assumption that $\pi^*(t, x, y)$ has at most linear growth w.r.t. y, we get that b is linear w.r.t. y.*

- *$\nabla_x H$: since b is linear w.r.t. (x, π), b_x is independent of π. By assumption, σ_x is independent of π and f_x has at most linear growth w.r.t. π. Together with the assumption that $\pi^*(t, x, y)$ has at most linear growth w.r.t. y, we get that $\nabla_x H$ has at most linear growth w.r.t. y and z.*

- *$\nabla_x H$: due to Conditions 10.1, f has at most quadratic growth w.r.t. x and it is continuous differentiable. Thus $\nabla_x H$ has at most linear growth w.r.t. x.*

Using these properties, we can apply Theorem 13.1 to the FBSDE system of the control problem (12.1), which is coupled through an optimality condition. Moreover, since $\pi^*(t, x, y)$ has at most linear growth w.r.t. y and we assume an exact inner optimization, we can transfer the convergence analysis from X_t^h and Y_t^h to π_t^h. qed.

Comments

When σ does not depend on π, $\nabla_\pi H$ does not depend on z and therefore the gradient error $\mathcal{E}^2(\text{gradient})$ in (13.10) disappers.

We already defined growth and global Lipschitz conditions in our Concave Problem 6 through Conditions 10.1 and 10.9. Considering these conditions together with the compactness and the convexity of the control set A, the additional linearity assumptions 2 and 3 in Theorem 13.3 are neglectable. In particular, when σ is independent of π, the coefficients b and σ do not depend on the adjoint variable Z_t. Moreover, [Delarue, Menozzi 2005]'s lemma 6.3 (and all connected statements) rely on the boundedness of the drift b w.r.t y. Since $\pi \in A$ and A is compact, b is already bounded in y.

Considering assumption 1 in Theorem 13.3, [Delarue, Menozzi 2005] stated (but did not prove) that their convergence analysis should hold for time dependent b, σ and f, too. Therefore, we make the following statement without a rigorous proof.

Statement 13.5 (Convergence uncontrolled diffusion 2) *Consider our Concave Problem 6 (Section 10.4) and let $p \geq 2$. Additionally assume that:*

1. *σ is independent of π,*
2. *π^* depends linearly on y.,*
3. *$\nabla_x f$ has at most linear growth w.r.t. π.*

Then the results of Theorem 13.3 hold. In particular, our CFB algorithm (12.36) - (12.41) in Section 12.6 obtain linear convergence when $h \to 0$ for $X_t^h \to X_t$ and $\pi_t^h \to \pi_t^*$ inside the spatial grid C_t^h.

Note that we did not analyze the effect of the optimization error $|\pi_{t_k}^* - \pi_k^h|$ to the overall convergence. Since b and $\nabla_x f$ have at most linear growth w.r.t. π, the above convergence result should hold until $|\pi_{t_k}^* - \pi_k^h| < ch$, where the constant $0 \leq c \leq 1$ is small enough.

13.4 Convergence for the general case

Little is known for the case when the diffusion is controlled, i.e., $\sigma(x, \pi)$ depends on π. We remark that a general proof for the existence and convergence in the controlled diffusion case would be analogous to proving the

existence, the uniqueness of a fully nonlinear PDE and the according convergence of the algorithm. Such a proof is far beyond the scope of this thesis.

Nevertheless, our CFB algorithm can be applied to controlled diffusion problems without any changes. Doing so, the CFB algorithm provides a strategy to calculate a best-known *optimized control* π^h that is likely to be close to an optimal control, if it exists.

13.5 Summary

- For a subclass of linear-concave stochastic optimal control problems with uncontrolled diffusion, the convergence of our proposed CFB algorithm (12.36) - (12.41) in Section 12.6 can be rigorously proved through Theorem 13.3.
- For the subclass of Concave Problems 6 (Section 10.4) with uncontrolled diffusion, the convergence analysis holds through Statement 13.5, too.
- Little is known about the convergence of the CFB algorithm for general Concave Problems with controlled diffusion but technically, it could be applied without any changes.
- In the convergence analysis, five different type of errors can be distinguished; the error due to:
  - time discretization $\mathcal{E}^2(\text{time})$,
  - space discretization $\mathcal{E}^2(\text{space})$,
  - truncation of the state space $\mathcal{E}^2(\text{trunc})$,
  - discrete quantization $\mathcal{E}^2(\text{quantiz})$,
  - gradient approximation $\mathcal{E}^2(\text{gradient})$.
- Using an appropriate choice of the discretization parameters as in (13.12), the convergence rate of our CFB algorithm to the optimal solution $(\hat{X}_t, \pi_t^*)$ (in the sense of (13.9) and (13.10)) is linear in h inside the boundaries of the spartial grid C^h when $h \to 0$.

13.5.1 Possible improvements

In order to improve the convergence we propose to implement an inner loop that calculates π_k^j, V_k^j, Y_k^j and Z_k^j in (12.37) - (12.41) multiple times for all $\xi^j \in \mathrm{C}_k^h$ in an iterative way, before continuing with the next time step $k-1$.

14 Comparison and advantages between the CFB scheme and the DP scheme

In this section we compare the complexity of our controlled forward-backward (CFB) scheme with the dynamic programming (DP) scheme through an operation count. We choose the DP scheme for comparison because it is a stochastic approach like the CFB scheme and therefore more comparable then a discrete approach, see Section 11.1.1. Moreover, using the right choice of discretization parameters (e.g. Δt and Δx), both schemes have the same convergence rate. We refer to [Pages, Pham, Printems 2004] for a general DP scheme using Quantization.

It may be clear how to compare the performance of different techniques of the same numerical scheme, but there is no standard way to compare the general complexity of different numerical schemes. We choose the operation count because it: 1) does not depend on a specific problem and parameters, 2) it is independent of the used hardware (CPU performance, number of cores, cash memory) and 3) it is independent of the used software library and coding technique. In contrast, an operation count is an adequate measure for complexity because it reveals the order, with which each discretization parameter contributes to the numerical effort (computational cost).

On one hand, we obviously cannot state that our CFB algorithm (SMP approach) superiors a DP algorithm (DPP approach) in general. On the other hand, we can definitely state that there is a structural difference in the number of operations between both schemes, according to the theoretical considerations in this section. In consequence, for certain problem classes, the SMP approach performs better than the DPP approach. A practical example of this performance improvement is presented in the next Section 15.

14.1 Operation count

We briefly recall the parameter settings. Let n denote the number of state dimensions, $X_t \in \mathbb{R}^n$, let d denote the number of uncertainties, $W_t \in \mathbb{R}^d$ and let r denote the dimension of the control space $\pi_t \in \mathbb{R}^r$. Furthermore, let N denote the number of time steps and let M denote the fixed number of discretization points of the spatial grid C_k^h for each dimension, $|I_k^n| = M^n$ for all $k = 1, ..., N$. Let L denote the fixed number of discretization points for the

Quantization grid Λ_k^h for each dimension, $|L_k^d| = L^d$ for all $k = 1, ..., N$.[53] Last but not least let K denote the average number of steps in the line-search that are necessary to reach the accuracy level ϵ^h for the optimal control values.

Since both schemes are stochastic approaches, we suppose the same Quantization method (Section 12.5.1) with the same number of Quantization points L in order to calculate the appearing expectation values. Moreover, we suppose the same optimization method (Section 12.3.1) with the same average number of iterations K for the inner optimization step in both schemes. Note that advanced techniques like iterative grids or advanced Monte Carlo methods can be applied to both schemes in a similar way. Thus, these advanced techniques do not affect our structural comparison of complexity.

14.1.1 Operation count for the DP scheme

Now we calculate the number of operations for a dynamic programming algorithm, where the Quantization method of Section 12.5.1 and the Newton line-search method of Section 12.3.1 are used. The discrete backward iteration of the DP scheme is given by:

For all $k = N - 1, ..., 0$,

for all $j = 1, ..., M^n$,

$$V_k^j = \max_{\pi \in A} \left(f(t_k, \xi^j, \pi) \Delta t_k^h + \sum_{l \in L^d} p^l \, V_{k+1}^h \left(\xi_{k+1}^{k,j}(\pi \,|\, y^l) \right) \right), \tag{14.1}$$

where:

$$\xi_{k+1}^{k,j}(\pi \,|\, y^l) = \Pi_{k+1} \left[\xi^j + b(\xi^j, \pi) + \sigma(\xi^j, \pi) \, y^l \right] \in \mathrm{C}_{k+1}, \quad \forall l \in \Lambda_k^h, \tag{14.2}$$

Equation (14.1) imposes (on average) K line-search steps in order to find an approximation of the optimal control π^*. One line-search is executed for all M^n grid points at each of the N time steps. Thus, the DP algorithm executes $N \cdot M^n \cdot K$ line-search steps to solve the optimization (14.1).

For each line search step, the exact Newton method builds up a system (12.32), where $H = f(t_k, \xi^j, \pi) \Delta t_k^h + \sum_{l \in L^d} p^l \, V_{k+1}^h \left(\xi_{k+1}^{k,j}(\pi \,|\, y^l) \right)$ is the objective function (see (14.1)). The gradient $\partial_\pi H$ in (12.32) has r entries and the Hessian $\partial_{\pi\pi}^2 H$ has r^2 entries. When each entry is approximated through

[53] There is no structural difference between CFB and DP through the choice of a Quantization method, a Markov-chain approximation or a Monte Carlo method.

finite differences of H, the DP algorithm evaluates H $2(r^2+r)$ times.

One evaluation of H imposes L^d simulations of $\xi_{k+1}^{k,j}(\pi\,|\,y^l)$ (see (14.2)) to evaluate the expectation $E[V_{k+1}^h]$ (the second term $\sum_{l\in L^d} p^l\, V_{k+1}^h$ on the right hand side of equation (14.1)). The number of operations to evaluate the first term $f\Delta t_k^h$ are neglectable. So one evaluation of H imposes cL^d $(c>1)$ operations. In summary, the DP algorithm imposes:

$$\#operations_{DP} \;=\; c\cdot N\cdot M^n\cdot L^d\cdot K\cdot r^2, \quad c>1. \tag{14.3}$$

Note that we do not consider the costs for the solution of the optimization i.e., the linear system (12.32). These costs should be of similar order for the DP and the CFB scheme. When a direct optimization (see Section 11.1.1) is used, the control space is discretized into $\tilde{K}^r$ points. For the operation count, we only need to replace $K\cdot r^2$ in (14.3) by $\tilde{K}^r$ then.

Now let us consider the choice of the discretization parameters N, M and L in order to obtain a linear convergence rate. In [Pages, Pham, Printems 2004], equation 5.3 shows the convergence rate of the DP scheme using quantization. When we set $M\sim N^2$ and $L\sim N^2$ (this is $\delta\sim 1/n^2$ and $N\sim n^2$ in the paper's notations) then the scheme converges linearly in $h=\Delta t=T/N$. In view of the operation count (14.3) we obtain:

$$\#operations_{DP}^{lin} \;=\; c\cdot N^{2n+1}\cdot L^d\cdot \tilde{K}^r. \tag{14.4}$$

14.1.2 Operation count for the CFB scheme

Now let us consider our controlled forward-backward algorithm of Section 12.6. As for the DPP approach, the backward iteration calculates expectation values $E[\cdot]$ (17.6) for all M^n grid points at each of the N time steps. There are $(1+n+nd)$ expectation values (for V, Y, Z) in (17.6) and each expectation imposes cL^d operations to simulate $\xi_{k+1}^{k,j}(\pi\,|\,y^l)$ (14.2). Thus, the CFB algorithm executes $c\cdot N\cdot M^n\cdot(1+n+nd)L^d$ operations so far.

The CFB algorithm also performs K line-search steps for all M^n grid points at each of the N time steps, in order to approximate the optimal control π^*. Again, each Newton line-search step builds up a linear system (12.38) with (r^2+r) entries. Each entry is calculated via (12.35) and thus imposes operations in the order of $n{\cdot}d$. So the optimization needs $c{\cdot}N{\cdot}M^n{\cdot}K{\cdot}(r^2{+}r){\cdot}nd$ operations. Adding both numbers, the CFB algorithm imposes:

$$\#operations_{CFB} \;=\; c\cdot N\cdot M^n\cdot nd(L^d+Kr^2), \quad c>1. \tag{14.5}$$

When using a direct optimization, we replace $K \cdot r^2$ in (14.5) by the number of points $\tilde{K}^r$ in the control space $A \subset \mathbb{R}^r$.

Now let us consider the discretization parameters N, M and L according to (13.12). This means that $M \sim N^2$ and $L \sim N$. By using these parameters, equation (13.13) shows that the CFB scheme converges linearly in $h = \Delta t = T/N$. In view of the operation count (14.5) we obtain:

$$\#operations^{lin}_{CFB} \;=\; c \cdot N^{2n+1} \cdot (L^d + \tilde{K}^r). \tag{14.6}$$

14.1.3 Comparison

By comparing (14.4) and (14.6) we see the difference between the DP and the CFB schemes in terms of complexity. The separate optimization of the Hamiltonian H (12.28) in the SMP scheme leads to an addition of $\tilde{K}^r$ for every grid point. In contrast, the optimization (14.1) of the value function's expectation $E[V]$ in the DP scheme leads to a multiplication with $\tilde{K}^r$ for every grid point.

In return, the CFB scheme needs to calculate $(1 + n + nd)$ expectation values for V, Y and Z. However, n and d are relatively small in comparison to the number of discretization points $\tilde{K}^r$. In summary, our comparison of complexity between the DP and the SMP schemes illustrates the advantages of the latter when $h \to 0$ ($N \to \infty$) if:

$$L^d + \tilde{K}^r \;<\; L^d \cdot \tilde{K}^r. \tag{14.7}$$

It is clear that the CFB scheme has significantly lower computational costs if a nontrivial number of optimization steps $\tilde{K}$ are required. Once again, the advantage of the CFB algorithm is that we do not need to optimize over the value function's expectation $E[V]$, which requires the recalculation of the expectations for each optimization step K. Instead, one calculates the expectations for V, Y and Z once and then optimize the Hamiltonian, which is a much simpler procedure. In particular, the optimization does not include the calculation of expectations.

14.2 Additional advantages

In many applications, the SMP approach has additional advantages compared to the DPP approach:

- The DPP approach deals with the value function V while the SMP approach deals with the adjoint functions Y and Z, the first and an

adjusted second derivative of V, respectively. Remember, V is not needed at all for the calculation of the optimal control π^* in the CFB algorithm.

If V is smooth but non-linear w.r.t x, the derivatives Y and Z may be simpler (flatter) functions w.r.t x than V itself. It is clear that the projection mapping Π_{k+1} in (12.40) - which projects the simulated $\xi_{k+1} \notin \mathrm{C}_{k+1}^h$ onto the grid C_{k+1}^h, also known as nearest neighbor method - yields better results for flatter functions. This is also true when an interpolation method is used instead of the projection mapping.

- Moreover, if the functions Y and Z are flatter (closer to linear) w.r.t x, then a (multi-dimensional) linear extrapolation for points outside the truncated grid C_{k+1}^h yields better approximations and the truncation error is smaller.

We do not show these advantages theoretically since they depend on the specific parameters of the stochastic optimal control problem. However, we will see the advantage of the linear inter- and extrapolation in our case studies on portfolio optimization in Part IV. In portfolio optimization and other applications in finance, the value function V is often increasing and concave w.r.t. x (wealth) due to the risk aversion of the agent. Thus, the derivatives of V are flatter (almost linear) functions.

15 Application and performance evaluation of the CFB-algorithm to Swing option valuation under market feedback

In this section, we compare the performance between our CFB scheme and the DP scheme for a problem from finance. In the previous Section 14, we showed through an operation count, that there is a structural difference in the complexity between both schemes. Now we support these theoretical analyses with results from application. We choose an easy example from finance such that the reader can keep the focus on the numerical scheme instead of dealing with issues of the problem. In particular, we consider the valuation of swing options with penalties. These options are standard instruments in commodity and energy markets, see Section 1.3.2 in the preliminaries.

The outline is as follows. In Section 15.1, the case of a non-controlled state process (small investor without market feedback) is presented where the DP algorithm performs quite well. Then, in Section 15.3, the case of a controlled process (large investor with market feedback) is presented where the CFB algorithm yields a huge performance improvement. This section also gives a first insight on how to apply the CFB algorithm in practice and how to deal with implementation issues.

15.1 Case 1: small investor

A Swing option gives the holder the opportunity to buy an underlying asset at time t with price S_t for a fixed price (strike) K over the whole exercise horizon $[0, T]$. The buyer is free to choose the amount of assets (volume) π_t to buy at each time point t. Nevertheless, the option has a local target volume a for each time t and a global target volume A (for the total amount of assets requested over $[0, T]$). If the requested volume differs from the targets, local and global penalties must be paid, respectively. For more details on Swing options see Section 1.3.2 in the preliminaries.

Now, let $[0, T]$ denote the exercise horizon of the Swing option, $S_t \in \mathbb{R}^+$ the underlying asset, $Q_t \in \mathbb{R}^+$ the total requested volume until time t and $\pi_t \geq 0$ the current requested rate. Let us assume that the underlying's price follows a geometric Brownian motion. In the case of a small investor, we assume that the request of assets π_t has no influence on the price changes dS_t.

Then, the states dynamics are given by:

$$\begin{aligned} dS_t &= \mu S_t dt + \sigma S_t dW_t, \qquad & S_0 &= x, \\ dQ_t &= \pi_t dt, & Q_0 &= 0, \end{aligned} \tag{15.1}$$

where $\mu, \sigma > 0$. Now, let $p(\pi), P(q) : \mathbb{R}_+ \to \mathbb{R}_+$ denote the local and global penalties, respectively. We assume that $p(\pi)$ and $P(q)$ are convex functions. Let Π denote the set of all admissible control policies:

$$\Pi = \{\{\pi_t, t \in [0,T]\} \,|\, \pi_t \text{ is feasible, } \pi_t > 0\}. \tag{15.2}$$

Then, for a fixed control policy $\pi \in \Pi$, the option's future payoff starting at time $t \in [0,T]$ with $S_t = x$ and $Q_t = q$ is given by:

$$J^\pi(t,x,q) = E_t^\pi \left[\int_t^T e^{-rs}[\pi_s(S_s - K) - p(\pi_s)]ds - e^{-rT}P(Q_T) \,\middle|\, S_t = x, Q_t = q \right], \tag{15.3}$$

where r denotes the "risk less" interest rate in the market and K denotes the strike. The price of the Swing option at time $t \in [0,T]$ with $S_t = x$ and $Q_t = q$ is given by:

$$v(t,x,q) = \sup_{\pi \in \Pi} J^\pi(t,x,q). \tag{15.4}$$

Now we define the local and global penalty functions of the swing option. Let $a(t)$ denote the local target volume and $A := \int_0^T a(t)dt$ denote the global target volume. Then the penalty functions are defined by:

$$p(\pi) = c(t)\ (\pi - a(t))^2, \quad P(q) = C\ (q - A)^2. \tag{15.5}$$

where $c(t), C > 0$. Note that the problem has no state constraints.

15.1.1 Applying the SMP

The problem satisfies the Basic Conditions 10.1 and the SMP Conditions 10.9. The only exception is the fact, that the terminal function $g(x,q) = -e^{-rT}C(q-A)^2$ is not global Lipschitz continuous w.r.t q but quadratic. However, the optimal control π_t^* stays bounded over the finite time $[0,T]$ because the gain functions $f, g \to -\infty$ if $\pi^* \to \infty$. In consequence, $Q_T = \int_0^T \pi_t^* dt$ is bounded and therefore $\partial_q g(X_T, Q_T)$ is bounded. Thus the Lipschitz continuity w.r.t. q is satisfied for the problem.

The Hamilton function is given by:

$$H(t,x,q,\pi,y^x,y^q,z^x) = \mu x y^x + \pi y^q + \sigma x z^x + e^{-rt}\left[\pi(x-K) - c(t)\ (\pi - a(t))^2\right], \tag{15.6}$$

and its derivatives are:

$$\begin{aligned}
\partial_x H &= \mu y^x + \sigma z^x + e^{-rt}\pi, \\
\partial_q H &= 0, \\
\partial_\pi H &= y^q + e^{-rt}\left[(x-K) - 2c(t)\,(\pi - a(t))\right], \\
\partial_{\pi\pi} H &= -2e^{-rt}\,c(t).
\end{aligned} \tag{15.7}$$

We see that H is concave w.r.t. π, linear w.r.t. x and constant in q, and the terminal function is concave. Thus the SMP Conditions 10.9 hold and we can apply the SMP, see Section 11.3. We obtain the following FBSDE system:

$$\begin{aligned}
dY_t^x &= \left(\mu Y_t^x + \sigma Z_t^x + e^{-rt}\pi_t^*\right)dt + Z_t^x dW_t, \qquad && Y_T^x = 0, \\
dY_t^q &= Z_t^q dW_t, && Y_T^q = 2\,e^{-rT}\,C\,(Q_T - A)\,.
\end{aligned} \tag{15.8}$$

Since $\partial_{\pi\pi} H < 0$, we obtain the optimal control for the Swing option by setting $\partial_\pi H = 0$:

$$\pi_t^* = \frac{e^{rt}Y_t^q + (S_t - K)}{2c(t)} + a(t). \tag{15.9}$$

The forward equations (15.1) and the backward equations (15.8) build a FBSDE system that is coupled through the optimality condition (15.9).

15.1.2 Applying the CFB algorithm

Now let us apply the CFB algorithm of Section 12.6 to system (15.1), (15.8) and (15.9). We define the approximation parameter h, discretize the time and state space according to Section 13.1 and use the discretization parameters according to (13.12). The optimal control π_t^* in (15.9) is linear w.r.t. S_t and Y_t. Thus the conditions of Theorem 13.3 are satisfied and the CFB algorithm converges linearly w.r.t. to the approximation parameter h. The CFB algorithm reads:

For all time steps $k = N, ..., 1$,

for all grid points $j = 1, ..., M^2$,

$$
\begin{aligned}
\pi_k^j &= \tfrac{e^{rk\Delta t}Y_{k+1}^{q,j}+(S_k^j-K)}{2c_k} + a_k,\\
S_{k+1}^{k,j} &= \mu S_k^j \Delta t + \sigma S_k^j \Delta W_t,\\
Q_{k+1}^{k,j} &= Q_k^j + \pi_k^j,\\
Z_k^x &= \tfrac{1}{\Delta t} E[Y_{k+1}^{x,j}\Delta W_k \,|\, S_k^j, Q_k^j],\\
Y_k^{x,j} &= \left(\mu Y_{k+1}^{x,j} + \sigma Z_k^{x,j} + e^{-rt}\pi_t^j\right)\Delta t + E[Y_{k+1}^x \,|\, S_k^j, Q_k^j],\\
Y_k^{q,j} &= E[Y_{k+1}^q \,|\, S_k^j, Q_k^j],\\
V_k^j &= e^{-rk\Delta t}\left(\pi_k^j(S_k^j - K) - c_k\left(\pi_k^j - a_k\right)^2\right)\Delta t + E[V_{k+1} \,|\, S_k^j, Q_k^j].
\end{aligned}
\tag{15.10}
$$

15.1.3 Implementation issues CFB

- Instead of using the projection mapping $\Pi(\cdot)$ in (13.4) (nearest neighbor method), we implement a linear inter- and extrapolation of the functions Y_{k+1}^x, Y_{k+1}^q and V_{k+1} to calculate their values at the simulated points $S_{k+1}^{k,j}$.

- The time step size is $\Delta t = h$. We choose a larger state grid step size $\Delta S = \Delta Q = h^{1/2}$ than in (13.12) ($h^{3/2}$) because of the linear inter- and extrapolation.

- We use the constant Quantization ΔW and the grid boundary extension ρ according to (13.12). The space and Quantization grids are equal for every time point $k = N, ..., 1$.

- We implement a second *inner iteration*, which we introduced in Section 13.5.1 in order to improve the convergence.

- The algorithm is executed on an eighth core machine. The parallelization is done by sharing the space grid between the cores.

- The values of Y_{k+1}^x and Z_{k+1}^x in (15.10) are not necessary to obtain the optimal control π_k^j. Therefore, the calculation of Y_{k+1}^x and Z_{k+1}^x could be omitted in the algorithm.

- The test problem's parameters are: $T = 2$, $r = 0.01$, $\mu = 0.05$, $\sigma = 0.2$, $a = 5$, $K = 3$, $c = 0.1$, $C = 0.2$.

15.1.4 Performance evaluation CFB

In order to compare the performance of the DP and the CFB algorithm, let us consider the $[4,6] \times [4,6]$ cube in the $S \times Q$ space as the area of interest. First, using a fine grid, we calculate approximative "true solutions" V^* and π^*. Then, we calculate approximations $\tilde{V}$ and $\tilde{\pi}$ for several coarser grids $i = 1, ..., 9$ with approximation parameters h^i and compare the maximum absolute error over $[4,6] \times [4,6]$ for time $t = 0$:

$$\begin{aligned} \epsilon_V &= \max_{(x,q)\in[4,6]\times[4,6]} \left|V_0^*(x,q) - \tilde{V}_0(x,q)\right|, \\ \epsilon_\pi &= \max_{(x,q)\in[4,6]\times[4,6]} |\pi_0^*(x,q) - \tilde{\pi}_0(x,q)|. \end{aligned} \tag{15.11}$$

Table 10 shows the calculation times and the errors of each test run.

i	h^i	time	ϵ_V	ϵ_π
1	0.1000	2.41s	3.1944	0.0364
2	0.0500	10.0s	0.5006	0.0238
3	0.0333	27.0s	0.0662	0.0145
4	0.0250	57.4s	0.0292	0.0093
5	0.0200	103s	0.0192	0.0060
6	0.0167	178s	0.0127	0.0038
7	0.0143	286s	0.0079	0.0022
8	0.0125	442s	0.0039	0.0010
*	0.0111	636s	-	-

Table 10: Calculation time and error ϵ_V, ϵ_π for the CFB algorithm on the cube $(S,Q) \in [4,6]$ at time $t = 0$ for the approximation parameter h^i, $i = 1, ..., 9$.

15.1.5 Dynamic programming

We presented the dynamic programming principle in Section 11.1 and discussed the numerical treatment in Section 11.1.1. Our implementaion of the DP algorithm is based on [Pages, Pham, Printems 2004], which we discussed in Section 14.1.1. Therefore, let us discretize an appropriate interval in the control space $\mathbb{R}_+$, where π^o, $o = 1, ..., O$ denote the grid points. Using

the same notation as in the previous paragraphs, we derive the DP algorithm:

For all time steps $k = N, ..., 1$,

for all space grid points $j = 1, ..., M^2$,

for all control grid points $o = 1, ..., O$,

$$\begin{aligned} S_{k+1}^{k,j} &= \mu S_k^j \Delta t + \sigma S_k^j \Delta W_t, \\ Q_{k+1}^{o,j} &= Q_k^j + \pi_k^o, \\ V_k^{o,j} &= e^{-rk\Delta t}\left(\pi^o(S_k^j - K) - c_k\,(\pi^o - a_k)^2\right)\Delta t + E[V_{k+1} \,|\, S_k^j,\, Q_k^j,\, \pi_k^o]. \end{aligned} \tag{15.12}$$

Then we set:

$$V_k^j = \max_o V_k^{o,j}. \tag{15.13}$$

15.1.6 Implementation issues DP

- [Pages, Pham, Printems 2004] used a projection mapping (nearest neighbor method) similar to the projection mapping $\Pi(\cdot)$ in (13.4). Here, we implement a linear inter- and extrapolation of the function V_{k+1} to calculate its values at the simulated points $S_{k+1}^{k,j}$, instead.

- The time step size is $\Delta t = h$. We choose a larger space grid step size $\Delta S = \Delta Q = h^{1/2}$ than [Pages, Pham, Printems 2004] recommended ($h^{3/2}$ see Section 14.1.1) because of the linear inter-, extrapolation.

- We use the constant Quantization ΔW and the grid boundary extension ρ according to (13.12). The space and Quantization grids are equal for every time point $k = N, ..., 1$.

- The algorithm is executed on an eighth core machine. The parallelization is done by sharing the space grid between the cores.

- We use the same parameters as for the CFB algorithm: $T = 2$, $r = 0.01$, $\mu = 0.05$, $\sigma = 0.2$, $a = 5$, $K = 3$, $c = 0.1$, $C = 0.2$.

- After a few test calculations we obtained that the optimal control $\pi_t^\star$ does not exceed the value $\pi_{max} := S_{max} * 6$ where $S_{max} := 6 + \rho$ is the maximal point of the spaced grid. Therefore we set the control interval to $[0, \pi_{max}]$.

- Since we choose to discretize the control space in a direct approach (see Section 11.1.1) in order to solve the optimization, we set the control grid step to $\Delta O = 2\Delta S$.

Iteration improvement for the DP

We implement the following iteration improvement for the DP algorithm (15.12). Let us assume that the backward algorithm is at time step t_k and grid point S^j. First we pre-calculate the expectation values $E_k^j[V^l] := E_k\left[V_{k+1} \,|\, S_k^j, Q_k^l, \pi = 0\right]$ for all grid points Q^l, $l = 1, ..., M_Q$. When we proceed with the above DP algorithm on every grid point Q^l $l = 1, ..., M_Q$ and every control grid point π_k^o, $o = 1, ..., O$, we can use the pre-calculated values $E_k^j[V^l]$ to interpolate $E_k\left[V_{k+1} \,|\, S_k^j, Q_k^l, \pi_k^o\right]$ in (15.12). This improvement is possible because only the ODE for the volume Q in (15.1) is controlled (it depends on π) but the SDE for the price process S_t is not controlled.

15.1.7 Performance evaluation DP

Again, let us consider the $[4, 6] \times [4, 6]$ cube in the $S \times Q$ space as the area of interest. We use the same approximative "true solutions" V^* and π^* as for the CFB algorithm in Section 15.1.4 and calculate approximations $\tilde{V}$ and $\tilde{\pi}$ for several grids $i = 1, ..., 6$ with approximation parameters h^i and compare the maximum absolute error (15.11) over $[4, 6] \times [4, 6]$ for time $t = 0$. Table 11 shows the calculation times and the errors of each test run.

i	h^i	time	ϵ_V	ϵ_π
1	0.1000	2.82s	11.5590	5.4185
2	0.0500	13.7s	1.5602	1.2718
3	0.0333	39.8s	0.2576	0.6084
4	0.0250	87.1s	0.0463	0.3901
5	0.0200	162s	0.0136	0.3363
6	0.0167	280s	0.0070	0.3283

Table 11: Calculation time and error ϵ_V, ϵ_π for the DP algorithm on the cube $(S, Q) \in [4, 6]$ at time $t = 0$ for the approximation parameter h^i, $i = 1, ..., 8$.

15.2 Performance comparison between CFB and DP

Figure 38 shows the error ϵ_V (15.11) in relation to the calculation time for the CFB and the DP algorithm on a logarithmic scale. We see that the performance of both algorithms are quite similar. In particular, their error - calculation-time relations have the same slope when h decreases. The DP algorithm performs slightly better for small h than the CFB algorithm.

Since the stochastic variable S_t does not depend on the control π, the algorithms do no need to repeat the path simulations of the stochastic variable S_t for different control values π^o, $o = 1, ..., O$. Therefore, we can implement an iteration improvement for the DP algorithm which we described in Section 15.1.6. The CFB performance would not be affected by a similar iteration improvement because it does not variate the control π but calculate it directly through (15.9). In consequence, the DP algorithm and the CFB algorithm have similar complexity (operation count), see Section 14.

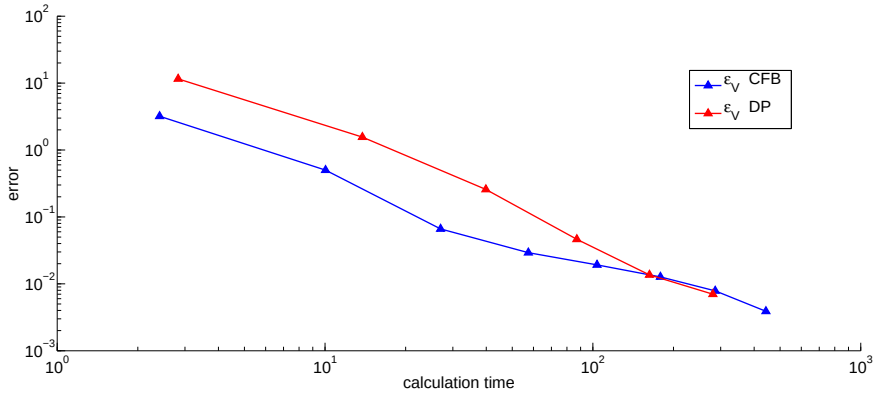

Figure 38: Comparison of the error ϵ_V in relation to the calculation time between the CFB and DP algorithm on the cube $(S, Q) \in [4, 6]$ at time $t = 0$ for the approximation parameter h^i, $i = 1, ..., 8$.

15.3 Case 2: big investor with market feedback

Now we assume that the investor's trading activity is influencing the market, i.e., the investors strategy has an impact on the evolution of the asset prices. This effect is called market feedback and especially observable in commodity markets because the prices of physical goods depend on supply and demand, see our discussion in Section 2 in the preliminaries. If demand is lower or

higher than the expected demand, the market has an over supply or a supply shortage, respectively. We assume that the underlying's price increases linearly when the option's holder requests more volume than the expected volume a. It decreases when the option's holder requests less volume than expected. Therefore, we replace the SDE (15.1) by:

$$\begin{aligned} dS_t &= \big(\mu S_t + \kappa(\pi_t - a)\big)\, dt + \sigma S_t dW_t, \qquad && S_0 = x, \\ dQ_t &= \pi_s ds, && Q_0 = 0. \end{aligned} \tag{15.14}$$

Here, $\kappa > 0$ is a fixed weight which describe the impact power of the agent to the market. We call it market impact or market strength of the trader. In the test runs we set $\kappa = 0.05$. Everything else is equal to case 1 of Section 15.1. The main difference is, that the stochastic variable S_t is controlled now, i.e., it depends on the control π.

15.3.1 Applying the CFB algorithm

The extended problem also satisfies the Basic Conditions 10.1 and the SMP Conditions 10.9. The optimal control π_t^* stays bounded if the penalty coefficient $c(t)$ in (15.5) is high enough compared to the market power coefficient κ in (15.14). Thus the conditions of Theorem 13.3 are satisfied and the CFB algorithm converges linearly w.r.t. to the approximation parameter h.

The Hamilton function of the extended problem is given by:

$$\begin{aligned} H(t, x, q, \pi, y^x, y^q, z_1) = \;& \mu\left(x + \kappa(\pi - a)\right) y^x + \pi y^q + \sigma x z_1 \\ & + e^{-rt}\left[\pi(x - K) - c(t)\,(\pi - a(t))^2\right], \end{aligned} \tag{15.15}$$

and thus we obtain the FBSDE system:

$$\begin{aligned} dY_t^1 &= \big(\mu Y_t^1 + \sigma Z_t^1 + e^{-rt}\pi_t^*\big)\, dt + Z_t^1 dW_t, \qquad && Y_T^1 = 0, \\ dY_t^2 &= Z_t^2 dW_t, && Y_T^2 = 2\, e^{-rT} C\,(Q_T - A)\,, \end{aligned} \tag{15.16}$$

with the optimal control given by:

$$\pi_t^* = \frac{e^{rt}\left(\mu\kappa Y_t^1 + Y_t^2\right) + (S_t - K)}{2c(t)} + a(t). \tag{15.17}$$

The CFB algorithm is equal to (15.10) but replacing π_k^* and S_{k+1} with:

$$\begin{aligned} \pi_k^* &= \frac{e^{rk\Delta t}\left(\mu\kappa Y_{k+1}^1 + Y_{k+1}^2\right) + (S_k - K)}{2c_k} + a_k, \\ S_{k+1} &= \mu\left(S_k + \kappa\left(\pi_k^* - a\right)\right)\Delta t + \sigma S_k \Delta W_t. \end{aligned} \tag{15.18}$$

15.3.2 Performance evaluation CFB

In order to compare the performance of the DP and the CFB algorithm, let us consider the $[4,6] \times [4,6]$ cube in the $S \times Q$ space as the area of interest. First, using a fine grid, we calculate approximative "true solutions" V^* and π^*. Then, we calculate approximations $\tilde{V}$ and $\tilde{\pi}$ for several coarser grids $i = 1, ..., 9$ with approximation parameters h^i and compare the maximum absolute errors ϵ_V and ϵ_π (15.11) over $[4,6] \times [4,6]$ for time $t = 0$. Table 12 shows the calculation times and the errors of each test run.

i	h^i	time	ϵ_V	ϵ_π	single op. time
1	0.1000	5.00s	5.0476	0.0206	9,154E-05
2	0.0500	24.4s	1.0460	0.0060	7,513E-05
3	0.0333	67.5s	0.2300	0.0033	7,082E-05
4	0.0250	145s	0.0463	0.0019	7,012E-05
5	0.0200	266s	0.0234	0.0014	7,163E-05
6	0.0167	457s	0.0162	0.0008	7,396E-05
7	0.0143	722s	0.0108	0.0005	7,526E-05
8	0.0125	1121s	0.0054	0.0002	7,831E-05
*	0.0111	1573s	-	-	8.031E-05

Table 12: Calculation time and error ϵ_V, ϵ_π for the CFB algorithm on the cube $(S, Q) \in [4,6]$ at time $t = 0$ for the approximation parameter h^i, $i = 1, ..., 9$.

The last column shows the calculation time divided by the number of operations (iterations) according to our operation count (14.5) in Section 14.1:

$$\text{single op. time} = \frac{\text{time}}{NM^2L}.$$

This is the time needed for a 'single' operation. Note that $K = 0$ since the optimal control is obtained directly through (15.17). We see that the values are of the same order (around $7.5E-05$) for all test runs $i = 1, ..., 9$. This supports the fact that our theoretical operation count is a good estimate of the algorithm's complexity and thus a performance indicator.

The calculation times for the market feedback case in Table 12 are approximately 2 to 2.5 times higher than the calculation times for the case without market feedback in Table 10. The reason it that the adjoint variables Y_t^x and Z_t^x can be neglected in the latter and only the variables V_t and Y_t^q are considered. In consequence, the number of variables is doubled in the market feedback case.

15.3.3 Dynamic programming

We use the same algorithm as in Section 15.1.5 but replacing (15.12) with:

$$\begin{aligned} S_{k+1}^{k,j} &= \mu\left(S_k^j + \kappa(\pi_k^o - a)\right)\Delta t + \sigma S_k^j \Delta W_t, \\ Q_{k+1}^{o,j} &= Q_k^j + \pi_k^o, \\ V_k^{o,j} &= e^{-rk\Delta t}\left(\pi^o(S_k^j - K) - c_k\,(\pi^o - a_k)^2\right)\Delta t + E[V_{k+1}\,|\,S_k^j,\, Q_k^j,\, pi_k^o]. \end{aligned} \tag{15.19}$$

Note that in the market feedback case, we can do an iteration improvement as in case 1, i.e., pre-calculate the expectations. The reason is that the stochastic variable S_t is controlled, i.e., it depends on the control π.

15.3.4 Performance evaluation DP

Again, let us consider the $[4,6]\times[4,6]$ cube in the $S\times Q$ space and the approximative "true solutions" V^* and π^* of Section 15.3.2. We compare the maximum absolute error (15.11) over $[4,6]\times[4,6]$ for time $t=0$ for several grids $i=1,...,6$. Table 13 shows the calculation times and the errors.

i	h^i	time	ϵ_V	ϵ_π	single op. time
1	0.1000	95s	28.6180	13.4540	1,15951E-05
2	0.0500	1,059s	4.6294	3.2666	1,23517E-05
3	0.0333	4,411s	0.9408	1.1027	1,26842E-05
4	0.0250	12,546s	0.2342	0.5093	1,26842E-05
5	0.0200	27,744s	0.0803	0.3838	1,30044E-05
6	0.0167	56,138s	0.0296	0.3300	1,34571E-05

Table 13: Calculation time and error ϵ_V, ϵ_π for the DP algorithm on the cube $(S,Q)\in[4,6]$ at time $t=0$ for the approximation parameter h^i, $i=1,...,8$.

The last column shows the calculation time divided by the number of operations (iterations) according to our operation count (14.3) in Section 14.1:[54]

$$\text{single op. time} = \frac{\text{time}}{NM^2LO}.$$

Again, all (single op. time) values are of the same order (around $1.27E-05$) for all test runs $i = 1, ..., 6$. This supports the fact that our theoretical operation count is a good estimate of the algorithm's complexity and thus a performance indicator.

The average time of a single operation (iteration) in the DP algorithm ($1.27E-05$) is about 6 times faster than the average time in the CFB algorithm ($7.5E-05$). The reason is that the CFB calculates multiple variables (V, Y, Z) (15.10) in each iteration, while the DP calculates only V (15.12).

15.4 Performance comparison between CFB and DP

Figure 39 shows the error ϵ_V (15.11) in relation to the calculation time for the CFB and the DP algorithm on a logarithmic scale.

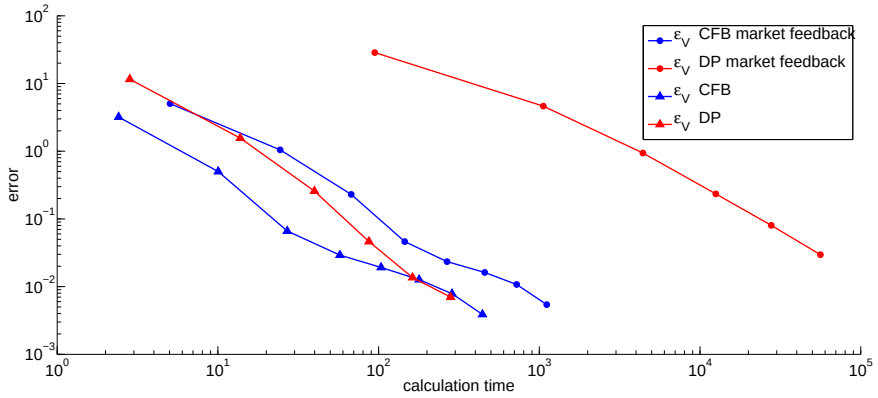

Figure 39: Comparison of the error ϵ_V in relation to the calculation time between the CFB and DP algorithm on the cube $(S, Q) \in [4, 6]$ at time $t = 0$ for the approximation parameter h^i, $i = 1, ..., 8$.

We see that the slope of the CFB algorithm is still similar to the slope of the DP algorithm but the CFB algorithm clearly superiors the DP algorithm in the market feedback case. In particular, to achieve an $\epsilon \approx 0.03$, the CFB

[54] Here, the number of control space points is O instead of K.

algorithm needs about $266s$, while the DP algorithm needs about $56138s$. This is a factor of 210.

15.5 Conclusions

The examples show the structural advantage of the CFB algorithm in terms of complexity (operation count) as we discussed in Section 14.1. In particular, when the stochastic variable is not controlled, the DP and the CFB perform similar. When the stochastic variable is controlled, the CFB algorithm superiors the DP algorithm due to lower complexity (operation count).

We also showed that the operation count is a good measure for performance. We even could measure the minor disadvantage of the CFB algorithm - the handling of multiple variables - in this example.

Part IV
Application of the controlled forward-backward algorithm to optimal portfolio allocation

In this part of the thesis, we present details about how we applied our controlled forward-backward (CFB) algorithm (12.36) - (12.41) to our model of optimal portfolio allocation of commodity related assets, in order to obtain the numerical results of Section 7 and 8. First, in Section 16, we show the applicability of the CFB algorithm to the stochastic optimal control Problem 2 (5.13) - (5.16). Then, in Section 17, we formulate the CFB algorithm for the specific case of Problem 2 and finally, in Section 18, we state some implementation issues. The notation in this part is based on the model's notation of Part II.

16 Applicability of the CFB algorithm

In this section, we show the applicability of the CFB algorithm (12.36) - (12.41) to our continuous time model for optimal portfolio allocation of production assets that is defined through Problem 2 (5.13) - (5.16). By applicability we mean that the model must be a Concave Problem 6 (see Section 10.4) in order to apply our CFB algorithm. In particular, the model must satisfy the Basic Conditions 10.1 and the SMP Conditions 10.9 of Section 5.

In Section 5.3.1 we already showed that the Basic Conditions 10.1 are satisfied. We proved the existence of a strong unique solution of the SDEs and the boundedness of the objective function through Theorem 5.2. In the following, we prove that the SMP Conditions hold too. For notational simplicity, we consider the one asset case with negative exponential utility. The results can be transfered to the multiple asset case without extensions and they also hold for other HARA utility function. Now, let us repeat the reduced continuous-time model for optimal portfolio allocation of production assets, Problem 2 (5.13) - (5.16) for the one asset case.

Problem 7 (Continuous-time portfolio alloc. of production assets) *We search for the $\mathbb{F}_t$-admissible, optimal control policy $(\alpha^*, \chi^*) \in \mathbb{R}^2 \times [0, T]$, that maximizes the objective function:*

$$J(0, a, b, c, \alpha, \chi) := E_0^{\alpha,\chi}\left[\int_0^T e^{-\delta t} \frac{1 - e^{-\gamma \chi_t}}{\gamma}\, dt + e^{-\delta T} \frac{1 - e^{\gamma(A_T + B_T)}}{\gamma} \,\middle|\, a, b, c \right], \tag{16.1}$$

under the state dynamics:

$$\begin{aligned} dA_t &= \left(\mu A_t + \alpha_t\right)\, dt \;+\; \sigma A_t\, dW_t, && A_0 = a \in \mathbb{R}_+, \\ dR_t &= \kappa\left(\bar{R} - R_t\right)\, dt \;+\; \sigma_R\, dW_t^R && R_0 = c \in \mathbb{R}, \\ dB_t &= \left(r B_t + (R_t - M_t) A_t \right. \\ &\quad \left. - \chi_t - \alpha_t - c_2 \alpha_t^2\right)\, dt, && B_0 = b \in \mathbb{R}, \end{aligned} \tag{16.2}$$

where:

$$dW_t\, dW_t^R \;=\; \rho\, dt. \tag{16.3}$$

Note that there are no state constraints and no control constraints in Problem 7 . In Section 7.6, we already showed that the state (short-selling) constraint and the control constraints of the original Problem 2 (5.13) - (5.16) are automatically fulfilled through the problem setup.

Note also that we do not consider the maximum operator $max(R_t, 0)$ on the right hand side of the thrid equation for dB_t in (16.2) because there is no theory for non differentiable terms. However, since the term $max(R_t, 0)$ is continuous, we can approximate it by a $C^2(\mathbb{R}, \mathbb{R})$-function in order to make the term differentiable. The numerical consequences are minimal. We already showed in Section 7.5 that we still can run the CFB algorithm when the term $max(R_t, 0)$ is considered because the derivative of the value function w.r.t. the return rate $\partial_c v$ is not needed in the algorithm.

16.1 Definition of the adjoint variables and the Hamilton function

Let us define $X_t := (A_t, R_t, B_t)$ and $\pi_t := (\alpha_t, \chi_t)$ as the state and control variables. Let $Y_t \in \mathbb{R}^3$ and $Z_t \in \mathbb{R}^{2\times 2}$ denote the first-order adjoint variables. Then, according to Definition 11.6 in Section 11.3, the Hamilton function

$H : [0,T] \times \mathbb{R}^3 \times \mathbb{R}^2 \times \mathbb{R}^3 \times \mathbb{R}^{2\times 2} \to \mathbb{R}$ for our Problem 7 is given by:

$$\begin{aligned} H(t,x,\pi,y,z) = \quad & (\mu a + \alpha)\, y^1 \; + \; \kappa(\bar{R} - c)\, y^2 \\ & + (rb - \chi + (c - M)a - \alpha - c_2\alpha^2)\, y^3 \\ & + \sigma a z^{11} + \sigma_R \rho z^{21} + \sigma_R \sqrt{1-\rho^2} z^{22} \\ & + e^{-\delta t} \tfrac{1-e^{-\gamma\chi}}{\gamma}, \end{aligned} \tag{16.4}$$

where $x = (a,c,b) \in \mathbb{R}^3$ and $\pi = (\alpha,\chi) \in \mathbb{R}^2$. The terminal function $g : \mathbb{R}^3 \to \mathbb{R}$ of Problem 7 is given by:

$$g(x) = e^{-\delta T} \frac{1 - e^{-\gamma\,(a+b)}}{\gamma}. \tag{16.5}$$

For later purpose, we define the discounted gain function $f : [0,T] \times \mathbb{R} \to \mathbb{R}$:

$$f(t,\chi) = e^{-\delta t} \frac{1 - e^{-\gamma\chi}}{\gamma}. \tag{16.6}$$

Lemma 16.1
For a fixed $c \in \mathbb{R}$, the Hamilton function H is concave (linear) w.r.t. the state pair $(a,b) \in \mathbb{R}^2$, it is concave w.r.t. the control $\pi = (\alpha,\chi) \in \mathbb{R}^2$ and linear w.r.t the adjoint variables y and z. Furthermore, the terminal function g is concave w.r.t. $x = (a,c,b) \in \mathbb{R}^3$.

Proof
First, the Gradient of the Hamilton function w.r.t (a,b) are given by:

$$\nabla_{(a,b)} H(t,x,\pi,y,z) = \begin{pmatrix} \mu y^1 + (c-M) y^3 + \sigma z^{11} \\ r y^3 \end{pmatrix}, \tag{16.7}$$

and the Hessian of the Hamilton function w.r.t. (a,b) are given by:

$$\nabla^2_{(a,b)(a,b)} H(t,x,\pi,y,z) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

So the Hamilton H is concave (linear) w.r.t. (a,b).

Second, the Jacobian of the Hamilton function w.r.t π is given by:

$$\nabla_\pi H(t,x,\pi,y,z) = \begin{pmatrix} y^1 - (1+2c_2\alpha)\, y^3 \\ e^{-\delta t} e^{-\gamma \chi} - y^3 \end{pmatrix}, \tag{16.8}$$

and the Hessian of the Hamilton function w.r.t. π is given by:

$$\nabla^2_{\pi\pi} H(t,x,\pi,y,z) = \begin{pmatrix} -2c_1^1 y^3 & 0 \\ 0 & -\gamma e^{-\delta t} e^{-\gamma \chi} \end{pmatrix}.$$

We showed in Theorem 11.14, that y^3 is the first derivative of the value function v w.r.t. the bank account. So $y^3 \geq 0$ naturally. Since $c_2 > 0$ too, all eigenvalues of $\nabla^2_{\pi\pi} H$ are negative (i.e., $\nabla^2_{\pi\pi} H$ is negative definite). Thus, H is concave w.r.t. π.

Third, the Hamilton function H is linear w.r.t. y and z by its definition in (16.4).

Fourth, the Hessian of the terminal function g w.r.t. $x = (a, c, b)$ is given by:

$$\nabla^2_{xx} g(a,c,b) = \begin{pmatrix} \bar{e} & 0 & \bar{e} \\ 0 & 0 & 0 \\ \bar{e} & 0 & \bar{e} \end{pmatrix}, \tag{16.9}$$

where

$$\bar{e} = -\gamma e^{-\delta T} e^{-\gamma(a+b)}.$$

The only non-zero eigenvalue of $\nabla^2_{xx} g$ is $2\bar{e} < 0$. Thus, $\nabla^2_{xx} g$ is negative-semidefinite and g is concave w.r.t x. qed.

16.2 SMP Conditions

Now we show that Problem 7 (16.1) - (16.3) satisfies the SMP Conditions 10.9, 6 - 8:

6. The drift terms $\mu a + \alpha$, $\kappa(\bar{R} - c)$ and $rb + (c - M)a - \chi - \alpha - c_2(\alpha^2)$, and also the diffusion terms σa and σ_R are all continuous differentiable w.r.t. t, x and π. This yields also for the gain function f in (16.6) and the terminal condition g in (16.5).

7. The first derivatives of the drift terms and the diffusion terms w.r.t $.x$, $\begin{pmatrix} \mu & 0 & 0 \\ 0 & -\kappa & 0 \\ c-M & a & r \end{pmatrix}$, $\begin{pmatrix} \sigma & 0 \\ 0 & 0 \end{pmatrix}$, respectively, are Lipschitz continuous w.r.t x. The gain function f, the terminal condition g and their derivatives w.r.t. x are only locally Lipschitz continuous w.r.t x. However, in Section 5.3.1 we showed that the return process R_t, $R_0 = c$ is bounded over the finite time interval. Additionally, we saw in Section 7.6 that the controls α and χ stays bounded and therefore the process X_t, with $X_0 = x$ stays bounded on a finite time interval. Note that a locally Lipschitz continuous function is globaly Lipschitz continuous on a bounded interval.

8. Through Lemma 16.1, we already showed that the Hamilton function H is concave w.r.t. the state pair (a, b), that H is concave w.r.t. the control pair $\pi = (\alpha, \chi)$ and that H is linear w.r.t the adjoint variables y and z. We also showed that the terminal condition g is concave w.r.t. x. Thus, condition 8 is almost satisfied. The only thing missing is, that H is not concave w.r.t. c. However, we do not need H to be concave w.r.t. c for the SMP theory, because the the return process R_t with $R_0 = c$ is not controlled (i.e., it does not depend on the controls α or χ) (see Section 5.3.1).

In summary, Problem 7 (16.1) - (16.3) satisfies only a local Lipschitz condition for f and g and the Hamilton is concave with respect to a and b only. However, the SMP Conditions 10.9 hold for both controlled state processes when they - starting at initial values $(a, c, b) \in \mathbb{R}^3$ - stay bounded over the finite time interval $[0, T]$. In consequence, our CFB algorithm (12.36) - (12.41) of Section 12.6 is applicable to Problem 7. Remember that Problem 7 is an adjusted version of the original Problem 2 (5.13) - (5.16).

17 Formulation of the CFB algorithm

In this section we formulate the specific CFB algorithm of Problem 2 (5.13) - (5.16) (optimal portfolio allocation of production assets) for the one-asset case according to the general CFB scheme (12.36) - (12.41).

17.1 Adjoint variables and optimal controls

Through Lemma 16.1, we know that H and g are concave w.r.t $(a, b) \in \mathbb{R}^2$ and $\pi = (\alpha, \chi) \in \mathbb{R}^2$. Therefore we can apply Theorem 11.12 (SMP's necessary conditions) and obtain the following backward SDE for the adjoint variables $Y_t \in \mathbb{R}^3$ and $Z_t \in \mathbb{R}^{2\times 2}$:

$$\begin{aligned} dY_t &= -\nabla_x H(t, X_t, \pi_t, Y_t, Z_t)dt + Z_t dW_t, \\ Y_T &= \nabla_x g(A_T, B_T). \end{aligned}$$

where the Gradient of the Hamilton function H w.r.t. $x = (a, c, b) \in \mathbb{R}^3$ is given by:

$$\nabla_x H(t, X_t, \pi_t, Y_t, Z_t) = \begin{pmatrix} \mu Y_t^1 + (\max(R_t, 0) - M)\, Y_t^3 + \sigma Z_t^{11} \\ A_t Y_t^3\, 1_{R_t \geq 0} \; - \; \kappa Y_t^2 \\ r Y_t^3 \end{pmatrix}, \tag{17.1}$$

and the terminal condition $\nabla_x g$ is given by:

$$\nabla_x g(A_T, B_T) = \begin{pmatrix} e^{-\delta T} e^{-\gamma (A_T + B_T)} \\ 0 \\ e^{-\delta T} e^{-\gamma (A_T + B_T)} \end{pmatrix}. \tag{17.2}$$

The adjoint variable Y_t^2 is not continuous at $R_t = 0$ when we use the optimized return $\max(R_t, 0)$ instead of R_t. However, the variable Y_t^2 is not needed for the calculation of the optimal controls (α_t^*, χ_t^*) via (17.3) and therefore we need not calculate Y_t^2 in the CFB algorithm.

According to Theorem 11.8 (SMP's sufficient conditions), we obtain the optimal controls by setting the Gradient of H_π (16.8) to zero:

$$\begin{aligned} \alpha_t^* &= \tfrac{1}{2c_2} \left(\tfrac{Y_t^1}{Y_t^3} - 1 \right), \\ \chi_t^* &= -\tfrac{1}{\gamma} \left(\delta t + \ln\left(Y_t^3\right) \right). \end{aligned} \tag{17.3}$$

17.2 The complete CFB algorithm

Now we state the specific CFB algorithm of Problem 2 (5.13) - (5.16). In the following, we use the notations of Section 12. In particular, $\mathrm{C}_k^h \subset \mathbb{R}^3$ denotes the space grids and $\Lambda_k^h \subset \mathbb{R}^2$ denotes the discrete Brownian motion, for all time points t_k, $k = 1, ..., N$ of the discretized time interval $[0, T]$. The grid C_k^h consists of the points (A^i, R^j, B^m), $i = 1, ..., I_k$, $j = 1, ..., J_k$, $q = 1, ..., Q_k$. The equidistant time step size is $\Delta t > 0$.

In order to keep notation clear, we change the notation for the adjoint variables through $(Y_k^A, Y_k^R, Y_k^B) := (Y_k^1, Y_k^2, Y_k^3)$ and simplify notation through $V_k^{i,j,m} := V_k(A^i, R^j, B^m)$. Then, the full CFB algorithm reads:

$\forall (A^i, R^j, B^m) \in \mathrm{C}_k^h$,

$$\begin{aligned}
V_N^{i,j,m} &= e^{-\delta T}\left(1 - e^{-\gamma(A^j + B^m)}\right), \\
{Y_N^A}^{i,j,m} &= e^{-\delta T} e^{-\gamma(A^j + B^m)} \\
{Y_N^B}^{i,j,m} &= e^{-\delta T} e^{-\gamma(A^j + B^m)} \\
\alpha_N^{i,j,m} &= \tfrac{1}{2c_2}\left(\tfrac{{Y_N^A}^{i,j,m}}{{Y_N^B}^{i,j,m}} - 1\right), \\
\chi_N^{i,j,m} &= -\tfrac{1}{\gamma}\left(\delta T + \ln\left(\tfrac{{Y_N^B}^{i,j,m}}{\gamma}\right)\right),
\end{aligned} \tag{17.4}$$

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

$\forall (A^i, R^j, B^m) \in \mathrm{C}_k^h$,

$\forall l \in \Lambda^h$,

$$\begin{aligned}
\bar{A}_{k+1}^i(y^l) &= \Pi_{k+1}\left[A^i + (\mu A^i + \alpha_{k+1}^{i,j,m})\Delta t + \sigma A^i y^{l1}\right], \\
\bar{R}_{k+1}^j(y^l) &= \Pi_{k+1}\left[A^i + \kappa(\bar{R} - R^j)\Delta t + \rho\sigma_R y^{l1} + \sqrt{1-\rho^2}\sigma_R y^{l2}\right], \\
\bar{B}_{k+1}^m(y^l) &= \Pi_{k+1}\left[B^m + \left(rB^m + \max(R^j, 0)A^i - MA^i - \chi_{k+1}^{i,j,m} - \alpha_{k+1}^{i,j,m}\right.\right. \\
&\qquad \left.\left. - c_2 {\alpha_{k+1}^{i,j,m}}^2\right)\Delta t\right].
\end{aligned} \tag{17.5}$$

and then set:

$$
\begin{aligned}
V_k^{i,j,m} &= e^{-\delta t_k}\frac{1-e^{-\gamma\chi_{k+1}^{i,j,m}}}{\gamma}\Delta t \;+\; \sum_{l\in L_k^d} p^l\, V_{k+1}\left(\bar{A}_{k+1}^i(y^l), \bar{R}_{k+1}^j(y^l), \bar{B}_{k+1}^m(y^l)\right),\\
Z_k^{AA^{i,j,m}} &= \frac{1}{\Delta t}\sum_{l\in L_k^d} p^l\, Y_{k+1}^A\left(\bar{A}_{k+1}^i(y^l), \bar{R}_{k+1}^j(y^l), \bar{B}_{k+1}^m(y^l)\right) y^{l1}.
\end{aligned}
\tag{17.6}
$$

$$
\begin{aligned}
Y_k^{A^{i,j,m}} = \;& \mu Y_{k+1}^{A\;\;i,j,m} + \left((\max(R^j,0)-M)Y_{k+1}^{A\;\;i,j,m} + \sigma Z_k^{AA^{i,j,m}}\right)\Delta t\\
&+ \textstyle\sum_{l\in L_k^d} p^l\, Y_{k+1}^A\left(\bar{A}_{k+1}^i(y^l), \bar{R}_{k+1}^j(y^l), \bar{B}_{k+1}^m(y^l)\right)\\
Y_k^{B^{i,j,m}} = \;& rY^{B\,i,j,m}_{k+1} + \textstyle\sum_{l\in L_k^d} p^l\, Y_{k+1}^B\left(\bar{A}_{k+1}^i(y^l), \bar{R}_{k+1}^j(y^l), \bar{B}_{k+1}^m(y^l)\right).
\end{aligned}
\tag{17.7}
$$

$$
\begin{aligned}
\alpha_k^{i,j,m} &= \tfrac{1}{2c_2}\left(\frac{Y^{A^{i,j,m}}_k}{Y^{B^{i,j,m}}_k}-1\right),\\
\chi_k^{i,j,m} &= -\tfrac{1}{\gamma}\left(\delta t_k + \ln\left(\frac{Y_k^{B^{i,j,m}}}{\gamma}\right)\right),
\end{aligned}
\tag{17.8}
$$

Through this backward algorithm, we obtain the values $V_k^{i,j,m}$ and $\pi_k^{i,j,m}$ for all $(A^i, R^j, B^m) \in \mathrm{C}_k^h$ for all $k = N-1, ..., 0$. Afterwards, we can start at time t_0 with $(A_0, R_0, B_0) = (a, c, b)$ and then simulate the processes (A_t, R_t, B_t) forward by projecting the stored optimal controls $\pi_k^{i,j,m}$ on C_k^h to the current state.

Note again that Y_k^R is not considered because this variable is not needed for the calculation of the optimal controls.

18 Implementation issues

In this section, we discuss implementation issues of the problem-specific CFB algorithm (17.4) - (17.8) of Section 17.2. As we mentioned above, all numerical results in Section 7 and 8 are obtained by using this algorithm.

In the implementation, we used an optimal grid by selecting the discretization parameters according to (13.12):

$$\rho = R\,h^{-\frac{1}{2}}, \qquad \delta = h^{\frac{3}{2}}, \qquad L^{-\frac{2}{d}} = h^{2+\frac{(d+2)}{p}}, \tag{18.1}$$

where h denotes the time step size, R denotes the radius of the truncated state space of interest, ρ denotes the extension of this radius to reduce the truncation error, δ denotes the state space step size in each dimension and L denotes the number of discretization points of the optimal Quantization of the Brownian motion in each dimension. Note that in the considered Problem 2, the number of Brownian motions is $d = 2$. Then we have:

1. The number of time steps $N = T/h$.
2. The radius of the state spaces $R = \frac{A_{max}-A_{min}}{2} = \frac{B_{max}-B_{min}}{2} = 10\frac{R_{max}-R_{min}}{2}$,
3. The radius extension $\rho = R/\sqrt{T/N}$. For example, the considered A-state space is $[A_{min} - \rho,\, A_{max} + \rho]$.
4. The number of space points $M = 2R\left((\frac{N}{T})^{3/2} + (\frac{N}{T})^2\right)$ in each dimension.
5. The number of Quantization points $L = (\frac{N}{T})$ in each dimension.

18.1 Linear Interpolation

When we use the projection mapping (nearest neighbor method) Π_k (17.5) in our example calculations, we obtain a rough surface for the value function $V_0(a, c, b)$ and the controls $\alpha_0(a, c, b)$, $\chi_0(a, c, b)$ over the state space $(a, c, b) \in [A_{min}, A_{max}] \times [R_{min}, R_{max}] \times [B_{min}, B_{max}]$. This is a crucial issue because the optimal allocation α_k is calculated through a division of $\frac{Y_k^A}{Y_k^B}$ in (17.8). The division amplifies the error. It is clear that the roughness is reduced when the approximation parameter h is smaller. However, when we replace the nearest neighbor method by a multi-dimensional linear inter- and extrapolation method for V and Y in (17.6) and (17.7), then the roughness disappears completely.

By using the linear inter- and extrapolation, we exploit an additional advantage of the SMP approach, see Section 14.2. In particular, the controls α and χ are calculated through Y. The first derivative Y is a simpler (flatter) functions w.r.t x than the value function V itself. In our cases, Y is almost linear. Therefore we can use the simple linear interpolation to obtain smooth solutions (surfaces). A linear extrapolation also reduces the truncation error. If we work only with the non-linear value function V like in Dynamic Programming, a higher order interpolation methods or finer grids may be necessary to obtain smooth solutions. Then we would also need to define a multi-dimensional, non-linear extrapolation method which is a difficult task.

18.2 Convergence rate

As an example, we show the error convergence for the CFB algorithm (17.4) - (17.8) with linear-interpolation for the one-asset case with fixed returns. The results of this case are presented and discussed in Section 7.2.

To measure the error we used the region of interest $(A, B) \in [0, 1]\times[-0.5, 0.5]$. The error ϵ^V_{max} is defined as the maximum absolute difference to the best numeric solution V_0^{36} over all points $(A^i, B^m) \in \mathrm{C}_0^h$ at time $t = 0$:

$$\epsilon^V_{max} := \max_{(A^i,B^m)\in \mathrm{C}_0^h} |V_0^{36}(A^i, B^m) \; - \; V_0^N(A^i, B^m)|, \tag{18.2}$$

where N is the refinement level i.e., number of time steps used in the approximation. The errors ϵ^{χ}_{max}, ϵ^{α}_{max}, $\epsilon^{Y_A}_{max}$, and $\epsilon^{Y_B}_{max}$ are defined in the same way. We used the discretization parameters according to (18.1) and the model parameters of Table 14:

T	δ	γ	r	μ	σ	$\bar{R}$	c_1	M
4	0.01	0.03	0.1	0.0	0.2	0.07	0.2	0.02

Table 14: One asset parameters

Table 15 shows the approximation errors according to the refinement level N. We obtain a linear convergence in dt for ϵ^{χ}_{max}, ϵ^{α}_{max}, $\epsilon^{Y_A}_{max}$, and $\epsilon^{Y_B}_{max}$ as we proposed in Section 13.

N	dt	dA	ϵ^V_{max}	ϵ^χ_{max}	ϵ^α_{max}	$\epsilon^{Y_A}_{max}$	$\epsilon^{Y_B}_{max}$
4	1.000	1.000	0.02336	0.56199	$2.912e^{-2}$	$7.42e^{-3}$	$5.79e^{-3}$
8	0.500	0.333	0.00165	0.10021	$0.746e^{-2}$	$1.38e^{-3}$	$1.01e^{-3}$
12	0.333	0.192	0.01314	0.05224	$0.452e^{-2}$	$0.76e^{-3}$	$0.54e^{-3}$
16	0.250	0.125	0.00338	0.03335	$0.273e^{-2}$	$0.47e^{-3}$	$0.34e^{-3}$
20	0.200	0.089	0.00542	0.02460	$0.187e^{-2}$	$0.34e^{-3}$	$0.25e^{-3}$
24	0.167	0.068	0.00620	0.01958	$0.145e^{-2}$	$0.27e^{-3}$	$0.20e^{-3}$
28	0.143	0.054	0.00798	0.01874	$0.109e^{-2}$	$0.25e^{-3}$	$0.20e^{-3}$
32	0.125	0.044	0.00354	0.00910	$0.050e^{-2}$	$0.12e^{-3}$	$0.09e^{-3}$
36	0.111	0.037	-	-	-	-	-

Table 15: Maximum absolute differences to the best numeric solution V_0^{36} over all points $(A^i, B^m) \in \mathrm{C}_0^h$ at time $t = 0$.

A Appendix

A.1 Basics about probability spaces and stochastic processes

In this subsection, we state a few basic definitions about probability spaces and stochastic processes. Since probability theory is not the focus of this thesis, we do not go into more detail here. For an appropriate introduction into probability theory we refer to [Grimmett, Stirzaker 1992]. The most definitions are taken from [Yong, Zhou 1999].

Definition A.1 (Probability space)
A probability space is usually defined by $(\Omega, \mathcal{F}, P)$. *Here,* Ω *denotes a non-empty sample space,* $\mathcal{F}$ *denotes a sigma-field on* Ω, *and* P *denotes a probability measure on* $(\Omega, \mathcal{F})$. *A point* $\omega \in \Omega$ *is called a sample.*

Definition A.2 (Probability measure)
A map $P : \mathcal{F} \to [0,1]$ *is called a probability measure on* $(\Omega, \mathcal{F})$ *if:*

$$\begin{gathered} P(\phi) = 0, \quad P(\Omega) = 1, \\ A_i \in \mathcal{F}, \;\; A_i \bigcap A_j = \phi, \;\; i, j = 1, 2, ..., \;\; i \neq j, \\ \Rightarrow P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P\left(A_i\right). \end{gathered} \tag{A.1}$$

Definition A.3 (Filtration)
A filtration $\mathbb{F} := \{F_t, t \geq 0\}$ *on a probability space* $(\Omega, \mathcal{F}, P)$ *is a monotone sequence of sub-sigma-fields* $F_t \subseteq \mathcal{F}$ *which satisfies* $F_t \subseteq F_\tau$, *for all* $t \leq \tau$.

Definition A.4 (Stochastic process)
Let I *be a non-empty index set. A family* $\{X(t), t \in I\}$ *of random variables from* $(\Omega, \mathcal{F}, P)$ *to* $\mathbb{R}^n$ *is called a stochastic process and denoted by* X_t. *For any* $\omega \in \Omega$, *the map* $t \to X(t, \omega)$ *is called a sample path.*

Definition A.5 (Measurable process)
A stochastic process $X : [0,T] \to \mathbb{R}^n$ *is called measurable, if the map* $(t, \omega) \to X(t, \omega)$ *is* $(B[0,T] \times \mathcal{F})/B(\mathbb{R}^n)$*-measurable:*

$$X_t^{-1}\left(B(\mathbb{R}^n)\right) \subseteq B[0,T] \times \mathcal{F}, \tag{A.2}$$

where $B(\cdot)$ *is the Borel* σ*-field.*

Definition A.6 ($\mathbb{F}$-adapted process)
A measurable process $X : [0,T] \to \mathbb{R}^n$ is called $\mathbb{F}$-adapted, if for all $t \in [0,T]$, the map $(s,\omega) \to X(s,\omega)$ from $[0,t] \times \Omega$ into $\mathbb{R}^n$ is $(B[0,t] \times F_t)/B(\mathbb{R}^n)$-measurable.

Definition A.7 (admissible process)
A $\mathbb{F}$-adapted process $\pi : [0,T] \to A \subset \mathbb{R}^r$ is called admissible, if:

$$E\left[\int_0^T |\pi_t|^m dt\right] < \infty \quad \text{for } m = 1, 2, \cdots. \tag{A.3}$$

If A is compact, then $|\pi_t| \leq M$ for some $M < \infty$ and A.3 automatically holds.

Definition A.8 (Brownian motion)
A $\mathbb{F}$-Brownian motion $W(\cdot)$ is a time-homogeneous, $\mathbb{F}$-adapted stochastic process with continuous paths. Furthermore, for all $0 \leq s < t$, $W(t) - W(s)$ is independent of F_t and normally distributed with mean 0 and covariance $(t-s)$. $W(\cdot)$ is called standard if $W(0) = 0$ almost surely. We usually say that $\{W_t, t \geq 0\}$ denotes a standard $\mathbb{F}$-Brownian motion over $[0,\infty)$.

Lemma A.9 (Generation of a filtration)
A Brownian motions W_t naturally generates a filtration $\mathbb{F}$ on the probability space $(\Omega, \mathcal{F}, P)$. The specific construction of this generated filtration is showed for example in [Yong, Zhou 1999] chapter 1, section 2.2. Loosely speaking, the set F_t contains all information up to time t.

Definition A.10 (Markov process)
A stochastic process X_t adapted to the filtration $\mathbb{F}$ is said to possess the Markov property w.r.t. $\mathbb{F}$ if, for each $x \in \mathbb{R}^m$ and each $s, t \in I$:

$$P\left(X_t = x | F_s\right) = P\left(X_t = x | X_s\right). \tag{A.4}$$

A stochastic process is called a Markov process if it satisfies the Markov property w.r.t. its natural filtration.

Definition A.11 (Markov control)
A control policy π is called Markov control if it has the following form: $\pi_t = p(t, X_t)$ where $p : [0,T] \times \mathbb{R}^n \to A$ is a measurable function and X_t is the stochastic process. Roughly speaking, π_t only depends on the current time t and the current state X_t. If the stochastic process X_t itself is controlled by π, then it is called feedback controlled process. Here, the feedback controlled process would still be a Markov process since it satisfies (A.4).

Definition A.12 (Orenstein-Uhlenbeck process)
An Ornstein-Uhlenbeck process X_t satisfies the following SDE:

$$dX_t = \kappa(\mu - X_t)dt + \sigma dW_t, \quad X_0 = x_0, \tag{A.5}$$

where $\kappa, \mu, \sigma, x_0 > 0$. The process is stationary and has:

$$\begin{aligned} E_0[X_t] &= X_0 e^{-\kappa t} + \mu\left(1 - e^{-\kappa t}\right), \\ Var[X_t] &= \frac{\sigma^2}{2\kappa}. \end{aligned} \tag{A.6}$$

The expectation goes to μ when time goes to infinity and the variance does not depend on time, see [Ornstein, Uhlenbeck 1930].

Theorem A.13 (Boundedness of the Orenstein-Uhlenbeck process)

Let X_t be a solution of (A.5). Then, for any $\alpha < \kappa$:

$$\limsup_{t\to\infty} \frac{|X_t|}{\sqrt{\ln t}} \leq \frac{\sigma}{\sqrt{\alpha}}, \quad \textit{almost surely.} \tag{A.7}$$

Proof A.14 (Boundedness of the Orenstein-Uhlenbeck process)
For a fixed $\alpha < \kappa$ and a fixed $\rho < \frac{\kappa^2\mu^2}{4(\kappa-\alpha)}$ we have:

$$x\kappa(\mu - x) \leq -\alpha x^2 + \rho. \tag{A.8}$$

Now we use [Li 2008][55] Theorem 1 to complete the proof. qed.

A.2 Analytical solution for negative exponential utiliy

The negative exponential utility function:

$$u(t,c) = e^{-\delta t}\frac{1 - e^{-\gamma c}}{\gamma}, \quad \gamma > 0,$$

and the terminal fuction:

$$u(T,x) = e^{-\delta T}\frac{1 - e^{-\gamma x}}{\gamma}, \quad \gamma > 0,$$

have constant absolute risk aversion $\lambda_1 = 1$. We suppose a solution of the value function v is of the form:

[55] Originally, the theorem was poposed by [Mao 1992].

$$v(t,x) = e^{-\delta t}\left[a(t)e^{b(t)x} + k(t)\right]. \tag{A.9}$$

with

$$a(T) = -\frac{1}{\gamma}, \quad b(T) = -\gamma, \quad k(T) = \frac{1}{\gamma}.$$

By using the HJB equation:

$$\begin{aligned} 0 = \partial_t v(t,x) + \max_{c\in\mathbb{A}}\left[e^{-\delta t}u(cx) + (r-c)x\,\partial_x v\right] \\ + \max_{\pi\in\mathbb{A}}\left[(\mu - r)\pi x\,\partial_x v + \tfrac{\sigma^2}{2}\pi^2 x^2\,\partial_{xx} v\right]. \end{aligned} \tag{A.10}$$

and plugging in our Ansatz (A.9), we get:

$$\begin{aligned} 0 = -\delta(ae^{bx} + k) + (\dot{a} + a\dot{b}x)e^{bx} + \tfrac{1}{\gamma} + \max_{c\in\mathbb{A}}\left[-\tfrac{e^{-\gamma cx}}{\gamma} + (r-c)xabe^{bx}\right] \\ + \max_{\pi\in\mathbb{A}}\left[(\mu - r)\pi xabe^{bx} + \tfrac{\sigma^2}{2}\pi^2 x^2 ab^2 e^{bx}\right]. \end{aligned} \tag{A.11}$$

Then, we can obtain the following optimal controls by setting the according derivatives to zero:

$$c^* x = -\frac{1}{\gamma}(\log(ab) + bx), \tag{A.12}$$

$$\pi^* x = -\frac{(\mu - r)}{\sigma^2 b}. \tag{A.13}$$

For k we get the following ODE:

$$\dot{k} + \frac{1}{\gamma} - \delta k = 0, \quad k(T) = \frac{1}{\gamma}.$$

thus we set:

$$k = \frac{1}{\gamma\delta} + (\frac{1}{\gamma} - \frac{1}{\gamma\delta})e^{-\delta(T-t)}).$$

which solves the ODE. Substituting these optimal control into the HJB equation and divide by e^{bx}we can obtain:

$$0 = -\delta a + (\dot{a} + a\dot{b}x) + \left[-\tfrac{ab}{\gamma} + rxab + \tfrac{ab}{\gamma}(\log(ab) + bx)\right] - \tfrac{(\mu-r)^2}{2\sigma^2}a. \tag{A.14}$$

In order to emiminate the x we must have:

$$\dot{b} + rb + \frac{b^2}{\gamma} = 0, \quad b(T) = -\gamma$$

We try the following ansatz:

$$b = -\frac{\gamma r}{1-(1-r)e^{-r(T-t)}},$$

which solve the above ODE. Plugging this into the HJB we get:

$$\frac{\dot{a}}{a} - r\log(ab) = \delta - \frac{r}{1-(1-r)e^{-r(T-t)}} + \frac{1}{2}\frac{(\mu-r)^2}{\sigma^2}, \quad a(T) = -\frac{1}{\gamma}$$

Therefore we try the following ansatz:

$$a = \frac{1}{b}e^{\Delta(t)/r}, \quad \Delta(T) = 0.$$

This solves the above ODE when:

$$\Delta = \left(r - \delta - \frac{1}{2}\frac{(\mu-r)^2}{\sigma^2}\right)(1-e^{-\phi(T-t)}) = q(1-e^{-\phi(T-t)}).$$

Because we get:

$$\left[\frac{-\dot{b}}{b^2}e^{\Delta(t)/r} + \frac{1}{b}e^{\Delta(t)/r}\dot{\Delta}(t)/r\right] / \frac{1}{b}e^{\Delta(t)/r}$$

$$-\Delta(t) - \frac{b}{\gamma} = \delta + \frac{1}{2}\frac{(\mu-r)^2}{\sigma^2},$$

which is:

$$\left[\frac{-\dot{b}}{b} + \dot{\Delta}(t)/r\right] - r + qe^{-\phi(T-t)} - \frac{b}{\gamma} = 0,$$

and

$$\left[\frac{-\dot{b}}{b} - q\phi e^{-\phi(T-t)}/r\right] - r + qe^{-\phi(T-t)} - \frac{b}{\gamma} = 0,$$

setting $\phi = r$ we get:

$$\dot{b} + rb + \frac{b^2}{\gamma} = 0,$$

which is exactly what b solves.

A.2.1 Solution

$$\pi^* x = \frac{(\mu - r)}{r\gamma\sigma^2}\left(1 - (1-r)e^{-r(T-t)}\right), \quad \pi^*(T)x = \frac{\mu - r}{\gamma\sigma^2} \tag{A.15}$$

$$c^* x = \frac{rx}{1-(1-r)e^{-r(T-t)}} + \left[\frac{\delta - r + \frac{(\mu-r)^2}{2\sigma^2}}{\gamma r}\right](1 - e^{-r(T-t)}), \quad c^*(T)x = x \tag{A.16}$$

and

$$\tilde{a} = \frac{1-(1-r)e^{-r(T-t)}}{r}\exp\left([(r-\delta) - \frac{1}{2}\frac{(\mu-r)^2}{\sigma^2}](1 - e^{-r(T-t)})/r\right)$$

$$v(t,x) = \tfrac{e^{-\delta t}}{\gamma}\left[\tfrac{1}{\delta} - \tilde{a}\exp\left(-\tfrac{\gamma r}{(1-(1-r)e^{-r(T-t)})}x\right)\right] + \tfrac{1}{\gamma}(1 - \tfrac{1}{\delta})e^{-\delta T}. \tag{A.17}$$

At Terminal time T we have:

$$v(T,x) = \frac{e^{-\delta T}}{\gamma}\left[\frac{1}{\delta} - \exp\left(-\gamma x\right)\right] + \frac{1}{\gamma}(1 - \frac{1}{\delta})e^{-\delta T}. \tag{A.18}$$

$$v(T,x) = \frac{e^{-\delta T}}{\gamma}\left[1 - \exp\left(-\gamma x\right)\right]. \tag{A.19}$$

A.3 Analytical solution of the ODE systems

In this section we derive the analytical solutions of the ODE system (6.30) for the investment-consumption model with stochastic return in Part II, Section 6.2. Remember that $\beta(t)$ is defined as:

$$\beta(t) := \frac{1 + (\delta - 1)e^{-\delta(T-t)}}{\delta}, \quad \beta(T) = 1. \tag{A.20}$$

The ODE system is given by:

$$\begin{aligned}
\dot{a}(t) &= (\delta + 2\kappa)\, a(t) - \tfrac{\beta(t)}{2\sigma^2}, && a(T) = 0,\\
\dot{b}(t) &= (\delta + 1)\, b(t) - 2\kappa\bar{R}a(t) - \tfrac{\beta(t)(\tilde{\mu}-r)}{\sigma^2}, && b(T) = 0,\\
\dot{c}(t) &= \delta c(t) - \sigma_R^2 a(t) - \kappa\bar{R}b(t) - \beta(t)\left(\tfrac{(\tilde{\mu}-r)^2}{2\sigma^2} + r - \delta\right), && c(T) = 0.
\end{aligned} \tag{A.21}$$

First: We suppose that the solution for $a(t)$ is of the following form:

$$\begin{aligned} a(t) &= y(t)e^{-\delta(T-t)} + z, \\ \dot{a}(t) &= (\dot{y}(t) + \delta y(t))e^{-\delta(T-t)}. \end{aligned}$$

Plugging this Ansatz into (A.21) we get:

$$\begin{aligned} &(\dot{y}(t) + \delta y(t))e^{-\delta(T-t)} \\ &= (\delta + 2\kappa)y(t)e^{-\delta(T-t)} \\ &+(\delta + 2\kappa)z - \tfrac{1}{2\delta\sigma^2} - \tfrac{1}{2\delta\sigma^2}(\delta - 1)e^{-\delta(T-t)}. \end{aligned}$$

Since this equation must hold for arbitrary t, we determine $z = \frac{1}{2\delta(\delta+2\kappa)\sigma^2}$. Dividing by $e^{-\delta(T-t)}$ we get:

$$\dot{y}(t) = 2\kappa\, y(t) - \frac{\delta - 1}{2\delta\sigma^2}, \quad y(T) = -z.$$

The unique solution for this ODE is:

$$y(t) = -(z + \frac{1}{2\kappa}\frac{\delta - 1}{2\delta\sigma^2})e^{-2\kappa(T-t)} + \frac{1}{2\kappa}\frac{\delta - 1}{2\delta\sigma^2}.$$

Now, if we define:

$$h_a := \frac{1}{2\delta(\delta + 2\kappa)\sigma^2}, \quad k_a := \frac{\delta - 1}{2\kappa\, 2\delta\sigma^2}, \tag{A.22}$$

then the solution for $a(t)$ is given by:

$$a(t) = \left(-(k_a + h_a)e^{-2\kappa(T-t)} + k_a\right) e^{-\delta(T-t)} + h_a. \tag{A.23}$$

Second: Since the ODE for $b(t)$ has a similar structure as the ODE for $a(t)$, we suppose that the solution for $b(t)$ is of the following form:

$$\begin{aligned} b(t) &= \left(we^{-(T-t)} + xe^{-2\kappa(T-t)} + y\right) e^{-\delta(T-t)} + z, \\ \dot{b}(t) &= \delta\left(we^{-(T-t)} + xe^{-2\kappa(T-t)} + y\right) e^{-\delta(T-t)} \\ &\quad + \left(2\kappa x e^{-2\kappa(T-t)} + uwe^{-(T-t)}\right) e^{-\delta(T-t)}. \end{aligned}$$

We plug this into (A.21) to get:

$$\begin{aligned}
&\delta\left(we^{-(T-t)} + xe^{-2\kappa(T-t)} + y\right)e^{-\delta(T-t)} \\
&+\left(2\kappa xe^{-2\kappa(T-t)} + we^{-(T-t)}\right)e^{-\delta(T-t)} \\
&=(\delta+1)\left[\left(we^{-(T-t)} + xe^{-2\kappa(T-t)} + y\right)e^{-\delta(T-t)} + z\right] \\
&-2\kappa\bar{R}\left[\left(-(k_a+h_a)e^{-2\kappa(T-t)} + k_a\right)e^{-\delta(T-t)} + h_a\right] \\
&-\tfrac{(\tilde{\mu}-r)}{\sigma^2}\tfrac{(\delta-1)}{\delta}e^{-\delta(T-t)} - \tfrac{(\tilde{\mu}-r)}{\delta\sigma^2}.
\end{aligned}$$

Since this equation must hold for arbitrary t, we determine $z = \frac{2\kappa\bar{R}h_a}{\delta+1} + \frac{(\tilde{\mu}-r)}{\delta(\delta+1)\sigma^2}$. Dividing by $e^{-\delta(T-t)}$ we get:

$$\begin{aligned}
&\delta\left(we^{-(T-t)} + xe^{-2\kappa(T-t)} + y\right) \\
&+\left(2\kappa xe^{-2\kappa(T-t)} + we^{-(T-t)}\right) \\
&=(\delta+1)\left[we^{-(T-t)} + xe^{-2\kappa(T-t)} + y\right] \\
&-2\kappa\bar{R}\left[-(k_a+h_a)e^{-2\kappa(T-t)} + k_a\right] \\
&-\tfrac{(\tilde{\mu}-r)}{\sigma^2}\tfrac{(\delta-1)}{\delta}.
\end{aligned}$$

Using the same argument we determine $y = 2\kappa\bar{R}k_a + \frac{(\tilde{\mu}-r)}{\sigma^2}\frac{(\delta-1)}{\delta}$. We get:

$$\begin{aligned}
&\left((\delta+1)x + 2\kappa\bar{R}(k_a+h_a) - \delta x - 2\kappa x\right)e^{-2\kappa(T-t)} \\
&=(\delta+1-\delta-1)\,we^{-(T-t)} = 0.
\end{aligned}$$

Since $e^t > 0$ we determine $x = \frac{2\kappa\bar{R}(k_a+h_a)}{2\kappa-1}$. Since $b(T) = 0$ we determine $w = -(x+y+z)$. Thus, if we define:

$$\begin{aligned}
h_b &:= \tfrac{2\kappa\bar{R}h_a}{\delta+1} + \tfrac{(\tilde{\mu}-r)}{\delta(\delta+1)\sigma^2}, \\
k_b &:= 2\kappa\bar{R}k_a + \tfrac{(\tilde{\mu}-r)}{\sigma^2}\tfrac{(\delta-1)}{\delta}, \\
l_b &:= \tfrac{2\kappa\bar{R}(k_a+h_a)}{2\kappa-1},
\end{aligned} \tag{A.24}$$

then the solution for $b(t)$ is given by:

$$b(t) = \left[-(l_b + k_b + h_b)e^{-(T-t)} + l_b e^{-2\kappa(T-t)} + k_b\right] e^{-\delta(T-t)} + h_b. \tag{A.25}$$

Third: Since the ODE for $c(t)$ has a similar structure as the ODE for $b(t)$, we suppose that the solution for $c(t)$ is of the following form:

$$\begin{aligned} c(t) = &\quad \left(we^{-(T-t)} + xe^{-2\kappa(T-t)} + y(t)\right) e^{-\delta(T-t)} + z, \\ \dot{c}(t) = &\quad \delta\left(we^{-(T-t)} + xe^{-2\kappa(T-t)} + y(t)\right) e^{-\delta(T-t)} \\ &+ \left(we^{-(T-t)} + 2\kappa xe^{-2\kappa(T-t)} + \dot{y}(t)\right) e^{-\delta(T-t)}. \end{aligned}$$

Plugging this Ansatz into (A.21) we get:

$$\begin{aligned} &\delta\left(we^{-(T-t)} + xe^{-2\kappa(T-t)} + y(t)\right) e^{-\delta(T-t)} \\ &+ \left(we^{-(T-t)} + 2\kappa xe^{-2\kappa(T-t)} + \dot{y}(t)\right) e^{-\delta(T-t)} \\ &= \delta\left[\left(we^{-(T-t)} + xe^{-2\kappa(T-t)} + y(t)\right) e^{-\delta(T-t)} + z\right] \\ &-\sigma_R^2\left[\left(-(k_a + h_a)e^{-2\kappa(T-t)} + k_a\right) e^{-\delta(T-t)} + h_a\right] \\ &-\kappa\bar{R}\left[\left[-(l_b + k_b + h_b)e^{-(T-t)} + l_b e^{-2\kappa(T-t)} + k_b\right] e^{-\delta(T-t)} + h_b\right] \\ &-\nu\frac{(\delta-1)e^{-\delta(T-t)}}{\delta} - \frac{\nu}{\delta}, \end{aligned}$$

where $\nu := \frac{(\tilde{\mu}-r)^2}{2\sigma^2} + r - \delta$. Since this equation must hold for arbitrary t, we determine $z = \frac{\sigma_R^2 h_a}{\delta} + \frac{\kappa\bar{R}h_b}{\delta} + \frac{\nu}{\delta^2}$. Dividing by $e^{-\delta(T-t)}$ we get:

$$\begin{aligned} &\left(we^{-(T-t)} + 2\kappa xe^{-2\kappa(T-t)} + \dot{y}(t)\right) \\ &= -\sigma_R^2\left[-(k_a + h_a)e^{-2\kappa(T-t)} + k_a\right] \\ &-\kappa\bar{R}\left[-(l_b + k_b + h_b)e^{-(T-t)} + l_b e^{-2\kappa(T-t)} + k_b\right] \\ &-\nu\frac{(\delta-1)}{\delta}. \end{aligned}$$

Using the same argument we determine $\dot{y}(t) = -\sigma_R^2 k_a - \kappa\bar{R}k_b - \nu\frac{(\delta-1)}{\delta}$. Since $c(T) = 0$, $y(t)$ is given by:

$$y(t) = \left[\sigma_R^2 k_a + \kappa\bar{R}k_b + \nu\frac{(\delta-1)}{\delta}\right](T-t) - (w+x+z).$$

Then the equation becomes:

$$\begin{aligned} &+\left(w - \kappa\bar{R}(l_b + k_b + h_b)\right)e^{-(T-t)} \\ &= (\sigma_R^2(k_a + h_a) - \kappa\bar{R}l_b - 2\kappa x)e^{-2\kappa(T-t)}. \end{aligned}$$

Again, this equation must hold for arbitrary t. Thus, $x = \frac{\sigma_R^2(k_a+h_a)}{2\kappa} - \frac{\kappa\bar{R}l_b}{2\kappa}$ and $w = \kappa\bar{R}(l_b + k_b + h_b)$. Now, if we define:

$$\begin{aligned} h_c &:= \frac{\sigma_R^2 h_a}{\delta} + \frac{\kappa\bar{R}h_b}{\delta} + \frac{1}{\delta^2}\left(\frac{(\tilde{\mu}-r)^2}{2\sigma^2} + r - \delta\right), \\ k_c &:= \sigma_R^2 k_a + \kappa\bar{R}k_b + \frac{(\delta-1)}{\delta}\left(\frac{(\tilde{\mu}-r)^2}{2\sigma^2} + r - \delta\right), \\ l_c &:= \frac{\sigma_R^2(k_a+h_a)}{2\kappa} - \frac{\kappa\bar{R}l_b}{2\kappa}, \\ m_c &:= \kappa\bar{R}(l_b + k_b + h_b), \end{aligned} \tag{A.26}$$

then the solution for $c(t)$ is given by:

$$c(t) = \left(m_c\, e^{-(T-t)} + l_c\, e^{-2\kappa(T-t)} + k_c\,(T-t) - (m_c + l_c + + h_c)\right)e^{-\delta(T-t)} + h_c. \tag{A.27}$$

References

[Akian, Sequier, Sulem 1995] Akian, M., Sequier, P., Sulem, A., *A finite horizon multidimensional portfolio selection problem with singular transactions*, Proceedings of the 34th Conference on Decision & Control, New Orleans, LA, December 1995

[Arrow 1965] Arrow, K.J., *Aspects of the Theory of Risk Bearing, The theory of risk aversion*, Yrjo Jahnssonin Saatio, Helsinki, 1965.

[Awerbuch, Burger 2003] Awerbuch, S., Berger, M., *Energy Security and Diversity in the EU: A Mean-Variance Portfolio Approach*, IEA Research Paper, Paris, February 2003.

[Awerbuch, Stirling, Jansen, Beurskens 2005] Awerbuch, S., Stirling, A.C., Jansen, J., Beurskens, *Full-spectrum portfolio and diversity analysis of energy technologies*, in Leggio, K.B., Bodde, D.L., Taylor, M.L., *Managing Enterprise Risk*, Elsevier Global Energy Policy and Economics Series, 2006

[Barles, Soner 1998] Barles, G., Soner, H. *Option pricing with transaction costs and a nonlinear Black-Scholes equation*, Springer, Finance and Stochastics 2, pp. 369-397, 1998

[Black, Scholes 1973] Black, F., Scholes, M., *The Pricing of Options and Corporate Liabilities*, The Journal of Political Economy, Vol. 81, No. 3 May - Jun., 1973, pp. 637-654

[Black 1976] Black, F., *The pricing of commodity contracts*, Journal of Financial Economics, 3, 1976

[Bloomberg New Energy Finance 2012] Bloomberg New Energy Finance, *Q1 2012 European Long-Term Generation Outlook*, February 2012

[Bellman 1957] Bellman R., *Dynamic Programming*, Princeton University Press, 1957

[Bender, Zhang 2008] Bender, C., Zhang, J., *Time Discretization and Markovian iteration for coupled FBSDEs*, The Annals of Applied Probability, Vol. 18, No. 1, pp. 143-177, 2008

[Bodea 2012] Bodea, A., *Valuation of Swing Options in Electricity Commodity Markets*, PhD Thesis, University of Heidelberg, 2012

[Broadie, Detemple 1994] Broadie, M., Detemple, J., *The Valuation of American Options on Multiple Assets*, Scientific Series, Montréal, September 1994

[Burger, Graeber, Schindlmayr 2007] Burger, M., Graeber, B., Schindelmayr, G., *Managing Energy Risk, An integrated View on Power and Other Energy Markets*, John Wiley & Sons Ltd, England, 2007

[Burkholder, Davis, Gundy 1972] Burkholder, D.L., Davis, B.J., Gundy, R.F., *Integral inequalities for convex functions of operators on martingales*, Univ. of Calif. Press, Proc. Sixth Berkeley Symp. on Math. Statist. and Prob., Vol. 2, pp. 223-240, 1972

[Campbell, Huisman, Koedijk 2001] Campbell, R., Huisman, R., Koedijk, K., *Optimal portfolio selection in a Value-at-Risk framework*, Elsevier, Journal of Banking & Finance 25, pp. 1789-1804, 2001

[Carmona 2008] Carmona, R., *Indifference pricing: theory and applications*, Princeton University Press, 2008

[Cetin, Jarrow, Protter 2004] Cetin, U., Jarrow,R., Protter, P., *Liquidity risk and arbitrage pricing theory*, Springer, Finance and Stochastics 8, 311-341, 2004

[Clewlow, Strickland 1999] Clewlow, L., Strickland., C., *A Multi-Factor Model for Energy Derivative*, By the authors, 1999

[Clewlow et al. 2009] Clewlow L., Meador D., Sobey R., Strickland C., *Valuing Generation Assets*, Masterclass series at energyrisk.com, June 2009

[Constantinides, Zariphopoulou 1999] Constantinides, G., Zariphopoulou, T. *Bounds on prices of contingent claims in an intertemporal economy with proportional transaction costs and general preferences*, Springer, Finance and Stochastics 3, pp. 345-369, 1999

[Coulon, Howison 2009] Coulon, M., Howison, S., *Modelling Price Dynamics Through Fundamental Relationships in Electricity and Other Energy Markets*, PhD Thesis, University of Oxford, 2009

[Dai, Jiang, Li, Yi 2009] Dai, M., Jiang, L., Li, P., Yi, F., *Finite Horizon Optimal Investment and Consumption with Transaction Costs*, SIAM J. Control Optim., Vol. 48, No. 2, pp. 1134-1154, 2009

[Davis 1997] Davis, M.H.A., *Option pricing in incomplete markets*, In: Mathematics of Derivatives Securities, By: Dempster, M.A.H, Pliska, S.R., pp. 216-226, Cambridge University Press, Cambridge, UK, 1997

[Davis, Norman 1990] Davis, M.H.A., Norman, A.R., *Portfolio selection with transaction costs*, Mathematics of Operations Research, Vol. 15, No. 4, November 1990

[Delarue 2010] Delarue, E., et al., *Applying portfolio theory to the electricity sector: Energy versus power*, Elsevier, Energy Economics, 2010

[Delarue, Menozzi 2005] Delarue, F., Menozzi, S., *A forward-backward stochastic algorithm for quasi-linear PDEs*, Universite Paris, December 2005

[Deng 2000] Deng, S., *Stochastic Models of Energy Commodity Prices and Their Applications: Mean-reversion with Jumps and Spikes*, University of California Energy Institute, Februar 2000

[Deng, Johnson, Sogomonian 2001] Deng, S. , Johnson, B., Sogomonian, A., *Exotic electricity options and the valuation of electricity generation and transmission assets*, Elsevier, Decision Support Systems 30, pp. 383-392, 2001

[Douglas, Ma, Protter 1996] Douglas, J., Ma, J., Protter, P., *Numerical Methods for Forward-Backward Stochastic Differential Equations*,The Annals of Applied Probability, Vol. 6, No. 3, pp. 940-968, 1996

[Evans 2002] Evans, L., *Partial differential equations*, Graduate Studies in Mathematics Volume 19, American Mathematical Society, 2001

[Eydeland, Geman 1995] Eydeland, A., Geman, H., *Asian options revisited: Inverting the Laplace Transform*, RISK, March 1995

[Eydeland, Geman 1998] Eydeland, A., Geman, H., *Some Fundamentals of Electricity Derivatives*, University Paris IX Dauphine and ESSEC, July 1998

[Fleming, Soner 2005] Fleming,W.H., Soner, H.M., *Controlled Markov Processes and Viscosity Solutions*, 2nd Edition, Springer, 2005

[Friend, Blume 1975] Friend, I., Blume, M., *The Demand for Risky Assets*, The American Economic Review, Vol. 65, No. 5, December 1975, pp. 900-922

[Geman, Yor 1993] Geman, H., Yor, M., *Bessel processes, Asian options and perpetuities*, Mathematical Finance, 4(3), 349-375, 1993

[Geman, Nguyen 2002] Geman, H., Nguyen, V.N., *Soybean inventory and forward curves dynamics*, Management Science, Vol. 51, No. 7, pp. 1076-1091, July 2005

[Geman 2005] Geman, H., *Commodities and commodity derivatives*, John Wiley & Sons Ltd, England, 2005

[Gennotte, Jung 1992] Gennotte, G., Jung, A., *Investment Strategies under Transaction Costs: The Finite Horizon Case*, Management Science, Vol 38, No. 11, 1992

[Graf, Luschgy 2000] Graf, S., Luschgy, H., *Foundations of Quantization for Random Vectors*, Springer, Berlin, 2000.

[Grimmett, Stirzaker 1992] Grimmett, G.R., Stirzaker, E.R., *Probability and Random Processes*, Oxford University Press, Oxford, 1992

[Hamilton 1989] Hamilton, J.D., *A new approach to the economic analysis of non-stationary time series*, Econometrica 57, 357-384, 1989

[Heath, Jarrow, Morton 1992] Heath, D., Jarrow, R.A., Morton, A., *Bond pricing and the term structure of interest raters: a new methology*, Econometrica, 60(1), pp. 77-105, 1992

[Huang, Wu, 2007] Huanga, Y-H., Wu, J-H., *A portfolio risk analysis on electricity supply planning*, Energy Policy, Volume 36, Issue 2, pp. 627-641, February 2008

[Ito 1951] Ito, K., *On stochastic differential equations*, Memoirs, American Mathematical Society 4, pp. 1-51, 1951.

[Janecek, Shreve 2005] Janecek, K., Shreve, S., *Asymptotic analysis for optimal investment and consumption with transaction costs*, Springer, Finance Stochast. 8, pp. 181-206, 2004

[Jeanne, Masson 2000] Jeanne, O., Masson, P., *Currency Crises, Sunspots, and Markov-Switching Regimes*, Journal of International Economics 50, pp. 327-350, 2000

[Kaldor 1939] Kaldor, N., *Speculation and economic stability*, The Review of Economic Studies, 7, pp. 1-27, 1939

[Kallberg, Ziemba 1984] Kallberg, J., Ziemba, W., *Mis-specifications in portfolio selection problems*, in G. Bamberg and K. Spremann, *Risk and Capital*, pp. 74-87, Springer, New York, 1984

[Kannan, Zavala 2010] Kannan, A., Zavala, V.M., *A Game-Theoretical Dynamic Model for Electricity Markets* Aswin Kannan and Victor M. Zavala, IEEE, Preprint ANL/MCS-P1792-1010, 2010

[Karaki, Chedid, Ramadan] Karaki, S.H., Chedid, R.B., Ramadan, R., *Probabilistic Performance Assessment of Autonomous Solar-Wind Energy Conversion Systems*, IEEE Transactions on Energy Conversion, Vol. 14, No. 3, September 1999

[Karatzas, Lehoczky, Shreve, Xu 1991] Karatzas, I., Lehoczky, J.P., Shreve, S.E., Xu G., *Martingale and duality methods for utility maximization in an incomplete market*, SIAM J. Control and Optimization, Vol. 29, No. 3, pp. 702-730, May 1991

[Klebaner 2005] Klebaner, F., *Introduction to Stochastic Calculus With Applications*, Imperial College Press, 2005

[Kleindorfer, Li 2004] Kleindorfer, P., Li, L., *Multi-Period VaR-Constrained Portfolio Optimization with Applications to the Electric Power Sector*, University of Pennsylvania, May 2004

[Kramer, 2009] Kramer, L., *Ein Fundamentalmodell zur stochastischen Modellierung von Elektrizitätspreisen*, Master thesis, University of Heidelberg, 2009

[Krylov 1980] Krylov, N.V., *Controlled Diffusion Processes*, Springer-Verlag, New-York, 1980

[Kushner, Dupuis 2001] Kushner, H.J., Dupuis, P.G. *Numerical Methods for Stochastic Control Problems in Continuous Time*, Springer US, 2nd. Edition, 2001

[Leland 1985] Leland, H.E., *Option pricing and replication with transaction costs*, Journal of Finance, 40, pp.1283-1301, 1985

[Li 2008] Li, Y., *On the almost surely asymptotic bounds of a class of Ornstein-Uhlenbeck processes in infinite dimensions*, Journal Syst Science & Complexity, Vol. 21, pp. 416-426, 2008

[Lindgren 1978] Lindgren, G., *Markov Regime Models for Mixed Distributions and Switching Regressions*, Scandinavian Journal of Statistics 5, pp. 81-91, 1978

[Lobo, Fazel, Boyd 2007] Lobo, M., Fazel, M., Boyd, S., *Portfolio optimization with linear and fixed transaction costs*, Annals of Operation Research, Volume 152, Number 1, pp.341-365, 2007

[Ludwig 2010] Ludwig, S., *Valuing power plants under emission reduction regulations and investing in new technologies: An exchange option on real options*, MSc Thesis, University of Oxford, 2010

[Ludwig et al. 2012] Ludwig, S., Sirignano, J., Huang, R., Papanicolaou, G., *A forward-backward algorithm for stochastic control problems, using the stochastic maximum principle as an alternative to dynamic programming*, Proceedings of the 1st International Conference on Operations Research and Enterprise Systems Vilamoura, Portugal, SciTePress, 2012

[Liu 2004] Liu, H., *Optimal Consumption and Investment with Transaction Costs and Multiple Risky Assets*, The Journal of Finance Vol. LIX, No.1, 2004

[Ma, Protter, Young 1994] Ma, J., Protter, P., Young , J., *Solving forward-backward stochastic differential equations explicitly - a four step scheme*, Prob. Th. & Rel. Fields, 98, pp. 339-359, 1994

[Madhavan 2011] Madhavan, A., *Exchange-Traded Funds, Market Structure and the Flash Crash*, Social Science Research Network, October 10, 2011

[Magill, Constantinides 1976] Magill, M.J.P., Constantinides, G.M., *Portfolio selection with ttransaction costs*, Journal of Economic Theory, 13, pp. 245-263, 1976

[Mao 1992] Mao, X., *Almost sure asymptotic bounds for a class of stochastic differential equations*, Stochastics and Stochastic Reports, Vol. 41(1-2), pp. 57-69, 1992

[Markowitz 1952] Markowitz, H.M., *Portfolio Selection*, The Journal of Finance 7, pp.77-91, 1952

[Merton 1969] Merton, R.C., *Lifetime Portfolio Selection under Uncertainty: The Continuous-Time Case*, The Review of Economics and Statistics, Vol. 51, No. 3, pp. 247-257, 1969

[Merton 1971] Merton, R.C., *Optimum Consumption and Portfolio Rules in a Continuous-Time Model*, Journal of Economic Theory 3, pp. 373-413, 1971

[Merton 1976] Merton, R.C., *Option pricing when underlying stock returns are discontinuous*, Journal of Financial Economics 3, pp. 125-144, 1976

[Milstein, Tretyakov 2006] Milstein, G.N., Tretyakov, M.V., *Numerical algorithms for forward-backward stochastic differential equations*, SIAM J. SCI. COMPUT. Vol. 28, No. 2, pp. 561-582, 2006

[Monoyios 2003] Monoyios, M., *Efficient option pricing with transaction costs*, J. Comp. Fin, 7, pp. 107-128, 2003

[Monoyios 2004] Monoyios, M., *Performance of utility-based strategies for hedging basis risk*, Quantitative Finance, 4, pp. 245-255, 2004

[Monoyios 2004b] Monoyios, M., *Option pricing with transaction costs using a Markov chain approximation*, J. Econ. Dyn. Cont., 28, pp. 889-913, 2004

[Muthuraman, Kumar 2006] Muthuraman, K., Kumar, S., *Multidimensional Portfolio Optimization with Proportional Transaction costs*, Mathematical Finance, Vol. 16, No. 2, 2006, pp. 301-335

[Nocedal, Wright 2006] Nocedal, J., Wright, S.J., *Numerical Optimization*, Springer Series in Operation Research and Financial Engineering, 2006

[Ornstein, Uhlenbeck 1930] Ornstein, L. S., Uhlenbeck G.E., *On the theory of Brownian Motion*, Physical Review. 36, pp. 823-841, 1930

[Oksendal, Sulem 2002] Oksendal, B., Sulem, A., *Optimal Consumption and Portfolio with Both Fixed and Proportional Transaction Costs*, SIAM J. Control Optim., 40(6), pp. 1765-1790, 2002

[Pages, Pham, Printems 2004] Pages, G., Pham, H., Printems, J., *An Optimal Markovian Quantization Algorithm for Multi-Dimensional Stochastic Control Problems*, World Scientific Publishing Company, Stochastics and Dynamics, Vol. 4, No. 4, pp. 501-545, 2004

[Papanicolaou, Sircar, 1998] Papanicolaou, G., Sircar, K., *General Black-Scholes Models Accounting for Increased Market Volatility from Hedging Strategies*, Appl. Math. Finance 5, pp. 45-82, 1998

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

[Pennanen, Penner, 2010] Pennanen, T., Penner, I., *Hedfing of Claims with Physical Delivery under Convex Transaction Costs*, SIAM, J. Financial Math., Vol. 1, pp. 158-178, 2010

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

[Pindyck 2001] Pindyck, R., *The dynamics of commodity spot and future markets: A primer*, Energy Journal, 22(3), pp. 1-29, 2001

[Pratt 1964] Pratt, J. W., *Risk aversion in the small and in the large*, Econometrica 32, January-April 1964, pp. 122-136.

[Ronn 2002] Ronn E. I., *Real Options and Energy Management*, Risk Waters Group Ltd, 2002

[Rockafellar, Uryasev 2000] Rockafellar, R., Uryasev, S., *Optimization of conditional value-at-risk*, Journal of Risk, Volume 2, Number 3, University of Washington, Spring 2000

[Rogers, Singh 2010] Rogers, L., Singh, S., *The Cost of Illiquidity and its Effects on Hedging*, Wiley, Mathematical Finance, Vol. 20, No. 4, pp. 597-615, 2010

[Roques, Newbery, Nutttall 2007] Roques, F.A., Newbery, D.M., Nuttall, W.J., *Fuel mix diversification incentives in liberalized electricity markets: A Mean-Variance portfolio theory approach*, Elsevier, Energy Economics, December 2007

[Schwartz 1997] Schwartz E.S., *The Stochastic Behavior of Commodity Prices: Implications for Valuation and Hedging*, The Journal of Finance, Vol. 52, No.3, July 1997

[Schwartz, Smith 2000] Schwartz, E.S, Smith, J.E., *Short-Term Variations and Long-Term Dynamics in Commodity Prices*, Management Science, Vol. 46, No. 7, pp. 893-911, Jul. 2000

[Sieczka, Holyst 2009] Sieczka, P., Holyst, J.A., *Correlations in commodity markets*, Elsevier, Physica A, 388, 2009, pp. 1621-1630

[Schönbucher, Wilmott 2000] Schönbucher, P., Wilmott, P., *The feedback effect of hedging in illiquid markets*, SIAM, J. Appl. Math., 61, 2000, pp. 232-272.

[Sorensen 2002] Sørensen, C., *Modeling seasonality in agricultural commodity futures*, The Journal of Futures Markets, Vol. 22, No. 5, pp. 393-426, 2002

[Stirling 1994] Stirling, A., *On the Economics and Analysis of Diversity*, Electronic Working Papers Series, Paper No 28, University of Sussex, 1994

[Szymanski, Bharadwaj, Varadarajan 1993] Szymanski, D.M., Bharadwaj, S.G., Varadarajan, P.R., *An Analysis of the Market Share-Profitability Relationship*, The Journal of Marketing, Vol. 57, No. 3, pp. 1-18, 1993

[Tourin, Zariphopoulou 1997] Tourin, A., Zariphopoulou, T., *Viscosity solutions and numerical schemes for investment/consumption models with transaction costs*, in Numerical Methods in Finance by Rogers and Talay, Cambridge University Press, 1997

[Touzi, Astic 2007] Touzi, N., Astic, F., *No arbitrage conditions and liquidity*, Elsevier, Journal of Mathematical Economics 43, pp. 692-708, 2007

[Treynor 1962] Treynor, J.L., *Toward a Theory of Market Value of Risky Assets*, Unpublished manuscript, 1962. Final version published in Asset Pricing and Portfolio Performance: Models, Strategy and Performance Metrics, R.A. Korajczyk, Risk Books, London, pp. 15-22, 1999

[Trolle, Schwartz 2009] Trolle, A.B., Schwartz, E.S., *Unspanned Stochastic Volatility and the Pricing of Commodity Derivatives*, Oxford University Press, The Society for Financial Studies, 2009

[Vasicek 1977] Vasicek, O., *An equilibrium characterization of the term structure*, Journal of Financial Economics, 5(3), pp. 177-188, 1977

[Wallstreet 2010] The Wallstreet Journal, *Nasdaq: Here's Our Timeline of the Flash Crash*, May 11, 2010

[Weber, Sunderkoetter 2009] Weber, Ch., Sunderkoetter, M., *Valuing fuel diversification in optimal investment policies for electricity generation portfolios*, Chair for Management Science and Energy Economics, Working Paper No. 6, 2009

[Weron, Bierbrauer, Trueck 2004] Weron, R., Bierbrauer, M., Trueck, S. *Modeling electricity prices: jump diffusion and regime switching*, Elsevier, Physica A, 336, pp. 39-48, 2004

[Working 1948] Working, H., *Theory of the inverse carrying charge in Futures markets*, Journal Farm Economics, 30, pp. 1-28, 1948

[Xu 2004] Xu, Z., *Stochastic Models for Gas Prices*, Thesis, University of Calgray, 2004

[Yong, Zhou 1999] Yong, J., Zhou, X.Y., *Stochastic Controls: Hamiltonian Systems and HJB Equations*, Springer, New-York, 1999

[Zhang, Elliott, Siu 2012] Zhang, X., Elliott, R.J., Siu, T.K., *A Stochastic Maximum Principle for a Markov Regime-Switching Jump-Diffusion Model and its Application to Finance*, Accepted by SIAM Journal on Control and Optimization, 2012

[CME Group, Webpage 2011] CME Group, Webpage, *http://www.cmegroup.com*, May 2011

[CME Group, Wheat Future 2011] CME Group, CBOT Wheat Future grades, *http://www.cmegroup.com/ rulebook/CBOT/II/14/104.html*, CBOT Rulebook, Chapter 14., Rule 14104., CME Group, Chicago, May 2011

[CME Group, WTI Future 2011] CME Group, NYMEX Light Sweet Crude Oil (WTI) Future,

http://www.cmegroup.com/ trading/energy/crude-oil/light-sweet-crude_contract_specifications.html, Contract Specifications, CME Group, Chicago, May 2011

[businessdictionary] *http://www.businessdictionary.com/definition/ market-fragmentation.html*, November 2012

[investopedia, market share] *http://www.investopedia.com/terms/ m/marketshare*, The online dictionary, August 2012

[the free dictionary] *http://financial-dictionary.thefreedictionary.com/ Commodity*, The free dictionary, Farlex, Inc., 2011

[wikipedia, risk aversion] *http://en.wikipedia.org/wiki/Risk_aversion*, July 2012

[wordpress] *http://unapologetic.wordpress.com/2011/05/04/ continuously-differentiable-functions-are-locally-lipschitz/*, July 2012