



HAL
open science

Asymptotic methods in stochastic control and applications in finance

Jiatu Cai

► **To cite this version:**

Jiatu Cai. Asymptotic methods in stochastic control and applications in finance. Computational Finance [q-fin.CP]. Université Paris Diderot, 2016. English. NNT: . tel-01297282

HAL Id: tel-01297282

<https://theses.hal.science/tel-01297282v1>

Submitted on 4 Apr 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Université Sorbonne Paris Cité
Université Paris Diderot



École Doctorale de Science Mathématiques de Paris Centre
Laboratoire de Probabilités et Modèles Aléatoires

THÈSE DE DOCTORAT
Discipline : Mathématiques appliquées

présentée par
Jiatu CAI

Méthodes asymptotiques en contrôle stochastique et applications à la finance

Asymptotic methods in stochastic control and applications in finance

co-dirigée par **Mathieu ROSENBAUM** et **Peter TANKOV**

Soutenue le **25 Mars 2016** devant le jury composé de :

Emmanuel GOBET	Ecole Polytechnique	rapporteur
Jean JACOD	Université Pierre et Marie Curie	examinateur
Johannes MUHLE-KARBE	University of Michigan	rapporteur
Huyên PHAM	Université Paris Diderot	examinateur
Mathieu ROSENBAUM	Université Pierre et Marie Curie	directeur
Mete SONER	ETH Zürich	examinateur
Peter TANKOV	Université Paris Diderot	directeur
Nizar TOUZI	Ecole Polytechnique	examinateur

Résumé

Dans cette thèse, nous étudions plusieurs problèmes de mathématiques financières liés à la présence d'imperfections sur les marchés. Notre approche principale pour leur résolution est l'utilisation d'un cadre asymptotique pertinent dans lequel nous parvenons à obtenir des solutions approchées explicites pour les problèmes de contrôle associés.

Dans la première partie de cette thèse, nous nous intéressons à l'évaluation et la couverture des options européennes. Nous considérons tout d'abord la problématique de l'optimisation des dates de rebalancement d'une couverture à temps discret en présence d'une tendance dans la dynamique du sous-jacent. Nous montrons que dans cette situation, il est possible de générer un rendement positif tout en couvrant l'option et nous décrivons une stratégie de rebalancement asymptotiquement optimale pour un critère de type moyenne-variance. Ensuite, nous proposons un cadre asymptotique pour la gestion des options européennes en présence de coûts de transaction proportionnels. En s'inspirant des travaux de Leland, nous développons une méthode alternative de construction de portefeuilles de réplication permettant de minimiser les erreurs de couverture.

La seconde partie de ce manuscrit est dédiée à la question du suivi d'une cible stochastique. L'objectif de l'agent est de rester proche de cette cible tout en minimisant le coût de suivi. Dans une asymptotique de coûts petits, nous démontrons l'existence d'une borne inférieure pour la fonction valeur associée à ce problème d'optimisation. Cette borne est interprétée en terme du contrôle ergodique du mouvement brownien. Nous fournissons également de nombreux exemples pour lesquels la borne inférieure est explicite et atteinte par une stratégie que nous décrivons.

Dans la dernière partie de cette thèse, nous considérons le problème de consommation et investissement en présence de taxes sur le rendement des capitaux. Nous obtenons tout d'abord un développement asymptotique de la fonction valeur associée que nous interprétons de manière probabiliste. Puis, dans le cas d'un marché avec changements de régime et pour un investisseur dont l'utilité est du type Epstein-Zin, nous résolvons explicitement le problème en décrivant une stratégie de consommation-investissement optimale. Enfin, nous étudions l'impact joint de coûts de transaction et de taxes sur le rendement des capitaux. Nous établissons dans ce cadre un système d'équations avec termes correcteurs permettant d'unifier les résultats de [ST13] et [CD13].

Mots-clefs

Couverture discrète, théorèmes limites, temps d'arrêt, optimalité asymptotique, coût de transactions, contrôle linéaire-quadratique, stratégie de Leland, variance conditionnelle, contrôle singulier, théorème limite central, mesure d'occupation, borne inférieure asymptotique, contrôle impulsionnelle, contrôle moyen en temps, programmation linéaire, problème de martingale, maximisation d'utilité, discrétisation des intégrales stochastiques, coût illiquidité, impact de marché temporaire, taxe de rendement des capitaux, utilité récursive, Epstein-Zin, homogénéisation.

Abstract

In this thesis, we study several mathematical finance problems related to the presence of market imperfections. Our main approach for solving them is to establish a relevant asymptotic framework in which explicit approximate solutions can be obtained for the associated control problems.

In the first part of this thesis, we are interested in the pricing and hedging of European options. We first consider the question of determining the optimal rebalancing dates for a replicating portfolio in the presence of a drift in the underlying dynamics. We show that in this situation, it is possible to generate positive returns while hedging the option and describe a rebalancing strategy which is asymptotically optimal for a mean-variance type criterion. Then we propose an asymptotic framework for options risk management under proportional transaction costs. Inspired by Leland's approach, we develop an alternative way to build hedging portfolios enabling us to minimize hedging errors.

The second part of this manuscript is devoted to the issue of tracking a stochastic target. The agent aims at staying close to the target while minimizing tracking efforts. In a small costs asymptotics, we establish a lower bound for the value function associated to this optimization problem. This bound is interpreted in term of ergodic control of Brownian motion. We also provide numerous examples for which the lower bound is explicit and attained by a strategy that we describe.

In the last part of this thesis, we focus on the problem of consumption-investment with capital gains taxes. We first obtain an asymptotic expansion for the associated value function that we interpret in a probabilistic way. Then, in the case of a market with regime-switching and for an investor with recursive utility of Epstein-Zin type, we solve the problem explicitly by providing a closed-form consumption-investment strategy. Finally, we study the joint impact of transaction costs and capital gains taxes. We provide a system of corrector equations which enables us to unify the results in [ST13] and [CD13].

Keywords

Discrete hedging, limit theorems, stopping times, asymptotic optimality, transaction costs, linear-quadratic control, Leland strategy, conditional variance, singular control, central limit theorem, occupation measure, asymptotic lower bound, impulse control, time-average control, linear programming, martingale problem, utility maximization, discretization of stochastic integrals, illiquidity cost, temporary market impact, capital gains taxes, recursive utility, Epstein-Zin, homogenization.

Contents

INTRODUCTION	1
Part I : Option pricing and hedging	3
1.1 Discrete hedging with directional views	3
1.2 Option replication with modified volatility	5
Part II : Asymptotic optimal tracking	8
2.1 Formulation of the tracking problem	8
2.2 Asymptotic framework	9
2.3 Main results	11
2.4 Relation with other asymptotic studies	12
Part III : Portfolio selection with capital gains taxes	14
3.1 Preliminary : the model of [BST10]	14
3.2 Expansion around tax-deflated model	16
3.3 Capital gains taxes with recursive utility and regime-switching	17
3.4 Joint impact of capital gains taxes and transaction costs	18
I OPTION PRICING AND HEDGING	21
1 Discrete hedging with directional views	23
1.1 Introduction	23
1.2 Assumptions and admissible strategies	26
1.2.1 Assumptions on the dynamics and admissibility conditions	26
1.2.2 Comments on the admissibility conditions	28
1.2.3 Examples of admissible discretization rules	29
1.3 Asymptotic optimality: a preliminary approach	31
1.4 Asymptotic expectation-error optimization	33
1.5 Black-Scholes model with time-varying coefficients	34
1.5.1 Explicit formulas	35
1.5.2 Numerical study	37
Appendix 1.A Proofs	39
1.A.1 Proof of Proposition 1.2.1	39
1.A.2 Proof of Proposition 1.2.2	44
1.A.3 Proof of Proposition 1.2.3	48
1.A.4 Proof of Theorem 1.4.1	52
Appendix 1.B Linear-quadratic optimal control	53
2 Option pricing with modified volatility and proportional costs	55
2.1 Introduction	55
2.2 Framework	60
2.2.1 Benchmark strategy	61
2.2.2 Continuous control strategies	62

2.2.3	Hedging error	63
2.3	Main results	63
2.3.1	Limit theorem of hedging error	63
2.3.2	Minimum conditional variance	66
2.3.3	Asymptotically optimal sequence	67
2.3.4	Implementation and numerical experiments	69
2.4	Heuristic derivation	71
2.4.1	Local probability model	73
2.4.2	Relation with Leland's strategy	75
2.4.3	Relation with Whalley and Wilmott's strategy	76
Appendix 2.A	Proof of Theorem 2.3.1	78
Appendix 2.B	Proof of Theorem 2.3.2	81
II ASYMPTOTIC OPTIMAL TRACKING		93
3	Asymptotic lower bounds	95
3.1	Introduction	95
3.2	Tracking with combined regular and impulse control	100
3.2.1	Asymptotic framework	101
3.2.2	Lower bound	104
3.3	Extensions of Theorem 3.2.1 to other types of control	105
3.3.1	Combined regular and singular control	105
3.3.2	Impulse control	106
3.3.3	Singular control	107
3.3.4	Regular control	108
3.4	Interpretation of lower bounds and examples	108
3.4.1	Martingale problem associated to controlled Brownian motion	109
3.4.2	Time-average control of Brownian motion	110
3.4.3	Explicit examples in dimension one	113
3.5	Proof of Theorem 3.2.1	119
3.5.1	Reduction to local time-average control problem	119
3.5.2	Proof of Theorem 3.2.1	121
3.5.3	Proof of Lemma 3.5.3	122
3.6	Proof of Theorem 3.3.1	124
3.7	Proof of Propositions 3.4.1-3.4.5	124
3.7.1	Verification theorem in \mathbb{R}^d	124
3.7.2	Verification of Proposition 3.4.4	126
Appendix 3.A	Kummer confluent hypergeometric function ${}_1F_1$	128
Appendix 3.B	Proof of Lemma 3.7.2	129
Appendix 3.C	Tightness function	133
Appendix 3.D	Convergence in probability, stable convergence	134
4	Feedback strategies and applications	137
4.1	Introduction	137
4.2	Combined regular and impulse control	138
4.2.1	Feedback strategies	139
4.2.2	Asymptotic performance	141
4.3	Extensions to other types of control	143
4.3.1	Combined regular and singular control	143
4.3.2	When only one control is present	145
4.4	Applications	147

4.4.1	Explicit optimal strategies in dimension one	147
4.4.2	Discretization of hedging strategies	148
4.4.3	Impacts of small market frictions	149
4.5	Proofs	152
4.5.1	Proof of Theorem 4.2.1	152
4.5.2	Proof of Theorem 4.3.1	154
4.5.3	Proof of Theorem 4.3.4	155
Appendix 4.A	Convergence of integral functionals	156
Appendix 4.B	Separability of (A, B)	158
4.B.1	Diffusion with Rebirth	159
4.B.2	Diffusion with Reflection	160
III PORTFOLIO SELECTION WITH CAPITAL GAINS TAXES		163
5	Recursive utility and regime switching	165
5.1	Introduction	165
5.2	Model formulation	167
5.3	Preliminary analysis	169
5.3.1	First properties of the value function	169
5.3.2	Certainty equivalent wealth loss (CEWL)	171
5.3.3	The value of deferring capital gains realization	171
5.3.4	HJB equation	171
5.3.5	Trading and no-trading regions	173
5.4	Asymptotic analysis	174
5.4.1	Small parameter	174
5.4.2	Asymptotic expansion	175
5.4.3	Economic analysis: single-regime case	176
5.4.4	Economic analysis: two-regime case	177
5.5	Numerical results	178
5.5.1	Single-regime case	178
5.5.2	Two-regime Case	182
5.6	Why is the order of deferral value $O(\varepsilon^{8/3})$?	185
5.6.1	Expansion around the tax-deflated model	185
5.6.2	Strategies based on barriers and reduced model	187
5.6.3	Solution of the optimal control problem (5.6.17)-(5.6.18)	189
Appendix 5.A	Formal derivation of Asymptotic Expansion 5.4.1	191
5.A.1	Expansion at solution of tax-deflated model	191
5.A.2	An approximation to the optimal consumption strategy	193
6	Joint impact of capital gains taxes and transaction costs	195
6.1	Main result	196
6.1.1	HJB equation	196
6.1.2	Benchmark model	196
6.1.3	Asymptotic expansion	197
6.2	Probabilistic interpretation	198
6.3	Homothetic case	200
6.4	Asymptotic development	201
6.4.1	Fast variables	201
6.4.2	Corrector equations	203
Bibliography		207

Introduction

This thesis is devoted to the study of three different problems in mathematical finance, which involve various important market features such as time discretization, transaction costs and capital gains taxes. Due to the presence of these market features, the optimal strategies deduced from most of the existing financial models cannot be implemented in practice. Therefore more complex models and advanced theoretical tools have been developed in order to deal with these market imperfections. However, the resulting stochastic control problems often become intractable. While it is possible to obtain numerical solutions in some cases, the required computational effort in reality is usually prohibitively large. In this thesis, we aim at proposing an asymptotic framework for the associated stochastic control problems, providing explicit and feasible optimal strategies. We focus both on the theoretical investigation of the related stochastic control methods, and on the economic analysis of the impact of market frictions.

In the first part, we study the pricing and hedging of option when taking into account discrete rebalancing and transaction costs.

In Chapter 1, we consider the hedging error of a derivative due to discrete trading. It turns out that, in the presence of a drift in the dynamics of the underlying asset, the trader can actually benefit from market trend. We suppose that the trader wishes to find rebalancing times for the hedging portfolio which enable him to keep the discretization error small while taking advantage of market tendency. Assuming that the portfolio is readjusted at high frequency, we introduce an asymptotic framework in order to derive optimal discretization strategies. More precisely, we formulate the optimization problem in terms of an asymptotic expectation-error criterion. In this setting, the optimal rebalancing times are given by the hitting times of two barriers whose values can be obtained by solving a linear-quadratic optimal control problem. In specific contexts such as in the Black-Scholes model, explicit expressions for the optimal rebalancing times can be derived.

In Chapter 2, we study the dynamic hedging of a European option under a general local volatility model with small proportional transaction costs. Extending the idea of Leland, which consists in modifying the volatility in the pricing PDE in order to compensate the costs incurred by discrete rebalancing, we consider instead a continuous version (with finite variation) of Leland's strategy that asymptotically replicates the payoff. In the limit of small proportional costs, an associated central limit theorem for hedging error is proved. The asymptotic variance is minimized by an explicit replication strategy. Depending on the transaction costs and the gamma of the option, the optimal replication strategy is given by either an absolutely continuous process or a singular process based on two barriers around a benchmark position. Numerical simulations demonstrate a significant improvement of our strategies over Leland's strategy in terms of conditional variance of the hedging error.

In the second part, we consider tracking problems which arise from the study of trading strategies under various market frictions and discretization effects. The aim is to minimize both deviation from the target and tracking efforts.

In Chapter 3, we propose an asymptotic framework and establish the existence of asymptotic lower bounds for the value functions of the corresponding control problems. These lower bounds can be related to the time-average control problem of Brownian motion. A key step is the use of a linear programming characterization of the lower bounds. Our probabilistic approach enables us to treat (the combination of) different control types such as (absolutely continuous) regular control, singular control and impulse control. Moreover, the lower bound are shown to hold pathwise. A comprehensive list of examples with closed-form solutions for the lower bounds is also provided.

In Chapter 4, we focus on strategies of feedback form for the problem of tracking and study their performance under our asymptotic framework. Depending only on the current state of the system, these strategies maintain the deviation from the target inside a time-varying domain. Although the dynamics of the target is non-Markovian, it turns out that they asymptotically attain the lower bounds previously established for a large list of examples. We apply our results to the analysis of discretization errors of stochastic integrals and impact of market frictions on portfolio management.

In the third part of this thesis, we perform asymptotic analysis for the problem of consumption-investment optimization with capital gains taxes.

In Chapter 5, we study the optimal consumption and portfolio decisions in the presence of capital gains tax and stochastic investment opportunity. The option to defer taxation of capital gain gives rise to an optimal investment strategy consisting of buy and sell boundaries around the tax-deflated Merton line. In a bull-bear switching market, the optimal investment strategy is affected by the investment opportunity in the other regime. Consequently, there is a cross-regime smoothing effect on the value of the deferral option. Moreover, regime switching has a greater impact under the bear regime and in the region of high capital gains. Depending on the level of capital gains, a sudden change of economic condition might not lead to an instantaneous jump in the optimal portfolio allocation. In contrast, for an investor with recursive utility, the EIS of the investor determines the optimal consumption rate but has little impact on the optimal investment strategy. Our asymptotic analysis is supported not only by numerical resolution of the corresponding PDEs but also by an underlying two time scales probabilistic model.

In Chapter 6, we are interested in the joint impact of capital gains taxes and transaction costs. Guided by the local probability model developed in the previous chapter, we develop a new system of corrector equations, unifying previous results on capital gains taxes and transaction costs via homogenization technique. In particular, we find that the presence of capital gains taxes has an equivalent effect of increasing selling costs for the non-transaction zone.

Part I : Option pricing and hedging

1.1 Discrete hedging with directional views

In order to manage the risks inherent to the derivatives they buy and sell, practitioners use continuous time stochastic models to compute their prices and hedging portfolios. In the simplest cases, notably in that of the so-called delta hedging strategy, the hedging portfolio obtained from the model is a time-varying self-financed combination of cash and the underlying. We denote the price at time t of the underlying asset by Y_t and assume it to be a one-dimensional semi-martingale. Hence, in such situations, the outputs of the model are the price of the option together with the number of shares in the underlying asset to be held in the hedging portfolio at any time t , denoted by X_t . The proportion invested in cash is then deduced from the self-financing property. Therefore, assuming zero interest rates, the theoretical value of the model-based hedging portfolio at the maturity of the option T is given by

$$\int_0^T X_t dY_t.$$

Typically, the process X_t derived from the model is a continuously varying semi-martingale, requiring continuous trading to be implemented in practice. This is of course physically impossible and would be anyway irrelevant because of the costs induced by microstructure effects. Hence practitioners do not use the strategy X_t , but rather a discretized version of it. This means that the hedging portfolio is only rebalanced a finite number of times and is held constant between these times. Let us denote by $(\tau_j^n)_{j \geq 0}$ an increasing sequence of rebalancing times over $[0, T]$. With respect to the target portfolio obtained by continuous rebalancing, the hedging error due to discrete trading Z_T^n is therefore given by

$$Z_T^n = \sum_{j=0}^{+\infty} X_{\tau_j^n} (Y_{\tau_{j+1}^n \wedge T} - Y_{\tau_j^n \wedge T}) - \int_0^T X_t dY_t.$$

Why market trends matter ?

When X and Y are Itô processes, the case of equidistant rebalancing dates $\tau_j^n = jT/n$ has been investigated in [BKL00, HM05, Roo80]. In these works, the following convergence in law is proved:

$$\sqrt{n} Z_T^n \xrightarrow{\mathcal{L}} \sqrt{\frac{T}{2}} \int_0^T \sigma_t^X \sigma_t^Y dB_t, \quad (1.1)$$

where σ^X and σ^Y are the volatilities of X and Y and B is a Brownian motion independent of the other quantities.

This asymptotic approach has also been recently used in [Fuk11c, RR10, GL14a, Lan13], where the rebalancing times are random stopping times. More precisely, for a given parameter n driving the asymptotic, one considers an increasing sequence of stopping times

$$0 = \tau_0^n < \tau_1^n < \dots < \tau_j^n < \dots$$

with $N_t^n := \max\{j \geq 0; \tau_j^n \leq t\} < \infty$ almost surely for any $t \geq 0$. When

$$\frac{\mathbb{E}[(X_{\tau_{j+1}^n} - X_{\tau_j^n})^4 | \mathcal{F}_{\tau_j^n}]}{\mathbb{E}[(X_{\tau_{j+1}^n} - X_{\tau_j^n})^2 | \mathcal{F}_{\tau_j^n}]} = \varepsilon_n^2 a_{\tau_j^n}^2 + o_p(\varepsilon_n^2), \quad \frac{\mathbb{E}[(X_{\tau_{j+1}^n} - X_{\tau_j^n})^3 | \mathcal{F}_{\tau_j^n}]}{\mathbb{E}[(X_{\tau_{j+1}^n} - X_{\tau_j^n})^2 | \mathcal{F}_{\tau_j^n}]} = -\varepsilon_n s_{\tau_j^n} + o_p(\varepsilon_n)$$

for a sequence $\varepsilon_n \rightarrow 0$ and left continuous adapted processes a and s , it is shown in [Fuk11c] that $\varepsilon_n^{-1}Z^n$ converges weakly to

$$\frac{1}{3} \int_0^\cdot s_t dY_t + \frac{1}{\sqrt{6}} \int_0^\cdot (a_t^2 - \frac{2}{3}s_t^2)^{1/2} \sigma_t^Y dB_t, \quad (1.2)$$

where B is a Brownian motion independent of all the other quantities.

One can remark a crucial difference between the deterministic discretization schemes associated to (1.1) and the random stopping times case leading to (1.2). For deterministic dates, the discretization error asymptotically behaves as a stochastic integral with respect to Brownian motion. Therefore, it is centered. In the case of random discretization dates, one may obtain a “biased” asymptotic hedging error because of the presence of the term

$$\int_0^T s_t dY_t.$$

Hence, if s does not vanish and Y has non zero drift, the asymptotic hedging error is no longer centered.

An asymptotic linear-quadratic criterion

A natural question is to determine the efficient frontier for the first and second moments of hedging errors and to find stopping times τ_j^n which attain the efficient frontier. More precisely, we are interested in

$$\inf_{(\tau_j^n)} (-\varepsilon_n^{-1} \mathbb{E}[Z_T^n], \varepsilon_n^{-2} \mathbb{E}[(Z_T^n)^2]), \quad (1.3)$$

as $\varepsilon_n \rightarrow 0$, meaning that the hedging frequency is high and the hedging error should be small. Our main contribution in Chapter 1 is the following result.

Main Result 1. *The efficient frontier of (1.3) can be explicitly determined from the solution of the following optimal control problem*

$$\inf_{(s_t)} -\mathbb{E}[Z_{s,T}^*] + \lambda \mathbb{E}[(Z_{s,T}^*)^2 + \frac{1}{2} \int_0^T s_t^2 (\sigma_t^Y)^2 dt], \quad (1.4)$$

where

$$Z_{s,T}^* = \int_0^T s_t dY_t.$$

Moreover, given an optimal strategy s_t^* for (1.4), one can construct explicitly two barriers \bar{l}_t^* and \underline{l}_t^* such that the sequence of rebalancing dates $(\tau_j^{n,*})$ defined by

$$\tau_{j+1}^{n,*} = \inf \left\{ t > \tau_j^{n,*} : X_t \notin (X_{\tau_j^{n,*}} - \varepsilon_n \underline{l}_t^*, X_{\tau_j^{n,*}} + \varepsilon_n \bar{l}_t^*) \right\} \wedge T$$

attains asymptotically the efficient frontier of (1.3).

Comparing to the classical continuous-time portfolio selection problem without hedging constraint and discrete rebalancing (see [ZL00])

$$\inf_{(s_t)} -\mathbb{E}[Z_{s,T}^*] + \lambda \mathbb{E}[(Z_{s,T}^*)^2],$$

we note that the discrete nature of the hedging strategy introduces an extra uncertainty on the final hedging error, which is approximately quantified in (1.4) by

$$\mathbb{E} \left[\frac{1}{2} \int_0^T s_t^2 (\sigma_t^Y)^2 dt \right],$$

as ε_n tends to zero.

1.2 Option replication with modified volatility

In Chapter 2, we develop a framework for option pricing and replication under proportional transaction costs. As we have seen in the previous section, without transaction costs the hedging error with respect to the target portfolio tends to zero as the rebalancing dates become more frequent. However, transaction costs increase as the rebalancing intervals decrease. Therefore, when pricing and hedging derivatives, one should include transaction costs. Inspired by the work of [Lel85], our approach has the following advantages :

1. The strategy replicates the option payoff under transaction costs, with an error whose distribution is explicitly determined in the limit of small costs.
2. We find explicit strategies minimizing the conditional variance of hedging errors.

In order to explain our framework and state the main results, we begin by a brief review of Leland's strategy in the Black-Scholes framework.

Leland's idea has two ingredients. The first is to find a good benchmark strategy by an enlargement of volatility, yielding certain surplus in the absence of transaction costs . Let p^α be a solution of the partial differential equation

$$\partial_t p^\alpha(s, t) + \frac{1}{2} \left(1 + \frac{2}{\alpha}\right) \sigma^2 s^2 \partial_s^2 p^\alpha(s, t) = 0, \quad p^\alpha(s, T) = f(s), \quad (1.5)$$

where $f(s) = (s - K)^+$ is the payoff of a European call option, T is the maturity of the option, σ is the volatility of Black-Scholes model. Here, α is an arbitrary positive constant that controls the enlargement of volatility, which should be determined by market supply-demand equilibrium. By Itô's formula, we have

$$f(S_T) = \Pi_0^\alpha + \int_0^T X_u^\alpha dS_u - \frac{1}{\alpha} \int_0^T \Gamma_u^\alpha d\langle S \rangle_u,$$

where

$$\Pi_t^\alpha = p^\alpha(S_t, t), \quad X_t^\alpha = \partial_s p^\alpha(S_t, t), \quad \Gamma_t^\alpha = \partial_s^2 p^\alpha(S_t, t). \quad (1.6)$$

This means that, without transaction costs and assuming zero interest rates, the self-financing strategy X^α with initial capital Π_0^α super-hedges the payoff $f(S_T)$ with surplus

$$\frac{1}{\alpha} \int_0^T \Gamma_t^\alpha d\langle S \rangle_t \geq 0, \quad (1.7)$$

where $\Gamma^\alpha \geq 0$ follows from the convexity of the payoff f .

The second step of Leland's strategy is to construct a good approximation of the benchmark X^α by a strategy with finite transaction cost so that the incurred costs are compensated by the surplus. Assume that the trader has to pay $\kappa|\Delta X|$ to buy or sell ΔX shares of stocks, where κ is a positive constant representing the proportional transaction costs. Leland considers an equidistant discretization of X^α defined by

$$X_t^{\alpha, \kappa} = X_{jh}^\alpha, \quad t \in (jh, (j+1)h], \quad j = 0, 1, 2, \dots, \quad (1.8)$$

with $h > 0$ the interval of rebalancing. Set the initial capital $\Pi_{0-}^{\alpha, \kappa}$, that is the price of the option, to be

$$\Pi_{0-}^{\alpha, \kappa} = \Pi_0^\alpha + \kappa S_0 |X_0^\alpha|.$$

The second term is to compensate the transaction cost at the inception. The associated wealth process $\Pi^{\alpha,\kappa}$ under transaction costs is then

$$\Pi_t^{\alpha,\kappa} = \Pi_0^\alpha + \int_0^t X_u^{\alpha,\kappa} dS_u - \kappa \sum_{0 < u \leq t} S_u |\Delta X_u^{\alpha,\kappa}|. \quad (1.9)$$

By choosing

$$h = \frac{2}{\pi} \frac{\kappa^2 \alpha^2}{\sigma^2}, \quad (1.10)$$

we have

$$\begin{aligned} \int_0^T X_u^{\alpha,\kappa} dS_u &\rightarrow \int_0^T X_u^\alpha dS_u, \\ \kappa \sum_{0 < u \leq T} S_u |\Delta X_u^{\alpha,\kappa}| &\rightarrow \frac{1}{\alpha} \int_0^T \Gamma_u^\alpha d\langle S \rangle_u, \end{aligned} \quad (1.11)$$

as $\kappa \rightarrow 0$. Consequently, the terminal wealth $\Pi_T^{\alpha,\kappa}$ is close to $f(S_T)$ when κ is small which is the case in liquid markets. In this sense, the self-financing strategy $X^{\alpha,\kappa}$ is an asymptotic replication strategy. The way to discretize X^α is essential. The first convergence in (1.11) holds in general as transactions are more and more frequent. On the other hand, if they are too frequent, then the total amount of transaction costs exceeds the surplus (1.7) and the second convergence of (1.11) fails. Therefore the frequency (1.10) results from a delicate balance.

Although the approach of Leland is consistent and easy to implement, it is still not fully satisfactory due to the lack of optimality. The strategy (1.8) is not the only choice as an approximation to X^α . For example, it is not necessary to match the benchmark strategy X^α after each rebalancing. In particular, it is possible to trade more frequently but with a smaller trading volume each time. Indeed, several results on related problems under the framework of utility maximization suggest that, under proportional transaction costs, the optimal strategy is to trade a minimal amount in continuous-time to keep the deviation from the benchmark inside a no trade zone (see [WW97, BS98, ST13]).

Our contribution in this context is two-fold. First, we introduce a reasonable class of continuous trading strategies with finite transaction costs, and provide a limit theorem for the corresponding replication error. In particular, we identify conditions for those strategies to (asymptotically) replicate or super-replicate the option. Second, we minimize the asymptotic variance of hedging error among those replicating strategies.

Under our framework, a candidate strategy $X^{b,c,\kappa}$ is indexed by two non-negative functions $b(s, t)$ and $c(z, s, t)$. Let $Z^\kappa = (X^\alpha - X^{b,c,\kappa})/\kappa$ be the normalized deviation of $X^{b,c,\kappa}$ from the benchmark position X^α . We consider $X^{b,c,\kappa}$ of the form

$$dX_t^{b,c,\kappa} = \frac{1}{\kappa} \text{sgn}(Z^\kappa) c(|Z_t^\kappa|, S_t, t) d\langle X^\alpha \rangle_t - \kappa dL_t^\kappa + \kappa dR_t^\kappa, \quad X_{0+}^{b,c,\kappa} = X_0^\alpha,$$

where L^κ and R^κ are non-decreasing processes such that

$$L_t^\kappa = \int_0^t 1_{\{Z_u^\kappa = -b(S_u, u)\}} dL_u^\kappa, \quad R_t^\kappa = \int_0^t 1_{\{Z_u^\kappa = b(S_u, u)\}} dR_u^\kappa, \quad |Z_t^\kappa| \leq b(S_t, t).$$

Intuitively, the regular control part $\kappa^{-1} \text{sgn}(Z^\kappa) c(|Z_t^\kappa|, S_t, t) d\langle X^\alpha \rangle_t$ pushes $X^{b,c,\kappa}$ toward X^α and is active when X^α moves. The singular control part $-\kappa dL_t^\kappa + \kappa dR_t^\kappa$ keeps Z^κ within the stochastic interval $[-b(S_t, t), b(S_t, t)]$, and is active only when Z^κ touches the boundary.

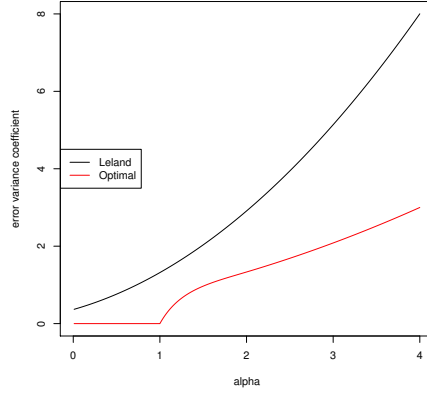


Figure 1 – Comparison between $\eta_L(\alpha)$ and $\eta_{\dagger}(\alpha)$.

Denoting by $\Pi^{b,c,\kappa}$ the wealth process associated with $X^{b,c,\kappa}$ and $\mathcal{E}^{b,c,\kappa}$ the process of tracking error associated with the strategy $X^{b,c,\kappa}$, we have

$$\begin{aligned} \mathcal{E}_t^{b,c,\kappa} &= \Pi_t^\alpha - \Pi_t^{b,c,\kappa} \\ &= \kappa \int_0^t Z_u^\kappa dS_u + \int_0^t S_u c(|Z_u^\kappa|, S_u, u) d\langle X^\alpha \rangle_u + \kappa^2 \int_0^t S_u [dL_u^\kappa + dR_u^\kappa] - \int_0^t \frac{1}{\alpha} \sigma^2 S_u^2 \Gamma_u^\alpha du. \end{aligned}$$

Our first main result in Chapter 2 is the following.

Main Result 2 (Limit distribution of normalized hedging errors). *Under general local volatility model for S and technical conditions for b and c , we have*

$$\kappa^{-1} \left(\mathcal{E}^{b,c,\kappa} - \int_0^t \delta^{b,c}(S_t, t) dt \right) \rightarrow W_{Q^{b,c}},$$

stably in law on $C[0, T]$ as $\kappa \rightarrow 0$, where W is an independent Brownian motion,

$$Q^{b,c} = \int_0^t \eta^{b,c}(S_u, u) d\langle S \rangle_u,$$

and $\delta^{b,c}$, $\eta^{b,c}$ are explicitly determined by b and c .

Fixing the conditional bias of hedging error $\delta^{b,c} = \delta$, it is natural to minimize the conditional variance $Q^{b,c}$. Our second contribution in this chapter is to provide an explicit expression for the infimum

$$Q^{\delta,*} := \operatorname{ess\,inf}_{(b,c) \text{ s.t. } \delta^{b,c} \equiv \delta} Q^{b,c},$$

among all candidate strategies, together with a sequence of strategies (b^*, c^*) attaining asymptotically the infimum $Q^{\delta,*}$. Comparing to the strategy of Leland, the hedging error is significantly reduced. Indeed, it is shown in [DK10] that, following the strategy of Leland,

$$\kappa^{-1}(\Pi^\alpha - \Pi^{\alpha,\kappa}) \rightarrow W_{Q_L}, \quad Q_L = \eta_L(\alpha) \int_0^t |\Gamma_u^\alpha S_u|^2 d\langle S \rangle_u,$$

where

$$\eta_L(\alpha) = \frac{1}{\pi} \alpha^2 + \frac{2}{\pi} \alpha + 1 - \frac{2}{\pi}.$$

On the other hand, taking $\delta = 0$, our optimal strategy (b^*, c^*) satisfies

$$\kappa^{-1}(\Pi^\alpha - \Pi^{b^*,c^*,\kappa}) \rightarrow W_{Q^{0,*}}, \quad Q^{0,*} = \eta_{\dagger}(\alpha) \int_0^t |\Gamma_u^\alpha S_u|^2 d\langle S \rangle_u,$$

where η_{\dagger} has closed-form expression and is much smaller than η_L , see Fig.1.

Part II : Asymptotic optimal tracking

To illustrate the type of problems treated in this part, we consider a government having two means of influencing the foreign exchange rate of its own currency:

1. By choosing the domestic interest rate. A higher interest rate encourages the investors to buy the domestic currency, and as a consequence this currency becomes more valuable and the exchange rate increases.
2. At selected times the government can intervene in the foreign exchange market by buying or selling large amounts of foreign currency. Such intervention is applied only at discrete time and can change the exchange rate instantaneously.

A fluctuating exchange rate is not suitable for the domestic economy due to the uncertainty that it creates. On the other hand, the application of the two policies to stabilize the exchange rate is also costly. The objective of the government is to keep the exchange rate close to a given central parity with minimal costs. See [MØ97, CZ00] for more details.

Similar tracking problems arise naturally in various situations such as the management of an index fund ([PS04, Kor99]), discretization of hedging strategies ([Fuk14, RT14, GL14a]), portfolio selection under transaction costs ([KMK15, ST13, PST15, AMKS15]), trading under market impact ([MMKS14, LMKW14, GW15a, GW15b, GW15c, BSV15]) or illiquidity cost ([RS10, NW11]).

These problems have two common components:

1. A target with stochastic evolution. Usually the target is a benchmark portfolio or index which fluctuates as function of market conditions.
2. A cost structure representing the tracking effort and deviation from the target. In general, greater tracking effort is needed to maintain smaller deviation from the target.

This leads us to formulate the tracking problem as follows.

2.1 Formulation of the tracking problem

We consider a target whose dynamics (X_t°) is modeled by a continuous Itô semi-martingale defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ with values in \mathbb{R}^d such that

$$dX_t^\circ = b_t dt + \sqrt{a_t} dW_t.$$

Here, (W_t) is a d -dimensional Brownian motion and (b_t) , (a_t) are predictable processes with values in \mathbb{R}^d and the set \mathcal{S}_d^+ of $d \times d$ symmetric positive definite matrices respectively. An agent observes X_t° and adjusts her position ψ_t in order to follow X_t° . However, she has to pay certain intervention costs for position adjustments. The objective of the agent is to stay close to the target X_t° while minimizing the tracking efforts. More precisely, let (X_t) be the deviation of the agent from the target (X_t°) , we have

$$X_t = -X_t^\circ + \psi_t. \tag{2.1}$$

Let $H_0(X)$ be a penalty functional for the deviation from the target and $H(\psi)$ the cost incurred by the control process (ψ_t) up to horizon T . Denote by \mathcal{A} the set of admissible strategies depending on the cost structure H_0 and H . Then the problem of tracking can be formulated as

$$\inf_{(\psi_t) \in \mathcal{A}} J(\psi) := H_0(X) + H(\psi). \tag{2.2}$$

Depending on the specific problem under consideration, the control process ψ can be (the combination of) regular control, singular control or impulse control.

To fix idea, we consider in this introduction the case of combined regular and impulse control, which corresponds to the management of exchange rate mentioned above. In that case, a tracking strategy $\psi = (u, \tau, \xi)$ is given by a progressively measurable process $u = (u_t)_{t \geq 0}$ with values in \mathbb{R}^d and $(\tau, \xi) = \{(\tau_j, \xi_j), j \in \mathbb{N}^*\}$, with (τ_j) an increasing sequence of stopping times and (ξ_j) a sequence of \mathcal{F}_{τ_j} -measurable random variables with values in \mathbb{R}^d . The process (u_t) represents the speed of the agent. The stopping time τ_j represents the timing of j th jump toward the target and ξ_j the size of the jump. The tracking error obtained by following the strategy (u, τ, ξ) is given by

$$X_t = -X_t^\circ + \int_0^t u_s ds + \sum_{j:0 < \tau_j \leq t} \xi_j.$$

At any time the agent is paying a cost for maintaining the speed u_t and each jump ξ_j incurs a positive cost. We are interested in the following type of cost functional

$$J(u, \tau, \xi) = \int_0^T (r_t D(X_t) + l_t^\circ Q(u_t)) dt + \sum_{j:0 < \tau_j \leq T} (k_{\tau_j}^\circ F(\xi_j) + h_{\tau_j}^\circ P(\xi_j)),$$

where (r_t) , (l_t°) , (k_t°) and (h_t°) are random weight processes. The cost functions D , Q , F , P are deterministic functions defined for example by

$$D(x) = \langle x, \Sigma^D x \rangle, \quad Q(u) = \langle u, \Sigma^Q u \rangle, \quad F(\xi) = \sum_{i=1}^d F_i \mathbb{1}_{\{\xi^i \neq 0\}}, \quad P(\xi) = \sum_{i=1}^d P_i |\xi^i|,$$

with $F_i, P_i \in \mathbb{R}_+$ such that $\min_i F_i > 0$ and Σ^D, Σ^Q are $d \times d$ positive definite matrices. Note that we have

$$D(\varepsilon x) = \varepsilon^{\zeta_D} D(x), \quad Q(\varepsilon u) = \varepsilon^{\zeta_Q} Q(u), \quad F(\varepsilon \xi) = \varepsilon^{\zeta_F} F(\xi), \quad P(\varepsilon \xi) = \varepsilon^{\zeta_P} P(\xi), \quad (2.3)$$

for any $\varepsilon > 0$ and

$$\zeta_D = 2, \quad \zeta_Q = 2, \quad \zeta_F = 0, \quad \zeta_P = 1.$$

2.2 Asymptotic framework

Our first contribution in this part is to introduce an asymptotic setting for the tracking problem (2.1)-(2.2). In general, the problem rarely admits explicit solution. Under our asymptotic framework of small tracking costs, we are able to establish an asymptotic lower bound for (2.2), which is related to the time average control of Brownian motion.

Assume that there exist $\varepsilon > 0$ and $\beta_Q, \beta_F, \beta_P > 0$ such that

$$l_t^\circ = \varepsilon^{\beta_Q} l_t, \quad k_t^\circ = \varepsilon^{\beta_F} k_t, \quad h_t^\circ = \varepsilon^{\beta_P} h_t. \quad (2.4)$$

Then the asymptotic framework of small tracking costs consists in considering the sequence of optimization problems indexed by $\varepsilon \rightarrow 0$

$$\inf_{(u^\varepsilon, \tau^\varepsilon, \xi^\varepsilon) \in \mathcal{A}^\varepsilon} J^\varepsilon(u^\varepsilon, \tau^\varepsilon, \xi^\varepsilon),$$

with

$$J^\varepsilon(u^\varepsilon, \tau^\varepsilon, \xi^\varepsilon) = \int_0^T (r_t D(X_t^\varepsilon) + \varepsilon^{\beta_Q} l_t Q(u_t^\varepsilon)) dt + \sum_{j:0 < \tau_j^\varepsilon \leq T} (\varepsilon^{\beta_F} k_{\tau_j^\varepsilon} F(\xi_j^\varepsilon) + \varepsilon^{\beta_P} h_{\tau_j^\varepsilon} P(\xi_j^\varepsilon)),$$

and

$$X_t^\varepsilon = -X_t^o + \int_0^t u_s^\varepsilon ds + \sum_{j:0 < \tau_j^\varepsilon \leq t} \xi_j^\varepsilon.$$

The key observation is that under such setting, the tracking problem can be decomposed into a sequence of local problems. More precisely, let $\{t_k^\varepsilon = k\delta^\varepsilon, k = 0, 1, \dots, K^\varepsilon\}$ be a partition of the interval $[0, T]$ with $\delta^\varepsilon \rightarrow 0$ as $\varepsilon \rightarrow 0$. Then we can write

$$J^\varepsilon(u^\varepsilon, \tau^\varepsilon, \xi^\varepsilon) = \sum_{k=0}^{K^\varepsilon-1} j_{t_k^\varepsilon}^\varepsilon(t_{k+1}^\varepsilon - t_k^\varepsilon),$$

with

$$j_t^\varepsilon = \frac{1}{\delta^\varepsilon} \left(\int_{t_k^\varepsilon}^{t_k^\varepsilon + \delta^\varepsilon} (r_t D(X_t^\varepsilon) + \varepsilon^{\beta_Q} l_t Q(u_t^\varepsilon)) dt + \sum_{j:t_k^\varepsilon < \tau_j^\varepsilon \leq t_k^\varepsilon + \delta^\varepsilon} (\varepsilon^{\beta_F} k_{\tau_j^\varepsilon} F(\xi_j^\varepsilon) + \varepsilon^{\beta_P} h_{\tau_j^\varepsilon} P(\xi_j^\varepsilon)) \right)$$

As ε tends to zero, we approximately have

$$J^\varepsilon(u^\varepsilon, \tau^\varepsilon, \xi^\varepsilon) \simeq \int_0^T j_t^\varepsilon dt.$$

Now consider the following rescaling of X^ε over the horizon $(t, t + \delta^\varepsilon]$:

$$\tilde{X}_s^{\varepsilon,t} = \frac{1}{\varepsilon^\beta} X_{t+\varepsilon^\alpha s}^\varepsilon, \quad s \in (0, T^\varepsilon],$$

with $T^\varepsilon = \varepsilon^{-\alpha\beta}\delta^\varepsilon$, where $\alpha = 2$ and $\beta > 0$ is to be determined (here $\alpha = 2$ is related to the scaling property of Brownian motion). On the one hand, we have

$$d\tilde{X}_s^{\varepsilon,t} = \tilde{b}_s^{\varepsilon,t} ds + \sqrt{\tilde{a}_s^{\varepsilon,t}} d\tilde{W}_s^{\varepsilon,t} + \tilde{u}_s^{\varepsilon,t} ds + d\left(\sum_{0 < \tilde{\tau}_j^{\varepsilon,t} \leq s} \tilde{\xi}_j^\varepsilon \right), \quad (2.5)$$

where

$$\begin{aligned} \tilde{b}_s^{\varepsilon,t} &= -\varepsilon^{(\alpha-1)\beta} b_{t+\varepsilon^\alpha s}, & \tilde{a}_s^{\varepsilon,t} &= a_{t+\varepsilon^\alpha s}, & \tilde{W}_s^{\varepsilon,t} &= -\frac{1}{\varepsilon^\beta} W_{t+\varepsilon^\alpha s}, \\ \tilde{u}_s^{\varepsilon,t} &= \varepsilon^{(\alpha-1)\beta} u_{t+\varepsilon^\alpha s}^\varepsilon, & \tilde{\xi}_j^\varepsilon &= \frac{1}{\varepsilon^\beta} \xi_j^\varepsilon, & \tilde{\tau}_j^{\varepsilon,t} &= \frac{1}{\varepsilon^\alpha \beta} (\tau_j^\varepsilon - t) \vee 0. \end{aligned}$$

On the other hand, using the homogeneity properties (2.3) of the cost functions, we obtain

$$\begin{aligned} j_t^\varepsilon &\simeq \frac{1}{T^\varepsilon} \left(\int_0^{T^\varepsilon} (\varepsilon^{\beta\zeta_D} r_t D(\tilde{X}_s^{\varepsilon,t}) + \varepsilon^{\beta_Q - (\alpha-1)\zeta_Q\beta} l_t Q(\tilde{u}_s^{\varepsilon,t})) ds \right. \\ &\quad \left. + \sum_{0 < \tilde{\tau}_j^{\varepsilon,t} \leq T^\varepsilon} (\varepsilon^{\beta_F - (\alpha-\zeta_F)\beta} k_t F(\tilde{\xi}_j^\varepsilon) + \varepsilon^{\beta_P - (\alpha-\zeta_P)\beta} h_t P(\tilde{\xi}_j^\varepsilon)) \right). \end{aligned}$$

The second approximation can be justified by the continuity of cost coefficients r_t , l_t , k_t and h_t .

If there exists $\beta > 0$ such that

$$\beta\zeta_D = \beta_Q - (\alpha-1)\zeta_Q\beta = \beta_F - (\alpha-\zeta_F)\beta = \beta_P - (\alpha-\zeta_P)\beta,$$

or equivalently,

$$\beta = \frac{\beta_F}{\zeta_D + \alpha - \zeta_F} = \frac{\beta_P}{\zeta_D + \alpha - \zeta_P} = \frac{\beta_Q}{\zeta_D + (\alpha-1)\zeta_Q}, \quad (2.6)$$

where $\alpha = 2$, then we have

$$j_t^\varepsilon \simeq \varepsilon^{\beta\zeta_D} I_t^\varepsilon,$$

with

$$I_t^\varepsilon = \frac{1}{T^\varepsilon} \left(\int_0^{T^\varepsilon} (r_t D(\tilde{X}_s^{\varepsilon,t}) + l_t Q(\tilde{u}_s^{\varepsilon,t})) ds + \sum_{0 < \tilde{\tau}_j^{\varepsilon,t} \leq T^\varepsilon} (k_t F(\tilde{\xi}_j^\varepsilon) + h_t P(\tilde{\xi}_j^\varepsilon)) \right). \quad (2.7)$$

It follows that

$$\varepsilon^{-\beta\zeta_D} J^\varepsilon \simeq \int_0^T I_t^\varepsilon dt. \quad (2.8)$$

We are hence led to study I_t^ε as $\varepsilon \rightarrow 0$, which is closely related to the time-average control problem of Brownian motion.

2.3 Main results

Lower bounds

In Chapter 3, we study I_t^ε and establish an asymptotic lower bound for (2.8). By suitably choosing δ^ε , we have $\delta^\varepsilon \rightarrow 0$ and $T^\varepsilon \rightarrow \infty$ as $\varepsilon \rightarrow 0$. Then we have $\tilde{b}_s^{\varepsilon,t} \simeq 0$ and $\tilde{a}_s^{\varepsilon,t} \simeq a_t$ for $s \in (0, T^\varepsilon]$. Therefore, the dynamics of (2.5) is approximately a controlled Brownian motion with fixed diffusion matrix a_t . Hence (2.5) together with (2.7) can be approximately bounded from below by the optimal cost of time-average control problem of a Brownian motion, that is

$$I_t^\varepsilon \gtrsim I(a_t, r_t, l_t, k_t, h_t), \quad (2.9)$$

with the term on the right hand side being defined as

$$I(a, r, l, k, h) = \inf_{(u, \tau, \xi)} \limsup_{S \rightarrow \infty} \frac{1}{S} \mathbb{E} \left[\int_0^S (r D(X_s) + l Q(u_s)) ds + \sum_{0 \leq \tau_j \leq S} (k F(\xi_j) + h P(\xi)) \right], \quad (2.10)$$

where

$$dX_s = \sqrt{a} dW_s + u_s ds + d \left(\sum_{0 < \tau_j \leq s} \xi_j \right). \quad (2.11)$$

Consequently, we obtain

$$\varepsilon^{-\beta\zeta_D} J^\varepsilon \simeq \int_0^T I_t^\varepsilon dt \gtrsim \int_0^T I_t dt,$$

as $\varepsilon \rightarrow 0$. The main contribution in Chapter 3 is to formulate rigorously the above result as the following.

Main Result 3 (Lower bound). *There exists β explicitly determined by (2.6) such that, for all $\delta > 0$ and any sequence of admissible strategies $\{\psi^\varepsilon \in \mathcal{A}^\varepsilon, \varepsilon > 0\}$, we have*

$$\lim_{\varepsilon \rightarrow 0^+} \mathbb{P} \left[\frac{1}{\varepsilon^{\beta\zeta_D}} J^\varepsilon(\psi^\varepsilon) \geq \int_0^T I_t dt - \delta \right] = 1, \quad (2.12)$$

where $I_t = I(a_t, r_t, k_t, h_t, l_t)$ is essentially the optimal cost of time-average control of Brownian motion (2.10)-(2.11) with parameters frozen at time t .

Various versions of Main Result 3 under different cost structures are also presented in Chapter 3, together with a comprehensive list of explicit examples for I .

A key step in the proof of Main Result 3 is to justify rigorously (2.9). More precisely, we show that

$$\liminf_{\varepsilon \rightarrow 0} \mathbb{P}[I_t^\varepsilon \geq I_t - \delta] = 1, \quad \forall \delta > 0, \quad (2.13)$$

where I_t^ε is given by (2.5)-(2.7) and $I_t = I(a_t, r_t, l_t, k_t, h_t)$ is given by (2.10)-(2.11). Inspired by [KM93] and [KS99], we apply weak convergence method on the empirical occupational measures and express the lower bound I_t as the solution of an infinite dimensional linear programming on a suitable space of measures. Such characterization is essentially equivalent to (2.10)-(2.11) if we formulate the controlled Brownian motion through a controlled martingale problem (see [KS01]).

Feedback strategies

Our goal in Chapter 4 is to build strategies attaining the lower bounds (2.12). Following classical approaches, we are interested in a class of feedback strategies which consists in keeping the deviation X_t inside a time-dependent domain. Consider for example the case of combined regular and impulse control. Let (G_t) be a moving open bounded domain associated with jump rule (ξ_t) from ∂G_t to G_t , and (u_t) be a continuous function from \bar{G}_t to \mathbb{R}^d . The sequence of feed-back strategies $(X^\varepsilon, u^\varepsilon, \tau^\varepsilon, \xi^\varepsilon)$ corresponding to the triplet (u_t, G_t, ξ_t) can be constructed in the following recursive way :

1. Let $\tau_0^\varepsilon = 0, X_0^\varepsilon = 0$.
2. For $t \geq \tau_{j-1}^\varepsilon$, let X_t^ε be defined by

$$dX_t^\varepsilon = -dX_t^o + u_t^\varepsilon dt,$$

with

$$u_t^\varepsilon = \varepsilon^{-(\alpha-1)\beta} u_t(\varepsilon^{-\beta} X_t^\varepsilon).$$

3. Set

$$\tau_j^\varepsilon = \inf\{t > \tau_{j-1}^\varepsilon, \frac{1}{\varepsilon^\beta} X_t^\varepsilon \notin G_t\}, \quad \xi_j^\varepsilon = \varepsilon^\beta \xi_{\tau_j^\varepsilon}(\varepsilon^{-\beta} X_{\tau_j^\varepsilon}^\varepsilon),$$

and

$$X_{\tau_j^\varepsilon}^\varepsilon = X_{\tau_j^\varepsilon-}^\varepsilon + \xi_j^\varepsilon.$$

We now give the main result for combined regular and impulse control in Chapter 4. The cases involving singular control are also studied in same chapter.

Main Result 4 (Feedback strategies). *Let $\{(X^\varepsilon, u^\varepsilon, \tau^\varepsilon, \xi^\varepsilon), \varepsilon > 0\}$ be the feedback strategy determined by an admissible triplet (u_t, G_t, ξ_t) , then we have*

$$\frac{1}{\varepsilon^{\beta\zeta_D}} J^\varepsilon(u^\varepsilon, \tau^\varepsilon, \xi^\varepsilon) \rightarrow_p \int_0^T c(u_t, G_t, \xi_t) dt,$$

where $c(u, G, \xi)$ can be explicitly determined via the stationary measures of Brownian motion in the domain G with drift u and jump rule ξ from the boundary.

For a wide range of examples, we show that there exists explicit triplet (u_t^*, G_t^*, ξ_t^*) verifying $c(u_t^*, G_t^*, \xi_t^*) = I_t$. Hence, the associated feedback strategy $\{(X^{\varepsilon,*}, u^{\varepsilon,*}, \tau^{\varepsilon,*}, \xi^{\varepsilon,*}), \varepsilon > 0\}$ satisfies

$$\frac{1}{\varepsilon^{\beta\zeta_D}} J^\varepsilon(u^{\varepsilon,*}, \tau^{\varepsilon,*}, \xi^{\varepsilon,*}) \rightarrow_p \int_0^T I_t dt.$$

In other words, the asymptotic lower bound in Main Result 3 is tight.

2.4 Relation with other asymptotic studies

Our final contribution in Chapter 4 is to establish a link among different asymptotic analysis in the literature on various topics such as optimal discretization of hedging strategies, discretization error of stochastic integrals and impact of small market frictions.

Main Result 3 and 4 enable us to revisit the asymptotic lower bounds for the discretization of hedging strategies in [Fuk11a, GL14a]. In these papers, the lower bounds are deduced by using subtle inequalities. We show that these bounds can be simply interpreted through the time average control problem of Brownian motion.

As a corollary of Main Result 4, we establish a weak convergence theorem for discretization errors of stochastic integrals with random stopping times. The limit law of the normalized discretization errors turns out to be a mixture of Gaussian distributions, with the conditional mean and variance being expressed as the first and second moments of the stationary measures of Brownian motion in a domain with jumps from the boundary. Our approach suggests that the results in [HM05, Fuk11c, Roo80, LR13] can be interpreted in a similar way.

The lower bound (2.12) appears also in the study of small market frictions under the framework of utility maximization. Indeed, we observe that utility maximization under small market frictions is heuristically equivalent to the tracking problem with the target being the optimal strategy under frictionless market. Consider, for example, the optimization of terminal wealth given by

$$u(w_0) = \sup_{\varphi} \mathbb{E}[U(w_T^{\varphi})],$$

where w_0 is the initial wealth and w_T^{φ} the terminal wealth following the trading strategy φ . In a market with proportional transaction costs, the portfolio dynamics is given by

$$w_t^{\varepsilon} = w_0 + \int_0^t \varphi_u^{\varepsilon} dS_u - \int_0^t \varepsilon h_u d\|\varphi^{\varepsilon}\|_u,$$

where S_t is the risky asset price, εh_t is a random weight process representing transaction costs, and φ_t^{ε} is a trading strategy with finite variation. Assume that φ_t^* is the optimal trading strategy in the frictionless market ($\varepsilon = 0$), and denote the equivalent martingale measure by \mathbb{Q} and the indirect risk-tolerance process by R_t . When ε is small, we can expect that φ_t^{ε} is close to φ_t^* . Then up to first order quantities, we have (see also [KMK15, KL13, Rog04])

$$\mathbb{E}[U(w_T^{\varepsilon})] - u(w_0) \simeq -u'(w_0) \mathbb{E}^{\mathbb{Q}} \left[\varepsilon \int_0^T h_t d\|\varphi^{\varepsilon}\|_t + \int_0^T \frac{a_t^S}{2R_t} (\varphi_t^{\varepsilon} - \varphi_t^*)^2 dt \right],$$

where a_t^S is the quadratic variation of the risky asset S_t . Hence the optimization of terminal wealth under small proportional costs can be reduced to a tracking problem with stochastic target φ_t^* and deviation penalty

$$r_t D(x) := \frac{a_t^S}{2R_t} x^2.$$

Denoting the value function under transaction costs by u^{ε} , and defining the certainty equivalent wealth loss Δ^{ε} by

$$u^{\varepsilon} =: u(w_0 - \Delta^{\varepsilon}),$$

then we have

$$\frac{1}{\varepsilon^{\beta} \zeta_D} \Delta^{\varepsilon} \simeq \mathbb{E}^{\mathbb{Q}} \left[\int_0^T I_t dt \right].$$

Similar correspondences can also be established for the cases with different cost structures ([AMKS15, MMKS14]), the indifference pricing of option ([WW97, WW99, KMK13]), the maximization of long term growth rate ([AW95, APW97, GW15a, GW15b, GW15c, LMKW14]), and optimization of consumption ([ST13, PST15]).

Part III : Portfolio selection with capital gains taxes

In contrast to transaction costs, the problem of portfolio selection under capital gains taxes received relatively limited attention. Capital gain taxes differ from transaction costs in the following aspects:

1. Investors pay taxes for capital gains but receive tax rebates for capital losses.
2. The amount of capital gains or losses taxed depends on the purchase price of stock holdings, known as the tax basis, which incurs strong path-dependency.

As a consequence, much of the existing literature on capital gain taxes has been restricted to discrete-time models with small number of time steps, see [Con83, Con84, DK96, DU05, GH06].

Using the average purchase price of stocks as an approximation for tax basis, [DSZ01, DSZ03] develop a binomial tree model that is able to effectively work with multi-step investment and consumption decisions. The advantage of the approximation is that the path dependency of the problem is considerably reduced, as the dynamics of the tax basis becomes Markovian. [GKT06] further extend the model to the multiple stocks case. In [BST10], the authors formulate a continuous-time version of the model introduced by [DSZ01].

In this part, we are interested in the joint impact of capital gains taxes with other market features such as regime-switching and transaction costs. Our work is mainly based on extensions of the models of [BST10, BST07]. We point out that the goal of this part is to provide a deep understanding of the optimal strategy and related probabilistic interpretation, although the main results remain to be rigorously proved.

3.1 Preliminary : the model of [BST10]

Let us describe the model of [BST10] in more details. Consider a financial market with two assets that the investor can trade without any transaction costs. The first asset is a bond with pre-tax interest rate r . The second asset is a risky stock whose price (S_t) evolves according to the Black-Scholes model:

$$dS_t = S_t(\mu dt + \sigma dW_t).$$

The investor is subject to capital gain taxes. The tax basis (B_t) used to evaluate capital gains is defined as the weighted average of past purchase prices. The amount of tax to be paid for each sale of risky asset is given by

$$\alpha(S_t - B_t),$$

where $\alpha \in [0, 1)$. When $S_t \geq B_t$, i.e. the current price of the risky asset is greater than the tax basis, the investor realizes a capital gain by selling the risky asset. When $S_t < B_t$, the sale of the risky asset corresponds to the realization of a capital loss.

Let X_t , Y_t , and K_t be the amount invested in the bond, the current dollar value of, and the cumulated purchase price of stock holdings, respectively. We introduce two càdlàg, non-negative, and non-decreasing \mathcal{F}_t -adapted processes L_t and M_t with $L_{0-} = M_{0-} = 0$, where dL_t represents the *dollar amount* transferred from the bank to the stock account at time t (corresponding to a purchase of stock), while dM_t represents the *proportion of shares* transferred from the stock account to the bank at time t (corresponding to a sale of stock). We assume the no-short-sales constraint such that $dM_t \leq 1$. Note that the cumulated purchase price of stock holding K_t is related to the tax basis B_t by

$$K_t = B_t \frac{Y_t}{S_t}.$$

Hence, when one sells stock at time t , the cumulated purchase price K_t declines by the same proportion dM_t as the dollar value of stock holdings does, and the realized capital gain is $(Y_{t-} - K_{t-})dM_t$. Then, the evolution processes of X_t, Y_t , and K_t are

$$\begin{aligned} dX_t &= [(1 - \alpha)rX_t - C_t]dt - dL_t + [Y_{t-} - \alpha(Y_{t-} - K_{t-})]dM_t, \\ dY_t &= Y_t(\mu dt + \sigma dW_t) + dL_t - Y_{t-}dM_t, \\ dK_t &= dL_t - K_{t-}dM_t, \end{aligned}$$

where (C_t) is the consumption stream.

The investor aims to find

$$\varphi(x, y, k) := \sup_{(C_t, L_t, M_t)} \mathbb{E} \left[\int_0^\infty e^{-\beta t} U(C_t, \gamma) dt \mid X_0 = x, Y_0 = y, K_0 = k \right], \quad (3.1)$$

where $\beta > 0$ is a constant discount factor and $U(\cdot, \gamma)$ is a power utility function with parameter γ .

Since the value function (3.1) admits no closed-form solution, [BST10] provide instead analytical upper and lower bounds. First, if there is no capital gains taxes ($\alpha = 0$), then the above problem reduces to the classical tax-free Merton problem. Denote the value function of the Merton problem by $\bar{\varphi}$ and the optimal consumption-investment strategy by $(\bar{c}, \bar{\xi})$. Second, consider the following tax-free model with parameters

$$\mu_\alpha = (1 - \alpha)\mu, \quad \sigma_\alpha = (1 - \alpha)\sigma, \quad r_\alpha = (1 - \alpha)r. \quad (3.2)$$

This is the so called tax-deflated model. Denote the value function of the corresponding Merton problem by $\bar{\varphi}_\alpha$ and the optimal consumption-investment strategy by $(\bar{c}_\alpha, \bar{\xi}_\alpha)$. Then it is shown in [BST10, Propostion 4.1, 4.2] that,

$$\bar{\varphi}_\alpha(z) \leq \varphi(x, y, k) \leq \bar{\varphi}(z), \quad (3.3)$$

where z is the liquidation wealth $z = x + y - \alpha(y - k)$.

While the upper bound is natural in that the investor cannot take advantage of tax rebates to do better than in a tax-free market, [BST10] provide an insight for the lower bound. More precisely, define the portfolio value after liquidation Z_t by

$$Z_t = X_t + Y_t - \alpha(Y_t - K_t).$$

Then we have

$$\begin{aligned} dZ_t &= dX_t + (1 - \alpha)dY_t + \alpha dK_t \\ &= [(1 - \alpha)rZ_t - C_t]dt + (1 - \alpha)Y_t(\mu dt + \sigma dW_t - (1 - \alpha)r dt) - (1 - \alpha)\alpha r K_t dt. \end{aligned}$$

Taking

$$\xi_t = \frac{Y_t}{Z_t}, \quad b_t = \frac{K_t}{Y_t}, \quad c_t = \frac{C_t}{Z_t}, \quad (3.4)$$

we obtain

$$dZ_t = Z_t[(r_\alpha - c_t)dt + \xi_t(\mu_\alpha dt + \sigma_\alpha dW_t - r_\alpha dt) + \alpha r_\alpha \xi_t(1 - b_t)dt]. \quad (3.5)$$

The process $b_t \geq 0$ is called the relative tax basis. Note that $b_t > 1$ corresponds to capital losses while $b_t < 1$ corresponds to capital gains. By constructing a sequence of strategies which keeps asymptotically $b_t \equiv 1$, [BST10] show that Z_t can be approximated by

$$dZ_t = Z_t[(r_\alpha - c_t)dt + \xi_t(\mu_\alpha dt + \sigma_\alpha dW_t - r_\alpha dt)],$$

which is exactly the wealth process under the tax-deflated model (3.2). The lower bound follows from the possibility to replicate (asymptotically) any consumption stream which is admissible for the tax-deflated model under the model of [BST10].

3.2 Expansion around tax-deflated model

Our first contribution in Chapter 5 is to improve the bounds in (3.3) by providing a first order correction in terms of a time-average control problem. The sub-optimal strategy $b_1 \equiv 1$ proposed by [BST10] consist in realizing both capital losses and gains immediately. We observe that keeping $b_t \equiv 1$ is apparently not the optimal choice. Indeed, as long as $1 - b_t \geq 0$, the last term in (3.5) is in the favor of the investor. Therefore, it is better for the investor to defer capital gains and keep $b_t \leq 1$. Meanwhile, such deferral will drive the portfolio away from the benchmark $\bar{\xi}_\alpha$ under the tax-deflated model. Hence the optimal strategy should be a delicate balancing between capital gains deferral and the utility loss due to deviation from the benchmark optimal strategy $\bar{\xi}_\alpha$ under the tax-deflated model.

The main idea is to consider the asymptotic setting where the interest rate or tax rate are small. Then the extra benefit from tax deferral $\alpha r_\alpha \xi_t (1 - b_t)$ is small and $\xi_t - \bar{\xi}_\alpha$ should also be small. Using heuristic Taylor expansion as in [Rog04], we obtain

$$\begin{aligned} & \mathbb{E} \left[\int_0^\infty e^{-\beta t} U(C_t, \gamma) dt \right] \\ & \simeq \bar{\varphi}_\alpha(z) + (1 - \gamma) U(\bar{c}_\alpha z, \gamma) \int_0^\infty e^{-\bar{c}_\alpha t} \mathbb{E}^{\mathbb{Q}_\alpha} \left[\int_0^t (\alpha r_\alpha \xi_s (1 - b_s) - \frac{\gamma \sigma_\alpha^2}{2} (\xi_s - \bar{\xi}_\alpha)^2) ds \right] dt. \end{aligned}$$

where \mathbb{Q}_α is a suitable change of probability.

In order to obtain extra welfare by tax-deferral, the investor needs essentially to maximize

$$\mathbb{E}^{\mathbb{Q}_\alpha} \left[\int_0^t (\alpha r_\alpha \xi_s (1 - b_s) - \frac{\gamma \sigma_\alpha^2}{2} (\xi_s - \bar{\xi}_\alpha)^2) ds \right], \quad (3.6)$$

where the first term represents the benefit of tax-deferral and the second the equivalent wealth loss due to deviation from the benchmark strategy $\bar{\xi}_\alpha$. Applying Itô formula to (3.4), it is not difficult to deduce the dynamics of ξ_t and b_t :

$$\begin{aligned} db_t & \simeq -\sigma b_t dW_t + \frac{1 - b_t}{\bar{\xi}} dD_t + dR_t, \\ d(1 - \alpha)\xi_t & \simeq \sigma \bar{\xi} (1 - \bar{\xi}) dW_t + dD_t - dU_t, \end{aligned} \quad (3.7)$$

where D_t, U_t and R_t are processes with finite variation determined by L_t and M_t .

An interesting property of the coupled system (3.7) is that

$$\xi_t - \bar{\xi}_\alpha = O(\delta^2) \Leftrightarrow 1 - b_t = O(\delta).$$

In other words, $1 - b_t$ and $\xi_t - \bar{\xi}_\alpha$ have different time scales. Defining

$$p_t = \frac{1 - b_t}{\delta}, \quad q_t = \frac{\xi_t - \bar{\xi}_\alpha}{\delta^2},$$

where

$$\delta = (A\varepsilon)^{2/3}, \quad \varepsilon = \sqrt{\frac{2\alpha r_\alpha \bar{\xi}_\alpha}{\gamma \sigma^2}}, \quad A = \frac{4}{[3\bar{\xi}(\bar{\xi} - 1)]^2},$$

and applying similar arguments as in the tracking problem, we are led to consider the following time-average control problem in which the fast variable q_t disappears:

$$I^\delta = \sup_{(l_t)} \liminf_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left[\int_0^T (p_t - \frac{1}{3} A^2 l_t^2) dt \right], \quad (3.8)$$

and the dynamics of p_t becomes

$$dp_t = \sigma(1 - \delta p_t)dW_t - \frac{\sigma^2}{3Al_t}p_t dt + dR_t, \quad (3.9)$$

where l_t is an adapted positive process and R_t a non-decreasing process representing the reflection of p_t at $p = 0$. Our first contribution in Chapter 5 is the following result, which is only heuristically proved.

Main Result 5. *Taking $z = x + y - \alpha(y - k)$ and defining the value of tax deferral w by*

$$\varphi(x, y, k) = \bar{\varphi}_\alpha(z e^w),$$

then we have

$$w = \frac{\gamma}{\bar{c}_\alpha} \frac{\sigma^2}{2} I^\delta A^{2/3} \varepsilon^{8/3} + o(\varepsilon^{8/3}),$$

where I^δ is given by the optimal cost of the time-average control problem (3.8)-(3.9).

Compare to the case of utility maximization with market frictions, we have identified a new local probabilistic model (3.8)-(3.9) in the first order expansion of the value function.

3.3 Capital gains taxes with recursive utility and regime-switching

Our second contribution in Chapter 5 is to study the impact of capital gains taxes in a regime-switching market for an investor with recursive utility of Epstein-Zin type. We not only provide asymptotic analysis based on the intuition developed in the previous section, but also perform extensive numerical study to validate our asymptotic analysis.

Let φ_i be the value function under regime $i \in \mathcal{I}$. As before, we introduce the tax-deflated regime-switching model for which the market parameters $r_{\alpha,i}$, $\mu_{\alpha,i}$ and $\sigma_{\alpha,i}$ are given by

$$r_{\alpha,i} = (1 - \alpha)r_i, \quad \mu_{\alpha,i} = (1 - \alpha)\mu_i, \quad \sigma_{\alpha,i} = (1 - \alpha)\sigma_i, \quad \forall i \in \mathcal{I}.$$

We denote the value function of tax-free problem under regime i by $\bar{\varphi}_{\alpha,i}$ and the optimal strategy by $(\bar{c}_{\alpha,i}, \bar{\xi}_{\alpha,i})$. The main result of Chapter 5 is the following heuristic expansion of value function.

Main Result 6. *Defining the value of deferral w_i under regime i by*

$$\varphi_i(x, y, k) = \bar{\varphi}_{\alpha,i}(z e^{w_i}),$$

then we have

$$w_i = \frac{\gamma}{\bar{c}_{\alpha,i}} \frac{\sigma_i^2}{2} m_i A_i^{2/3} \varepsilon_i^{8/3} + o(\varepsilon_i^{8/3}),$$

where $\{m_i, i \in \mathcal{I}\}$ are explicitly determined in term of $\{I^{\delta_i}, i \in \mathcal{I}\}$ via a linear system.

The key step to obtain the above expansion is first suggested in [CD13], where the authors use the following “fast variables”

$$p_i = \frac{1 - k/y}{\delta_i}, \quad q_i = \frac{y/z - \bar{\xi}_{\alpha,i}}{\delta_i^2},$$

and postulate in the associated HJB equation that

$$w_i = \gamma \varepsilon_i^2 \left(\delta_i \frac{\sigma_i^2}{2} \frac{1}{\bar{c}_{\alpha,i}} m_i + \delta_i^3 g_i^\delta(p_i) + \delta_i^5 v_i^\delta(p_i, q_i) \right).$$

This asymptotic setting is now fully supported by our probabilistic analysis in Section 3.2.

The asymptotic expansion allows us to perform fruitful economic analysis of the impact of capital gains taxes. In particular, we obtain the impact of capital gains taxes in function of the elasticity of inter-temporal substitution (EIS) and regime transition intensities, together with explicit trading boundaries. Moreover, these asymptotic analysis are all validated by a direct numerical computation of the associated HJB equation based on penalty method.

3.4 Joint impact of capital gains taxes and transaction costs

Our goal in Chapter 6 is to derive a system of corrector equations for the value function of portfolio selection problem under proportional transaction costs and capital gains taxes, extending both [ST13] and [CD13]. Models with both transaction costs and capital gains taxes have also been studied in [CP99, BCP05, Lel99] but under quite different settings.

For simplicity, we use the model of [BST07] and keep the same notation as in Section 3.1. Given any strategy (C_t, L_t, M_t) , the dynamics of Y_t and K_t remain the same and we have

$$dX_t = ((1 - \alpha)rX_t - C_t)dt - (1 + \lambda_B)dL_t + (1 - \lambda_S)[(1 - \alpha)Y_{t-} + \alpha K_{t-}]dM_t,$$

where $\lambda_B, \lambda_S \in [0, 1)$ represent the buy/sell costs and $\alpha \in [0, 1)$ is the tax rate. In the asymptotic setting of small transaction costs and interest rate, we replace the transaction cost coefficients λ_B, λ_S and the interest rate r by

$$\lambda_B^\delta = \lambda_B \delta^6, \quad \lambda_S^\delta = \lambda_S \delta^6, \quad r^\delta = r \delta^3, \quad \delta > 0.$$

and denote the corresponding value function by $\varphi^\delta(x, y, k)$. The associated tax-deflated model is defined by

$$\mu_\alpha = (1 - \alpha)\mu, \quad \sigma_\alpha = (1 - \alpha)\sigma, \quad r_\alpha^\delta = (1 - \alpha)r^\delta.$$

Denote the value function of Merton problem under the tax-deflated model (without transaction costs and taxes) by $\bar{\varphi}_\alpha^\delta(z)$, and the corresponding optimal consumption-investment strategy by $(\bar{C}_\alpha^\delta(z), \bar{y}_\alpha^\delta(z))$. The main contribution of Chapter 6 is the formal derivation of the following corrector equations.

Main Result 7. *Taking $z = x + y - \alpha(y - k)$, we have*

$$\varphi^\delta(x, y, k) = \bar{\varphi}_\alpha^\delta(z) + u_\alpha^\delta(z)\delta^4 + o(\delta^4),$$

where u_α^δ is given by

$$-\frac{1}{2}\sigma_\alpha^2(\bar{y}_\alpha^\delta(z))^2\partial_{zz}^2 u_\alpha^\delta(z) - [(\mu_\alpha - r_\alpha^\delta)\bar{y}_\alpha^\delta(z) + r_\alpha^\delta z - \bar{C}_\alpha^\delta(z)]\partial_z u_\alpha^\delta(z) + \beta u_\alpha^\delta(z) - a(z) = 0.$$

The constant $a(z)$ (depending on z) is determined by

$$\min \left\{ -\frac{1}{2}\sigma^2(1 - \delta p)^2\partial_{pp}^2 G_\alpha^\delta(z, p) - (\alpha r_\alpha \partial_z \bar{\varphi}_\alpha^\delta(z)\bar{y}_\alpha^\delta(z)p - b(z, p)) + a(z), \right. \\ \left. G_\alpha^\delta(z, p) + (\lambda_S + \lambda_B)\partial_z \bar{\varphi}_\alpha^\delta(z)\bar{y}_\alpha^\delta(z) - G_\alpha^\delta(z, 0) \right\} = 0,$$

and

$$\min \left\{ \frac{1}{2}\sigma^2(\bar{y}_\alpha^\delta(z))^2[1 - (1 - \alpha)\partial_z \bar{y}_\alpha^\delta(z)]^2\partial_{qq}^2 w_\alpha^\delta(z, p, q) + \frac{1}{2}\sigma_\alpha^2(-\partial_{zz}^2 \bar{\varphi}_\alpha^\delta(z))q^2 - b(z, p), \right. \\ \left. \frac{p}{\bar{y}_\alpha^\delta(z)}\partial_p G_\alpha^\delta(z, p) + \lambda_B \partial_z \bar{\varphi}_\alpha^\delta(z) - \partial_q w_\alpha^\delta(z, p, q), \right. \\ \left. -\lambda_S \partial_z \bar{\varphi}_\alpha^\delta(z) - \partial_q w_\alpha^\delta(z, p, q) \right\} = 0.$$

In particular, we recover the corrector equations in [ST13, Definition 3.1] if $\alpha = 0$, and [CD13, Equation (A.13)] if $\lambda_B = \lambda_S = 0$.

In [ST13, Remark 3.3], the equation for $w_\alpha^\delta(q)$ is represented by the time-average control of Brownian motion with proportional costs. We find that the equation for $G_\alpha^\delta(p)$ is also closely related to a time-average control problem like (3.8)-(3.9) and $a(z)$ can be interpreted as the corresponding optimal cost I^δ .

Part I

**OPTION PRICING AND
HEDGING**

Chapter 1

Optimal dicretization of hedging strategies with directional views

Contents

1.1	Introduction	23
1.2	Assumptions and admissible strategies	26
1.2.1	Assumptions on the dynamics and admissibility conditions	26
1.2.2	Comments on the admissibility conditions	28
1.2.3	Examples of admissible discretization rules	29
1.3	Asymptotic optimality: a preliminary approach	31
1.4	Asymptotic expectation-error optimization	33
1.5	Black-Scholes model with time-varying coefficients	34
1.5.1	Explicit formulas	35
1.5.2	Numerical study	37
Appendix 1.A	Proofs	39
1.A.1	Proof of Proposition 1.2.1	39
1.A.2	Proof of Proposition 1.2.2	44
1.A.3	Proof of Proposition 1.2.3	48
1.A.4	Proof of Theorem 1.4.1	52
Appendix 1.B	Linear-quadratic optimal control	53

1.1 Introduction

In order to manage the risks inherent to the derivatives they buy and sell, practitioners use continuous time stochastic models to compute their prices and hedging portfolios. In the simplest cases, notably in that of the so-called delta hedging strategy, the hedging portfolio obtained from the model is a time-varying self-financed combination of cash and the underlying. We denote the price at time t of the underlying asset by Y_t and assume it to be a one-dimensional semi-martingale. Hence, in such situations, the outputs of the model are the price of the option together with the number of shares in the underlying asset to be held in the hedging portfolio at any time t , denoted by X_t . The proportion invested in cash is then deduced from the self-financing property. Therefore, assuming zero interest rates, the theoretical value of the model-based hedging portfolio at the maturity of the option T is given by

$$\int_0^T X_t dY_t.$$

Typically, the process X_t derived from the model is a continuously varying semi-martingale, requiring continuous trading to be implemented in practice. This is of course physically impossible and would be anyway irrelevant because of the costs induced by microstructure effects. Hence practitioners do not use the strategy X_t , but rather a discretized version of it. This means that the hedging portfolio is only rebalanced a finite number of times and is held constant between these times. Let us denote by $(\tau_j^n)_{j \geq 0}$ an increasing sequence of rebalancing times over $[0, T]$ (the meaning of the parameter n will be explained below). With respect to the target portfolio obtained by continuous rebalancing, the hedging error due to discrete trading Z_T^n is therefore given by

$$Z_T^n = \sum_{j=0}^{+\infty} X_{\tau_j^n} (Y_{\tau_{j+1}^n \wedge T} - Y_{\tau_j^n \wedge T}) - \int_0^T X_t dY_t.$$

Thus, some important questions in practice are:

- What is the order of Z_T^n for commonly used discretization strategies ?
- For a given criterion, how to optimize the rebalancing times ?

The most widely studied rebalancing scheme is that of equidistant trading dates of the form

$$\tau_j^n = jT/n, \quad j = 0, \dots, n,$$

where n represents the total number of trades during the time period $[0, T]$. In this setting, the first question has been addressed in detail. There are two popular approaches to quantify the hedging error Z_T^n , both of them being asymptotic, assuming the rebalancing frequency n/T tends to infinity (that is n tends to infinity since T is fixed). A first possibility is to use the L^2 -norm, where one typically looks for asymptotic bounds of the form

$$\mathbb{E}[(Z_T^n)^2] \leq cn^{-\theta}, \quad n \rightarrow \infty.$$

Many authors have explored various aspects of this problem in this deterministic rebalancing dates framework. For European call and put options in the Black-Scholes model, it is shown in [BKL00] and [Zha99] that the L^2 -error has a convergence rate $\theta = 1$. For other options, the convergence rate depends on the regularity of the payoff. For example, it is shown in [GT01] that for binary options, the convergence rate is $\theta = 1/2$. However, in this context, the convergence rate $\theta = 1$ can be recovered by choosing a suitable non equidistant deterministic rebalancing grid, see [Gei02]. An asymptotic lower bound for the L^2 -error is given in [Fuk11a, Fuk14] for a general class of rebalancing schemes.

The second way to assess the hedging error is through the weak convergence of the sequence of the suitably rescaled random variables Z_T^n . When X and Y are Itô processes, the case of equidistant rebalancing dates has been investigated using this approach in [BKL00, HM05, Roo80]. In these references, with varying degree of generality, the following convergence in law is proved:

$$\sqrt{n}Z_T^n \xrightarrow{\mathcal{L}} \sqrt{\frac{T}{2}} \int_0^T \sigma_t^X \sigma_t^Y dB_t, \quad (1.1.1)$$

where σ^X and σ^Y are the volatilities of X and Y and B is a Brownian motion independent of the other quantities. The case where X and Y are processes with jumps is treated in [TV09], and [GT09] allow for non-equidistant time nets.

This asymptotic approach has also been recently used in the context where the rebalancing times are random stopping times. Some specific hitting times based schemes derived from a microstructure model are investigated in [RR10]. In [Fuk11c], the author works with quite general sampling schemes based on stopping times. More precisely, for a given parameter n driving the asymptotic, one considers an increasing sequence of stopping times

$$0 = \tau_0^n < \tau_1^n < \dots < \tau_j^n < \dots$$

with $N_t^n := \max\{j \geq 0; \tau_j^n \leq t\} < \infty$ almost surely for any $t \geq 0$. When

$$\frac{\mathbb{E}[(X_{\tau_{j+1}^n} - X_{\tau_j^n})^4 | \mathcal{F}_{\tau_j^n}]}{\mathbb{E}[(X_{\tau_{j+1}^n} - X_{\tau_j^n})^2 | \mathcal{F}_{\tau_j^n}]} = \varepsilon_n^2 a_{\tau_j^n}^2 + o_p(\varepsilon_n^2), \quad \frac{\mathbb{E}[(X_{\tau_{j+1}^n} - X_{\tau_j^n})^3 | \mathcal{F}_{\tau_j^n}]}{\mathbb{E}[(X_{\tau_{j+1}^n} - X_{\tau_j^n})^2 | \mathcal{F}_{\tau_j^n}]} = -\varepsilon_n s_{\tau_j^n} + o_p(\varepsilon_n)$$

for a sequence $\varepsilon_n \rightarrow 0$ and left continuous adapted processes a and s , under some additional regularity conditions, it is shown in [Fuk11c] that $\varepsilon_n^{-1} Z^n$ converges weakly to

$$\frac{1}{3} \int_0^\cdot s_t dY_t + \frac{1}{\sqrt{6}} \int_0^\cdot (a_t^2 - \frac{2}{3} s_t^2)^{1/2} \sigma_t^Y dB_t, \quad (1.1.2)$$

where B is a Brownian motion independent of all the other quantities. Further,

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\varepsilon_n^2 N_t^n < \int_0^t \frac{(\sigma_u^X)^2}{a_u^2 - s_u^2} du - \delta \right) = 0 \quad (1.1.3)$$

for all $\delta > 0$ and $t \geq 0$.

One can remark a crucial difference between the deterministic discretization schemes associated to (1.1.1) and the random stopping times case leading to (1.1.2). For deterministic dates, the discretization error asymptotically behaves as a stochastic integral with respect to Brownian motion. Therefore, it is (essentially) centered. In the case of random discretization dates, one may obtain a “biased” asymptotic hedging error because of the presence of the term

$$\int_0^T s_t dY_t.$$

Hence, if s does not vanish and Y has non zero drift, the asymptotic hedging error is no longer centered.

From a practitioner viewpoint, this is quite an interesting property. Indeed, it shows that in the presence of market trends, the trader, whose principal goal is to hedge the option, may actually be compensated for the extra risk arising from discrete trading, provided that the rebalancing dates are chosen in an appropriate way. Intuitively, the idea is to wait longer before rebalancing the portfolio if the market is moving in a direction which is favorable for the trader. Of course one may say this is not the option trader’s job to try to get a positive expected return with the hedging strategy. However, knowing that there is anyhow a hedging error, it seems reasonable to optimize it to the trader’s benefit.

Hence, we place ourselves in the asymptotic high frequency regime where n is large and therefore

$$\sup_j (\tau_{j+1}^n - \tau_j^n)$$

is small, meaning that the hedging error should be small. In this setting we address the second question raised above, that is finding the optimal times to rebalance the portfolio. To do so, we simply use an asymptotic expectation-error type criterion. More precisely, we wish to maximize the expectation of the hedging error under a constraint on its L^2 -norm. This is somewhat related to [Sep13], where the author aims at finding an optimal hedging frequency to optimize the Sharpe ratio. However, in this reference, the market trend is assumed to be zero and the hedging error is not centered due to the presence of transaction costs. Observe that in our context, the L^2 -norm is more meaningful than the variance since the primary goal of the trader is to make the hedging error small. Our asymptotic approach goes as follows. First, we approximate the law of the renormalized hedging error by that in Equation (1.1.2). Then we find the processes a_t^* and s_t^* which correspond to optimality in terms of our expectation-error criterion for the family

of laws given by (1.1.2). Finally, we show that we can indeed build a discretization rule which leads to the optimal a_t^* and s_t^* in the limiting distribution of the hedging error.

Using an asymptotic framework to design optimal discretizations of hedging strategies has been quite a popular approach in the recent years. This method (although in a slightly different context) is in particular used in [Fuk11a, Fuk11c, GL14a] in the continuous setting whereas the case with jumps is investigated in [RT14]. All these works aim at minimizing some form of transaction costs (typically the number of trades) under some constraint on the L^2 -norm of the hedging error. Here we also put a constraint on the L^2 -norm of the hedging error. However, instead of minimizing transaction costs, we maximize the expectation of the hedging error. Thus our viewpoint is that of a trader giving himself a lower bound on the quality of his hedge (the L^2 -norm of the hedging error), but allowing himself to try to take advantage of market trends provided the constraint is satisfied.

In practice, our optimal strategy should probably only be considered as a benchmark. Indeed, its computation requires precise knowledge of the market trend over the hedge horizon. This is of course not realistic since any kind of statistical estimation of the drift is irrelevant in this high frequency setting. However, some practitioners have views on the market based, for example, on economic analysis, and our work gives them a way to incorporate their beliefs in their hedging strategies. To assess the robustness of our approach with respect to drift misspecification, in Section 1.5.2 we compare numerically the performance of the optimal strategy to the performance of alternative strategies based on misspecified values of the drift. We find that the optimal strategies (with misspecified drift) can still recover a large part of the extra return if the directional bet is good.

The paper is organized as follows. In Section 1.2, we investigate the set of admissible discretization rules, that is those leading to a limiting law of the form (1.1.2). In particular, we extend the examples provided in [Fuk11c] by showing that the discretization rules based on hitting times of stochastic barriers are admissible. In Section 1.3, we consider a first criterion for optimizing the trading times: the modified Sharpe ratio. It enables us to carry out very simple computations. However, the relevance of the modified Sharpe ratio being in fact quite arguable, a more suitable approach in which we consider an expectation-error type criterion is investigated in Section 1.4. Using tools from linear-quadratic optimal control theory, explicit developments are provided in the Black-Scholes model in Section 1.5. Numerical illustrations are given in Section 1.5.2 and the longest proofs are relegated to an appendix.

1.2 Assumptions and admissible strategies

In this section, we detail our assumptions on the processes X and Y together with the admissibility conditions for the sampling schemes.

1.2.1 Assumptions on the dynamics and admissibility conditions

Let $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ be a filtered probability space satisfying the usual conditions. We write Y for the price dynamics of the underlying risky asset. Let $T > 0$ stand for the maturity of the derivative to be hedged. We assume that the benchmark hedging strategy deduced from a theoretical model simply consists in holding a certain number of units of the underlying asset, denoted by X , and some cash in a self-financed way, under zero interest rates. Throughout the paper, we assume both Y and X are Itô processes of the form

$$dY_t = b_t^Y dt + \sigma_t^Y dW_t^Y, \quad dX_t = b_t^X dt + \sigma_t^X dW_t^X \quad (1.2.1)$$

on $[0, T]$, where W^X and W^Y are \mathbb{F} -Brownian motions which may be arbitrarily correlated, and the coefficients of X and Y satisfy the following technical assumptions.

Assumption 1.2.1.

1. The processes b^Y , b^X , σ^Y , and σ^X are adapted and continuous on $[0, T]$ almost surely.
2. The volatility process σ^Y of Y is positive on $[0, T]$ almost surely.
3. The volatility process σ^X of X is positive on $[0, T]$ almost surely.
4. The instantaneous Sharpe ratio $\rho = b^Y/\sigma^Y$ satisfies

$$\mathbb{E}\left[\int_0^T \rho_t^2 dt\right] < +\infty.$$

Remark 1.2.1. Up to some technicalities, the requirement on the positivity of the volatility processes σ^X and σ^Y might be weakened. See for example the perturbation arguments in [GL14b].

Example 1.2.1 (The Black-Scholes model). The case of $b_t^Y = bY_t$ and $\sigma_t^Y = \sigma Y_t$ with constants b and $\sigma > 0$ corresponds to the Black-Scholes model. The instantaneous Sharp ratio $\rho = b/\sigma$ is a constant. To hedge a call option with payoff $(Y_T - K)_+$ and strike $K > 0$, the standard theory suggests to use the so-called Delta hedging strategy:

$$X_t = \Phi(d_1(t, Y_t)), \quad d_1(t, y) = \frac{\log(y/K) + \sigma^2(T-t)/2}{\sigma\sqrt{T-t}},$$

where Φ stands for the distribution function of a standard Gaussian random variable. By Itô's formula, we see that X is an Itô process of the form (1.2.1) with $W^X = W^Y$ and

$$\begin{aligned} b_t^X &= \phi(d_1(t, Y_t)) \left\{ \frac{\partial d_1}{\partial t}(t, Y_t) + \frac{\sigma^2}{2} \frac{\partial^2 d_1}{\partial y^2}(t, Y_t) Y_t^2 + b \frac{\partial d_1}{\partial y}(t, Y_t) Y_t \right\} + \frac{\sigma^2}{2} \left(\frac{\partial d_1}{\partial y}(t, Y_t) \right)^2 \phi'(d_1(t, Y_t)) Y_t^2, \\ \sigma_t^X &= \sigma \phi(d_1(t, Y_t)) \frac{\partial d_1}{\partial y}(t, Y_t) Y_t, \end{aligned}$$

with ϕ being the density of a standard Gaussian random variable. Almost surely, $Y_T \neq K$ and therefore both b^X and σ^X are continuous on $[0, T]$ and $b_T^X = \sigma_T^X = 0$. Furthermore σ^X is positive on $[0, T]$. Hence Assumption 1.2.1 is satisfied.

As explained in the introduction, in practice, the trader cannot realize the theoretical strategy X_t which typically implies continuous trading. Hence the quantity

$$\int_0^T X_s dY_s$$

only represents a benchmark terminal wealth and X_t is a benchmark hedging strategy. Thus, we discretize this strategy over the stopping times

$$0 = \tau_0^n < \tau_1^n < \dots < \tau_j^n < \dots,$$

so that for given n , almost surely, τ_j^n attains T for j large enough. Such array of stopping times is called a discretization rule. Consequently, if we define the discretized process X^n by

$$X_t^n = X_{\tau_j^n}, \quad t \in [\tau_j^n, \tau_{j+1}^n),$$

the hedging error Z_T^n with respect to the benchmark strategy writes as

$$Z_T^n = \int_0^T (X_{s-}^n - X_s) dY_s.$$

We now define the admissibility conditions for our discretization rules which we comment in the next subsection.

Condition 1.2.1 (Admissibility conditions). *A discretization rule (τ_j^n) is admissible if there exist continuous \mathbb{F} -adapted processes a and s satisfying*

$$\mathbb{E} \left[\int_0^T (1 + (\rho_t)^2) (a_t^2 + s_t^2) (\sigma_t^Y)^2 dt \right] < \infty, \quad (1.2.2)$$

and a positive sequence ε_n tending to zero such that:

- The first two moments of the renormalized hedging error $\varepsilon_n^{-1} Z_T^n$ converge to those of a random variable of the form

$$Z_{a,s}^* = \frac{1}{3} \int_0^T s_t dY_t + \frac{1}{\sqrt{6}} \int_0^T \left(a_t^2 - \frac{2}{3} s_t^2 \right)^{1/2} \sigma_t^Y dB_t, \quad (1.2.3)$$

that is,

$$\mathbb{E}[\varepsilon_n^{-1} Z_T^n] \rightarrow \mathbb{E}[Z_{a,s}^*], \quad \mathbb{E}[(\varepsilon_n^{-1} Z_T^n)^2] \rightarrow \mathbb{E}[(Z_{a,s}^*)^2], \quad (1.2.4)$$

where B is a Brownian motion, independent of all the other quantities.

- Almost surely, the processes a_t and s_t satisfy $a_t^2 \geq s_t^2$, for all $t \in [0, T]$.

1.2.2 Comments on the admissibility conditions

Equation 1.2.2 is simply a technical integrability condition. We now give the interpretation of the sequence ε_n . Recall that for fixed n , we deal with an increasing sequence of stopping times (τ_j^n) over $[0, T]$. Typically, ε_n^2 will represent the order of magnitude of the interarrival time $\tau_{j+1}^n - \tau_j^n$. For example, in the case of equidistant trading times with frequency n/T , ε_n can simply be taken equal to $n^{-1/2}$. In the case of the hitting times based scheme consisting in rebalancing the portfolio each time the process X has varied by ν_n , where ν_n is a deterministic sequence tending to zero, one can choose $\varepsilon_n = \nu_n$ (since the order of magnitude of the time interval between two hitting times is ν_n^2).

The specific form (1.2.3) may appear rather ad hoc at first sight. However, it is in fact quite natural. Indeed, Proposition 1.2.1 below, which is proved in the Appendix and used to show the main result of the next subsection, indicates that as soon as the quadratic covariations $\varepsilon_n^{-2} \langle Z^n \rangle$ and $\varepsilon_n^{-1} \langle Z^n, Y \rangle$ have regular limits, the form (1.2.3) appears for the weak limit of the renormalized hedging error. So the idea for this admissibility condition is that in our asymptotic approach, we want to work in regular cases where the renormalized hedging error can be approximated by a random variable of the form (1.2.3). However, our asymptotic optimality criterion will be based on the first two moments of the renormalized hedging error only. Therefore, we just require these first two moments to be asymptotically close to those of a random variable of the form (1.2.3) (in particular we do not impose the convergence in law of the renormalized hedging error towards $Z_{a,s}^*$, although this is the underlying idea behind this admissibility condition). We now give Proposition 1.2.1.

Proposition 1.2.1. *If there exists a sequence $\varepsilon_n \rightarrow 0$ and continuous processes s and a such that*

$$\varepsilon_n^{-2} \langle Z^n \rangle \rightarrow \frac{1}{6} \int_0^\cdot a_u^2 (\sigma_u^Y)^2 du, \quad (1.2.5)$$

$$\varepsilon_n^{-1} \langle Z^n, Y \rangle \rightarrow \frac{1}{3} \int_0^\cdot s_u (\sigma_u^Y)^2 du, \quad (1.2.6)$$

in probability uniformly on $[0, T]$, then $\varepsilon_n^{-1} Z^n$ converges in law to

$$\frac{1}{3} \int_0^\cdot s_t dY_t + \frac{1}{\sqrt{6}} \int_0^\cdot \left(a_t^2 - \frac{2}{3} s_t^2 \right)^{1/2} \sigma_t^Y dB_t \quad (1.2.7)$$

in $C[0, T]$. In particular, the convergence in law of $\varepsilon_n^{-1} Z_T^n$ to $Z_{a,s}^*$ defined by (1.2.3) holds. If in addition,

$$\varepsilon_n^{-4/3} \sup_{j \geq 0} (\tau_{j+1}^n \wedge T_0 - \tau_j^n \wedge T_0) \rightarrow 0 \quad (1.2.8)$$

in probability, for all $T_0 \in [0, T]$, then almost surely $a_t^2 \geq s_t^2$ for all $t \in [0, T]$.

We now consider the processes a_t^2 and s_t appearing in the admissibility conditions. We place ourselves in the situation where Proposition 1.2.1 can be applied. In that case, an inspection of the proof of this proposition shows that the inequality $a_t^2 \geq s_t^2$ essentially follows from the elementary fact that $\mathbb{E}[\Delta^2]\mathbb{E}[\Delta^4] \geq \mathbb{E}[\Delta^3]^2$ for a general random variable Δ . Indeed, a_t^2 and s_t are respectively related to the local third and fourth conditional moments of the increments of X . Proposition 1.2.2 below, which is proved in the Appendix and used to show the main result in the next subsection, somehow illustrates the connections between a_t^2 and s_t on the one hand and the conditional moments on the other hand. To that end, let $\Delta_{j,n} = X_{\tau_{j+1}^n} - X_{\tau_j^n}$ be the increment of X between τ_j^n and τ_{j+1}^n and N_t^n be the number of rebalancing times until time t :

$$N_t^n = \max \{j \geq 0 \mid \tau_j^n \leq t\}.$$

The following proposition holds.

Proposition 1.2.2. *Let ε_n be a positive sequence tending to 0 and s and a be continuous processes. Assume the following:*

- *The family of random variables*

$$\varepsilon_n^{-4} \sup_{t \in [0, T]} |X_t^n - X_t|^4 \quad (1.2.9)$$

is uniformly integrable.

- *The following uniform convergences in probability on $[0, T_0]$ hold for all $T_0 \in [0, T]$:*

$$\begin{aligned} \varepsilon_n^{-1} \sum_{j=0}^{N_t^n} \kappa_{\tau_j^n} \mathbb{E}[\Delta_{j,n}^3 \mid \mathcal{F}_{\tau_j^n}] &\rightarrow - \int_0^{\cdot} s_u (\sigma_u^Y)^2 du, \\ \varepsilon_n^{-2} \sum_{j=0}^{N_t^n} \kappa_{\tau_j^n} \mathbb{E}[\Delta_{j,n}^4 \mid \mathcal{F}_{\tau_j^n}] &\rightarrow \int_0^{\cdot} a_u^2 (\sigma_u^Y)^2 du, \end{aligned} \quad (1.2.10)$$

where $\kappa_u = (\sigma_u^Y / \sigma_u^X)^2$.

Then the convergences (1.2.5), (1.2.6) and (1.2.8) hold.

Proposition 1.2.2 is useful to obtain the convergences (1.2.5) and (1.2.6) for a given discretization rule since it is usually easy to have approximate values of the conditional moments of the increments. We actually apply this approach in the proof of the main result of the next subsection.

1.2.3 Examples of admissible discretization rules

We show in this section that the most common discretization rules are admissible. We start with hitting times based schemes. We have the following result.

Proposition 1.2.3 (Hitting times based discretization rule). *Let ε_n be a positive sequence tending to zero and \underline{l} and \bar{l} be two adapted processes which are positive and continuous on $[0, T]$ almost surely with*

$$\mathbb{E} \left[\int_0^T (1 + (\rho_t)^2) (\bar{l}_t \vee \underline{l}_t)^2 (\sigma_t^Y)^2 dt \right] < \infty. \quad (1.2.11)$$

The discretization rule based on the hitting times of $\varepsilon_n \underline{l}_t$ or $\varepsilon_n \bar{l}_t$ by the process X :

$$\tau_{j+1}^n = \inf \left\{ t > \tau_j^n : X_t \notin (X_{\tau_j^n} - \varepsilon_n \underline{l}_t, X_{\tau_j^n} + \varepsilon_n \bar{l}_t) \right\} \wedge T \quad (1.2.12)$$

is admissible with

$$s_t = \underline{l}_t - \bar{l}_t, \quad a_t^2 = s_t^2 + \underline{l}_t \bar{l}_t. \quad (1.2.13)$$

Moreover,

$$\varepsilon_n^2 N_t^n \rightarrow \int_0^t \frac{(\sigma_u^X)^2}{\underline{l}_u \bar{l}_u} du \quad (1.2.14)$$

in probability for all $t \in [0, T]$.

It is interesting to note here that the limit $Z_{a,s}^*$ does not depend on the structure of X .

This result is particularly important since many traders monitor the values of the increments of their so-called delta (which corresponds to the process X) in order to decide when to rebalance their portfolio. Thus, they are indeed using hitting times based strategies. The use of time varying barriers (which correspond to \underline{l} and \bar{l}) is more natural in financial practice than that of constant barriers (between τ_j^n and τ_{j+1}^n) treated in [Fuk11c, Lan13], although the arguments therein might be adapted to certain cases in our current situation.

Now remark that under the condition $a_t^2 > s_t^2$, we can always find some positive processes \bar{l}_t and \underline{l}_t such that (1.2.13) is satisfied. Indeed, it is easy to see that the real numbers \bar{l}_t and $-\underline{l}_t$ can be taken as the roots of the quadratic equation $x^2 + s_t x + s_t^2 - a_t^2 = 0$. Under the condition $a_t^2 > s_t^2$, this equation admits two nonzero roots with different signs. Therefore, another interesting property of hitting times based schemes is the following.

Lemma 1.2.1. *For any pair of limiting processes s and a satisfying (1.2.2) and $a_t^2 > s_t^2$, we can always build a corresponding admissible discretization rule based on hitting times as in (1.2.12)-(1.2.13).*

Consequently, if one has some processes a_t and s_t as targets, Lemma 1.2.1 implies that a strategy giving rise to these processes in the limiting distribution (1.2.3) can be found. We will work in this framework in Section 1.4. Remark that there are infinitely many strategies for which the renormalized hedging error converges in law to $Z_{a,s}^*$ with the same a and s due to the degree of freedom in choosing the normalizing sequence ε_n . The hitting time strategy is an efficient one among them in the sense that it attains the asymptotic lower bound (1.1.3).

Another classical discretization rule is given by equidistant trading times. Here, the integrability property (1.2.4) in the admissibility conditions does not hold in full generality. Compared to the hitting times setting, this is because the deviations of the benchmark strategy are not explicitly controlled by the barriers. Nevertheless, the following example describes a reasonable framework under which such a discretization rule is admissible.

Proposition 1.2.4 (Equidistant sampling discretization rule). *Consider the hedging strategy of a European option with payoff $h(Y_T)$ where the underlying Y_t follows a diffusion process of the form*

$$dY_t = b(t, Y_t)Y_t dt + \sigma(t, Y_t)Y_t dW_t,$$

with b , σ and h deterministic functions satisfying the following regularity assumptions:

- b and σ are bounded, of class $C^{1,\infty}([0, T] \times \mathbb{R}_+)$, such that $x \frac{\partial b}{\partial x}$, $x \frac{\partial \sigma}{\partial x}$, $x^2 \frac{\partial^2 \sigma}{\partial x^2}$ and $x^3 \frac{\partial^3 \sigma}{\partial x^3}$ are bounded;
- σ is Hölder continuous in x and bounded away from zero uniformly in $(t, x) \in [0, T] \times \mathbb{R}_+$.

— h is almost everywhere differentiable with

$$|h| + |h'| \leq L(1 + |x|^\alpha)$$

almost everywhere, for positive constants L and α .

Define the delta hedging portfolio:

$$X_t = \frac{\partial P}{\partial y}(t, Y_t), \quad \text{with } P(t, y) = \mathbb{E}_{(t,y)}^{\mathbb{Q}}[h(Y_T)],$$

where $\mathbb{E}^{\mathbb{Q}}$ denotes the expectation operator under the risk neutral probability. Let ε_n be a positive sequence tending to zero. Then the equidistant trading times discretization rule:

$$\tau_j^n = j\varepsilon_n^2, \quad j = 0, \dots, n, \dots$$

is admissible (under the original measure). Moreover, we can take

$$s_t = 0, \quad a_t^2 = 3(\sigma_t^X)^2.$$

Proof. The proof of Proposition 1.2.4 follows easily from the results in [Zha99]. Indeed, with the notation of this proposition,

$$\sigma_t^X = Y_t \sigma(t, Y_t) \frac{\partial^2 P}{\partial y^2}(t, Y_t),$$

and by Proposition 2.2.31 in [Zha99], using the Hölder inequality and the assumptions on σ and b , we have that

$$\mathbb{E}[Z_{a,s}^*] = 0, \quad \mathbb{E}[(Z_{a,s}^*)^2] = \frac{1}{2} \mathbb{E} \left[\int_0^T (\sigma_t^Y)^2 (\sigma_t^X)^2 dt \right]$$

and (1.2.2) is satisfied. For the convergence of the second moment it suffices to use Theorem 2.4.1 in [Zha99]. The convergence of the first moment to zero, on the other hand, is shown on pages 86–88 of this reference. □

Note that the discretization rule based on equidistant trading times will not be of interest for us since the associated s_t process vanishes and so the expectation of the limiting variable is zero.

1.3 Asymptotic optimality: a preliminary approach

Our viewpoint is that the trader's priority is to get a small hedging error. However, once this error is suitably controlled, he may try to take advantage of the directional views he has on the market. Hence, adopting the asymptotic approximation under which the first two moments of the renormalized hedging error are given by those of $Z_{a,s}^*$, we aim at maximizing $\mathbb{E}[Z_{a,s}^*]$ while keeping $\mathbb{E}[(Z_{a,s}^*)^2]$ reasonably small. This very problem is treated in Section 1.4.

Here, as a first step, we consider the approximation for $\mathbb{E}[(Z_{a,s}^*)^2]$ given by $\mathbb{E}[(Z_{a,s}^{*,c})^2]$, where $Z_{a,s}^{*,c}$ denotes the sum of the two integrals with respect to the Brownian motions W^Y and B in the definition of $Z_{a,s}^*$ in Equation (1.2.3), that is

$$Z_{a,s}^{*,c} = \frac{1}{3} \int_0^T s_t \sigma_t^Y dW_t^Y + \frac{1}{\sqrt{6}} \int_0^T \left(a_t^2 - \frac{2}{3} s_t^2 \right)^{1/2} \sigma_t^Y dB_t.$$

To do so, we place ourselves in this section under the additional admissibility condition that the renormalized hedging error weakly converges in the sense of (1.2.7) and we take s_t and a_t^2 as

the processes in the limit (1.2.7) (so s_t and a_t^2 are uniquely defined). Replacing $\mathbb{E}[(Z_{a,s}^*)^2]$ by $\mathbb{E}[(Z_{a,s}^{*,c})^2]$ is technically very convenient but in practice quite arguable since this approximation is meaningful only when the drift is small. However, our aim here is only to have a first rough idea about the form of the optimal discretization rules. Since we wish to get the moment of order one large while that of order two remains controlled, we consider that we want to maximize the so-called modified Sharpe ratio S defined by

$$S = S(a, s) = \frac{\mathbb{E}[Z_{a,s}^*]}{\sqrt{\mathbb{E}[(Z_{a,s}^{*,c})^2]}}.$$

This ratio is said to be modified since we use $\mathbb{E}[(Z_{a,s}^{*,c})^2]$ instead of the variance of $Z_{a,s}^*$.

Hence we are looking for strategies which maximize S . To do so, we now introduce the notion of nearly efficient (modified) Sharpe ratio.

Definition 1.3.1 (Nearly efficient Sharpe ratio). *The value $S^* \in \mathbb{R}$ is said to be a nearly efficient Sharpe ratio if:*

1. *For any admissible discretization rule with associated limiting processes a and s , the associated modified Sharpe ratio $S(a, s)$ satisfies*

$$S(a, s) \leq S^*.$$

2. *For any $\eta > 0$, there exists a discretization rule with associated limiting processes a and s such that*

$$S(a, s) \geq S^* - \eta.$$

We only consider nearly efficient ratios since our strategies will not enable us to attain exact efficiency (which would correspond to $\eta = 0$ in the previous definition). Of course, the slight difference between efficient and nearly efficient ratios has no importance in practice.

In our setting, for any limiting variable $Z_{a,s}^*$, we have

$$S(a, s) = \frac{\mathbb{E}\left[\frac{1}{3} \int_0^T s_t b_t^Y dt\right]}{\left(\mathbb{E}\left[\frac{1}{9} \int_0^T s_t^2 (\sigma_t^Y)^2 dt + \frac{1}{6} \int_0^T (a_t^2 - \frac{2}{3} s_t^2) (\sigma_t^Y)^2 dt\right]\right)^{1/2}}.$$

Now, the admissibility condition $a_t^2 \geq s_t^2$ implies

$$S(a, s) \leq \frac{\sqrt{6}}{3} \frac{\mathbb{E}\left[\int_0^T s_t b_t^Y dt\right]}{\left(\mathbb{E}\left[\int_0^T s_t^2 (\sigma_t^Y)^2 dt\right]\right)^{1/2}}$$

and Cauchy-Schwarz inequality gives

$$S(a, s) \leq \frac{\sqrt{6}}{3} \left(\mathbb{E}\left[\int_0^T \left(\frac{b_t^Y}{\sigma_t^Y}\right)^2 dt\right]\right)^{1/2}.$$

This provides an upper bound for the modified Sharpe ratio. We now wish to find a discretization rule enabling to (almost) attain this upper bound. To achieve this, our rule must be such that for the associated processes a_t and s_t , the inequalities used above ($a_t^2 \geq s_t^2$ and Cauchy-Schwarz) become almost equalities. This means that a_t should be close to s_t and s_t should essentially be proportional to $b_t^Y / (\sigma_t^Y)^2$. Furthermore, we want the product $s_t b_t^Y$ to be essentially positive in order to get a positive modified Sharpe ratio. If we look for this rule among the hitting times based schemes specified by two processes $(\underline{l}_t, \bar{l}_t)$, Lemma 1.2.1 implies that

- the difference $l_t - \bar{l}_t$ should be essentially proportional to $b_t/(\sigma_t^Y)^2$,
- the product $l_t \bar{l}_t$ should be negligible compared to $(l_t - \bar{l}_t)^2$,
- the term $(l_t - \bar{l}_t)b_t^Y$ should be essentially positive.

From these remarks together with Proposition 1.2.3, we easily deduce the following theorem.

Theorem 1.3.1. *Suppose that for all $t \leq T$, $b_t^Y \neq 0$. Then the value*

$$\frac{\sqrt{6}}{3} \left(\mathbb{E} \left[\int_0^T \left(\frac{b_t^Y}{\sigma_t^Y} \right)^2 dt \right] \right)^{1/2}$$

is a nearly efficient Sharpe ratio. It is approximately attained by the discretization rule defined for $\lambda > 0$ by

$$\tau_{j+1}^{n,\lambda} = \inf \left\{ t > \tau_j^{n,\lambda}; X_t - X_{\tau_j^{n,\lambda}} = -\frac{b_t^Y}{(\sigma_t^Y)^2} e^{\lambda \varepsilon_n} \text{ or } \frac{b_t^Y}{(\sigma_t^Y)^2} e^{-\lambda \varepsilon_n} \right\}, \quad \tau_0^n = 0. \quad (1.3.1)$$

Indeed,

$$\lim_{\lambda \rightarrow +\infty} S(\lambda) = \frac{\sqrt{6}}{3} \left(\mathbb{E} \left[\int_0^T \left(\frac{b_t^Y}{\sigma_t^Y} \right)^2 dt \right] \right)^{1/2},$$

where $S(\lambda)$ denotes the modified Sharpe ratio obtained for the law of the variable $Z_{a,s}^$ associated to the discretization rule (1.3.1) with parameter λ .*

This result provides simple and explicit strategies for optimizing the modified Sharpe ratio. It is also very easy to interpret. Indeed, we see that in order to take advantage of the drift, one needs to consider asymmetric barriers. The limitation is that we do not really control accurately the magnitude of the hedging error at maturity.

The asymptotic setting simply means that we require λ to be quite large while $e^{\lambda \varepsilon_n}$ is small. When using such discretization rule in practice, it is reasonable to consider that the trader fixes a maximal value for the asymmetry between the barriers controlled by λ . This way he can choose the parameter λ . Then ε_n is set to match the bound on $\mathbb{E}[(Z_{a,s}^{*,c})^2]$ that the trader does not want to exceed.

1.4 Asymptotic expectation-error optimization

In this section, we now consider a natural expectation-error type criterion in order to optimize our discretization rules. To do so, we work in an asymptotic setting where we are looking for discretization rules which are optimal in the expectation-error sense for their associated limiting random variable $Z_{a,s}^*$. We start with a definition inspired by the classical portfolio theory.

Definition 1.4.1 (Efficient couples). *A couple $(m, v) \in (\mathbb{R}^+)^2$ is said to be nearly efficient if there exists no admissible discretization rule such that its associated limiting random variable $Z_{a,s}^*$ satisfies*

$$\mathbb{E}[Z_{a,s}^*] \geq m, \quad \mathbb{E}[(Z_{a,s}^*)^2] < v.$$

and for any $\eta > 0$, there exists an admissible discretization rule such that its associated limiting random variable $Z_{a,s}^$ satisfies*

$$\mathbb{E}[Z_{a,s}^*] = m, \quad \mathbb{E}[(Z_{a,s}^*)^2] \leq v + \eta.$$

It is said to be efficient if we can take $\eta = 0$.

We introduce the set \mathcal{Z}_T of random variables of the form

$$Z_{T,s} = \frac{1}{3} \int_0^T s_t dY_t + \frac{1}{3\sqrt{2}} \int_0^T s_t \sigma_t^Y dB_t, \quad (1.4.1)$$

where B is a Brownian motion independent of \mathcal{F} and s_t is an adapted continuous process such that

$$\mathbb{E} \left[\int_0^T (1 + (\rho_t)^2) s_t^2 (\sigma_t^Y)^2 dt \right] < \infty. \quad (1.4.2)$$

We can now state our main result which enables us to compute efficient discretization rules.

Theorem 1.4.1. *The following results hold:*

- Let (m^*, v^*) be an efficient couple. Then there exists an adapted continuous process s^* satisfying (1.4.2), such that $\mathbb{E}[Z_{T,s^*}] = m^*$ and $\mathbb{E}[(Z_{T,s^*})^2] = v^*$, and for no other process s satisfying (1.4.2), one has $\mathbb{E}[Z_{T,s}] \geq m^*$ and $\mathbb{E}[(Z_{T,s})^2] < v^*$.
- Conversely, let s^* be an adapted continuous process satisfying (1.4.2), such that $\mathbb{E}[Z_{T,s^*}] = m^*$ and $\mathbb{E}[(Z_{T,s^*})^2] = v^*$. Then, (m^*, v^*) is a nearly efficient couple. More precisely, let $\delta > 0$ and $(\underline{l}_t^\delta, \bar{l}_t^\delta)$ be defined by

$$\underline{l}_t^\delta - \bar{l}_t^\delta = s_t^*, \quad (\underline{l}_t^\delta)^2 - \underline{l}_t^\delta \bar{l}_t^\delta + (\bar{l}_t^\delta)^2 = (s_t^*)^2 + \frac{6\delta}{(\sigma_t^Y)^2},$$

that is

$$\underline{l}_t^\delta = \sqrt{\frac{(s_t^*)^2}{4} + \frac{6\delta}{(\sigma_t^Y)^2}} + \frac{s_t^*}{2}, \quad \bar{l}_t^\delta = \sqrt{\frac{(s_t^*)^2}{4} + \frac{6\delta}{(\sigma_t^Y)^2}} - \frac{s_t^*}{2}. \quad (1.4.3)$$

Then the hitting times based discretization rule specified through the barriers $(\underline{l}_t^\delta, \bar{l}_t^\delta)$ satisfies

$$\mathbb{E}[Z_{a,s}^*] = m^*, \quad \mathbb{E}[(Z_{a,s}^*)^2] = v^* + \delta T.$$

Therefore, we have reduced the impulse control problem of finding the optimal rebalancing times to a classical expectation-error optimization with continuous dynamics. The solutions of this problem can be obtained by solving for $\mu > 0$

$$\inf_{(s_t)} \{ -\mathbb{E}[Z_{T,s}] + \mu \mathbb{E}[(Z_{T,s})^2] \},$$

for which we can apply the theory of linear-quadratic optimal control, see for example [LZ02] and [ZL00]. As shown in the next section, we can even obtain closed formulas in the case where the underlying has deterministic drift and volatility. Note that again, our barriers strategies enable us to attain only nearly efficient couples. Indeed, reaching efficient couples would lead to the use of degenerate barriers with $\delta = 0$. This does not make sense in practice, however δ can of course be selected small.

1.5 Black-Scholes model with time-varying coefficients

In this section, we explain how our method can be applied in practice through the simple example of the Black-Scholes model with time-varying coefficients. So, we assume the underlying risky asset follows the dynamics

$$dY_t = Y_t(b_t dt + \sigma_t dW_t),$$

where b_t and σ_t are continuous deterministic functions. We also assume b_t and σ_t do not vanish. Using the theory of linear-quadratic optimal control, we give an explicit solution for the problem of designing optimal rebalancing times in this specific setting.

1.5.1 Explicit formulas

We aim at finding the efficient couples for the controlled random variables of the form $Z_{T,s}$ as in (1.4.1). Following [ZL00], such problem is classically recast as follows: solving for any $\mu > 0$ the optimization problem

$$\inf_{(s_t, 0 \leq t \leq T)} -\mathbb{E}[Z_{T,s}] + \mu \mathbb{E}[(Z_{T,s})^2] = \inf_{(s_t, 0 \leq t \leq T)} \mu \mathbb{E}[(Z_{T,s} - \frac{1}{2\mu})^2] - \frac{1}{4\mu}.$$

Let us define the family of processes of the form

$$d\check{Z}_t = s_t Y_t (\check{b}_t dt + \check{\sigma}_t dW_t), \quad \check{Z}_0 = 0,$$

with $\check{b}_t = b_t/3$, $\check{\sigma}_t = \sigma_t/3$ and s_t being adapted continuous. Using obvious computations, the independence between the process B in Equation (1.4.1) and \mathcal{F} , and the fact that s_t is \mathcal{F} -adapted, we get $\mathbb{E}[\check{Z}_T] = \mathbb{E}[Z_{T,s}]$ and

$$\mu \mathbb{E}[(\check{Z}_T - \frac{1}{2\mu})^2] = \mu \mathbb{E}[(Z_{T,s} - \frac{1}{2\mu})^2] - \frac{\mu}{18} \mathbb{E}[\int_0^T (s_t \sigma_t Y_t)^2 dt].$$

Hence, we can equivalently solve

$$\inf_{(s_t, 0 \leq t \leq T)} \mathbb{E}[\mu \check{Z}_T^2 + \frac{\mu}{2} \int_0^T (s_t \check{\sigma}_t Y_t)^2 dt],$$

with

$$d\tilde{Z}_t = s_t Y_t (\tilde{b}_t dt + \tilde{\sigma}_t dW_t), \quad \tilde{Z}_0 = -\frac{1}{2\mu}.$$

Using the results of [ZL00] which are summarized in Theorem 1.B.1 in Appendix 1.B, the optimal control s_t^* and the optimally controlled process \tilde{Z}_t^* satisfy

$$s_t^* Y_t = -\frac{1}{\tilde{b}_t} \frac{\dot{P}_t}{P_t} \tilde{Z}_t^*,$$

where P_t is the solution of the (ordinary) differential equation

$$\dot{P}_t = \rho_t^2 \frac{P_t^2}{P_t + \mu}, \quad P_T = 2\mu,$$

with $\rho_t = b_t/\sigma_t$. The solution of this equation is given by

$$P_t = \frac{\mu}{L\left(\frac{1}{2} \exp\left(\int_t^T \rho_s^2 ds + \frac{1}{2}\right)\right)},$$

with $L(\cdot)$ being the inverse function of $x \mapsto xe^x$. Moreover, the optimal process \tilde{Z}^* satisfies

$$\frac{d\tilde{Z}_t^*}{\tilde{Z}_t^*} = -\frac{\dot{P}_t}{P_t} (dt + \frac{1}{\rho_t} dW_t), \quad \tilde{Z}_0^* = -\frac{1}{2\mu}.$$

Therefore, we obtain

$$\mathbb{E}[\tilde{Z}_T^*] = -\frac{1}{2\mu} \frac{P_0}{P_T}.$$

Using Theorem 1.B.1, we get

$$\mathbb{E}\left[(\tilde{Z}_T^*)^2 + \frac{1}{2} \int_0^T (s_t^* Y_t \tilde{\sigma}_t)^2 dt\right] = \left(\frac{1}{2\mu}\right)^2 \frac{P_0}{P_T}.$$

Consequently, we have that the optimal variable Z_{T,s^*} satisfies

$$\mathbb{E}[Z_{T,s^*}] = \frac{1}{2\mu} \left(1 - \frac{P_0}{P_T}\right)$$

and

$$\mathbb{E}\left[\left(Z_{T,s^*} - \frac{1}{2\mu}\right)^2\right] = \left(\frac{1}{2\mu}\right)^2 \frac{P_0}{P_T}.$$

Hence

$$\mathbb{E}[(Z_{T,s^*})^2] = \left(\frac{1}{2\mu}\right)^2 \left(1 - \frac{P_0}{P_T}\right).$$

We have thus proved the following proposition.

Proposition 1.5.1. *In the Black-Scholes model with time-varying coefficients, the efficient points are the couples of the form*

$$\left(\frac{1}{2\mu}, \frac{1}{4\mu^2}\right) \left(1 - \frac{P_0}{P_T}\right), \quad (1.5.1)$$

with $\mu > 0$ (notice that the ratio $\frac{P_0}{P_T}$ does not depend on μ). Furthermore, the associated process s_t^* enabling us to compute optimal rules according to Theorem 1.4.1 is given by

$$\frac{1}{3} s_t^* Y_t = -\frac{1}{b_t} \frac{\dot{P}_t}{P_t} \tilde{Z}_t^*,$$

with

$$\frac{d\tilde{Z}_t^*}{\tilde{Z}_t^*} = -\frac{1}{b_t} \frac{\dot{P}_t}{P_t} \frac{dY_t}{Y_t}, \quad \tilde{Z}_0^* = -\frac{1}{2\mu}.$$

Note that, in practice, \tilde{Z}^* is not observable. However, it can of course be approximated by a process $\tilde{Z}^{(*)}$ thanks to historical data, using for example a scheme of the form

$$\tilde{Z}_{t_{i+1}}^{(*)} = \tilde{Z}_{t_i}^{(*)} \left(1 - \frac{1}{b_{t_i}} \frac{\dot{P}_{t_i}}{P_{t_i}} \frac{Y_{t_{i+1}} - Y_{t_i}}{Y_{t_i}}\right), \quad \tilde{Z}_0^{(*)} = -\frac{1}{2\mu},$$

where the t_i are the observation times of market data.

To implement our strategy in practice, the trader must choose three parameters: the value μ which determines a point on the efficient frontier, the asymmetry parameter δ and finally the value ε_n which determines how close to the limit we are. However, in the Black-Scholes model the sequence of renormalized hedging error processes corresponding to an optimal discretization rule satisfies

$$\mathbb{E}[\mu \varepsilon_n^{-1} Z_T^n] \rightarrow \frac{1}{2} \left(1 - \frac{P_0}{P_T}\right) \quad \text{and} \quad \mathbb{E}[(\mu \varepsilon_n^{-1} Z_T^n)^2] \rightarrow \frac{1}{4} \left(1 - \frac{P_0}{P_T}\right).$$

This means that (in the asymptotic regime) ε_n and μ are redundant parameters, and one can simply take $\mu = 1$. This leaves the trader with the choice of ε_n and δ . These parameters will typically be chosen to meet two conflicting objectives: keep the tracking error below a certain threshold, and limit the number of transaction dates. With the above choice of $\mu = 1$, using the suboptimal strategy of Theorem 1.4.1, we see that

$$\mathbb{E}[(\varepsilon_n)^{-1} Z_T^n]^2 \rightarrow \frac{P_T - P_0}{P_T} + \delta T$$

as $n \rightarrow \infty$. Therefore, if δ is small, its effect on the tracking error is negligible. However, it has a strong effect on the total number of transaction dates, which tends to infinity as $\delta \rightarrow 0$. In view of these observations, we propose the following approach for choosing ε_n and δ :

- First, choose ε_n to meet the tracking error target for the optimal strategy (assuming $\delta = 0$).
- Then, choose the smallest δ which meets the target of the total expected number of transaction dates (e.g., by simulation). The transaction dates target should be large enough so that the loss of optimality due to nonzero δ is negligible.

1.5.2 Numerical study

In this section, we present a numerical example illustrating the performance of our asymptotically optimal strategy for finite values of ε_n , as well as its robustness to drift misspecification.

The dynamics of the underlying Y is given by the Black-Scholes model with constant drift b and volatility σ :

$$dY_t = Y_t(bdt + \sigma dW_t),$$

and the reference strategy X by the continuous delta hedging strategy of European calls. The values of the model parameters are given in the following table:

b	σ	Y_0	T	K	X_0
{0.10, 0.20 }	{0.15, 0.30}	100.0	1.0	100	0.53

In the numerical implementation, we choose:

μ	ε_n
1.0	{1.0, 1.2, 1.4, 1.6, 1.8, 2.0, 2.4, 2.8, 3.2, 3.6, 4.0}

For each set of parameter values, and for each ε_n , we simulate 40000 trajectories of the asset price with 80000 discretization points. Then we compute the rebalancing dates according to the optimal strategy determined from formula (1.4.3), and evaluate the expectation and the second moment of the hedging error.

Remark 1.5.1. *To simplify the implementation, in the simulations, we have taken $\delta = 0$ in formula (1.4.3), which corresponds to a degenerate barrier. In practice, δ should be significantly positive in order to avoid microstructure effects.*

Validity of the asymptotic framework. We perform the computation for different values of ε_n in order to evaluate the speed of convergence of the strategies built with finite ε_n towards the asymptotic efficient frontier. Following the results of Section 1.5.1, for given ε_n , the moments of the hedging error $\mathbb{E}[Z_T^n]$ and $\mathbb{E}[(Z_T^n)^2]$ should be close to $\varepsilon_n \mathbb{E}[Z_{T,s^*}]$ and $\varepsilon_n^2 \mathbb{E}[(Z_{T,s^*})^2]$. Figure 1.1 shows the empirical counterparts of $\mathbb{E}[Z_T^n]$ and $\mathbb{E}[(Z_T^n)^2]$, together with the couples of the form $(x\mathbb{E}[Z_{T,s^*}], x^2\mathbb{E}[(Z_{T,s^*})^2])$, with $x \geq 0$ (red curve). The patterns being very similar in all cases, we only show one set of parameters in Figure 1.1. We see that indeed, the empirical first and second moments are close to the theoretical frontier, for quite a large choice of ε_n .

Performance under different market scenarios. We now fix ε_n and consider the performance of the optimal strategy under different scenarios for the market. Table 1 gives the first and second moments of hedging errors obtained by Monte Carlo simulations under these scenarios.

$(\mathbb{E}[Z_T^n], \mathbb{E}[(Z_T^n)^2])$	low drift ($b = 0.10$)	high drift ($b = 0.20$)
low volatility ($\sigma = 0.15$)	(0.043, 0.0085)	(0.136, 0.0316)
high volatility ($\sigma = 0.30$)	(0.010, 0.0019)	(0.041, 0.0083)

Table 1.1 – Hedging errors under different markets with $\varepsilon_n = 2$.

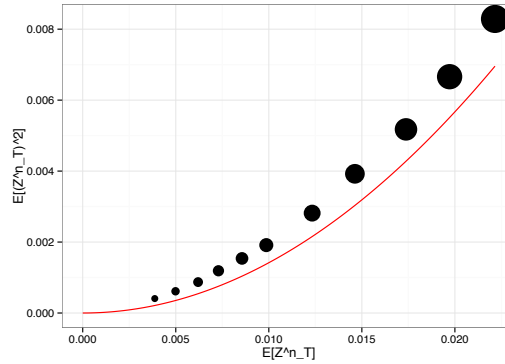


Figure 1.1 – Convergence to efficient frontier (point size prop. to ε_n , $b = 0.10$, $\sigma = 0.30$).

To give an idea of the order of magnitude of extra return which can be generated thanks to our approach, we place ourselves in the scenario with high drift and low volatility in Table 1.1. In this case, the call option premium is approximately equal to 7.97. The second moment $\mathbb{E}[(Z_T^n)^2] = 0.0316$ corresponds to a hedging error less than 2.3% of the option's premium (which is a rather tight bound), while the expected extra return $\mathbb{E}[Z_T^n] = 0.136$ is slightly more than 1.7% of the option's premium.

Robustness against drift misspecification. In practice, it is difficult to obtain an accurate estimation \hat{b} of the drift. Table 1.2 illustrates the robustness of the strategies which are deduced from misspecified drift \hat{b} , where the last column is the standard Sharpe ratio (expectation divided by standard deviation) of the hedging error. In particular, it shows that a right view on the market trend plays an essential role in order to obtain a positive expected gain.

\hat{b}/b	$\mathbb{E}[Z_T^n]$	$\mathbb{E}[(Z_T^n)^2]$	Sharpe ratio
1	0.136	0.0316	1.19
2	0.191	0.0646	1.14
0.5	0.070	0.0085	1.15
-1	-0.481	0.7129	-0.69

Table 1.2 – Performance with misspecified drift and $b = 0.20$, $\sigma = 0.15$, $\varepsilon_n = 2$.

Appendix 1.A Proofs

In the following C denotes a constant which may vary from line to line. Note that we use several localization procedures in the proofs. We often give them in detail since some of them are slightly unusual, in particular because of the fact that σ^X may vanish at maturity.

1.A.1 Proof of Proposition 1.2.1

We start by proving in a very standard way the stable convergence of $\varepsilon_n^{-1}Z^n$ in $C[0, T]$, which is stronger than the weak convergence. More precisely, we show that for any bounded continuous function f on $C[0, T]$ and bounded random variable U defined on $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$,

$$\lim_{n \rightarrow \infty} \mathbb{E}[Uf(\varepsilon_n^{-1}Z^n)] = \mathbb{E}[Uf(Z^*)],$$

where Z^* is defined by

$$Z_t^* = \frac{1}{3} \int_0^t s_u dY_u + \frac{1}{\sqrt{6}} \int_0^t \left(a_u^2 - \frac{2}{3} s_u^2 \right)^{1/2} \sigma_u^Y dB_u,$$

on an extension of $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ on which B is a Brownian motion independent of all the other quantities. For $K > 0$, we set

$$\alpha^K = \inf\{t > 0; |\rho_t| \geq K\} \wedge T.$$

Since ρ is continuous on $[0, T]$ almost surely,

$$\lim_{K \rightarrow \infty} \mathbb{P}[\alpha^K < T] = 0.$$

Now observe that

$$\begin{aligned} & |\mathbb{E}[Uf(\varepsilon_n^{-1}Z^n)] - \mathbb{E}[Uf(Z^*)]| \\ & \leq |\mathbb{E}[Uf(\varepsilon_n^{-1}Z^n)] - \mathbb{E}[Uf(\varepsilon_n^{-1}Z_{\cdot \wedge \alpha^K}^n)]| + |\mathbb{E}[Uf(\varepsilon_n^{-1}Z_{\cdot \wedge \alpha^K}^n)] - \mathbb{E}[Uf(Z_{\cdot \wedge \alpha^K}^*)]| \\ & \quad + |\mathbb{E}[Uf(Z_{\cdot \wedge \alpha^K}^*)] - \mathbb{E}[Uf(Z^*)]| \\ & \leq 4\|f\|_\infty \|U\|_\infty \mathbb{P}[\alpha^K < T] + |\mathbb{E}[Uf(\varepsilon_n^{-1}Z_{\cdot \wedge \alpha^K}^n)] - \mathbb{E}[Uf(Z_{\cdot \wedge \alpha^K}^*)]|. \end{aligned}$$

Consequently, it suffices to show that for any $K > 0$,

$$\lim_{n \rightarrow \infty} |\mathbb{E}[Uf(\varepsilon_n^{-1}Z_{\cdot \wedge \alpha^K}^n)] - \mathbb{E}[Uf(Z_{\cdot \wedge \alpha^K}^*)]| = 0.$$

Let

$$\mathcal{E} = \exp \left\{ - \int_0^{\alpha^K} \rho_t dW_t^Y - \frac{1}{2} \int_0^{\alpha^K} \rho_t^2 dt \right\}.$$

Since $\mathbb{E}[\mathcal{E}] = 1$, the measure \mathbb{Q} defined by

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \mathcal{E}$$

is a probability measure under which $Z_{\cdot \wedge \alpha^K}^n$ is a local martingale. Under \mathbb{Q} , the uniform convergences in probability on $[0, T]$ given in (1.2.5) and (1.2.6) remain true. Therefore by Theorem IX.7.3 of [JS13], we have the stable convergence of $\varepsilon_n^{-1}Z_{\cdot \wedge \alpha^K}^n$ to $Z_{\cdot \wedge \alpha^K}^*$ under \mathbb{Q} . Note that $\tilde{U} = U/\mathcal{E}$ is a \mathbb{Q} -integrable positive random variable and moreover, for all $A > 0$,

$$\begin{aligned} & |\mathbb{E}[Uf(\varepsilon_n^{-1}Z_{\cdot \wedge \alpha^K}^n)] - \mathbb{E}[Uf(Z_{\cdot \wedge \alpha^K}^*)]| \\ & \leq |\mathbb{E}^{\mathbb{Q}}[\tilde{U}f(\varepsilon_n^{-1}Z_{\cdot \wedge \alpha^K}^n)] - \mathbb{E}^{\mathbb{Q}}[\tilde{U}f(Z_{\cdot \wedge \alpha^K}^*)]| \\ & \leq |\mathbb{E}^{\mathbb{Q}}[(\tilde{U} \wedge A)f(\varepsilon_n^{-1}Z_{\cdot \wedge \alpha^K}^n)] - \mathbb{E}^{\mathbb{Q}}[(\tilde{U} \wedge A)f(Z_{\cdot \wedge \alpha^K}^*)]| + 2\|f\|_\infty \mathbb{E}^{\mathbb{Q}}[\tilde{U}\mathbb{1}_{\{\tilde{U} \geq A\}}]. \end{aligned}$$

The second term tends to 0 uniformly in n as $A \rightarrow \infty$. The first term converges to 0 due to the stable convergence under \mathbb{Q} , where it is used that $(\tilde{U} \wedge A)$ is a bounded random variable.

Now, we prove $a^2 \geq s^2$ under the additional condition (1.2.8). Since a and s are continuous, it suffices to show $a_t^2 \geq s_t^2$ for all $t \in [0, T]$. Fix $T_0 < T$ and let

$$\hat{\alpha}^K = \inf\{u > 0; |b_u^X| \vee \sigma_u^X \vee \sigma_u^Y \geq K \text{ or } \sigma_u^X \leq 1/K\} \wedge T_0 \quad (1.A.1)$$

for $K > 0$. Since σ^X is positive and continuous on $[0, T_0]$, we have

$$\lim_{K \rightarrow \infty} \mathbb{P}[\hat{\alpha}^K < T_0] = 0. \quad (1.A.2)$$

Therefore, it suffices to show

$$a_{u \wedge \hat{\alpha}^K}^2 \geq s_{u \wedge \hat{\alpha}^K}^2 \quad (1.A.3)$$

for all $u \geq 0$ and $K > 0$. Fix K and define the probability measure $\hat{\mathbb{Q}}$ by

$$\frac{d\hat{\mathbb{Q}}}{d\mathbb{P}} = \exp\left\{-\int_0^{\hat{\alpha}^K} \frac{b_u^X}{\sigma_u^X} dW_u^X - \frac{1}{2} \int_0^{\hat{\alpha}^K} \left(\frac{b_u^X}{\sigma_u^X}\right)^2 du\right\}.$$

Under $\hat{\mathbb{Q}}$, $X_{\cdot \wedge \hat{\alpha}^K}$ is a martingale with bounded quadratic variation. Since $\hat{\mathbb{Q}}$ is equivalent to \mathbb{P} , it suffices to show (1.A.3) under $\hat{\mathbb{Q}}$.

By (1.2.8), there exists a subsequence $\{n(k)\}$ such that

$$\hat{\mathbb{Q}}\left[\varepsilon_{n(k)}^{-4/3} \sup_{j \geq 0} (\tau_{j+1}^{n(k)} \wedge T_0 - \tau_j^{n(k)} \wedge T_0) > \frac{1}{k}\right] < \frac{1}{k}.$$

Let

$$T_k = \inf\left\{u > 0, \varepsilon_{n(k)}^{-4/3} \sup_{j \geq 0} (\tau_{j+1}^{n(k)} \wedge u - \tau_j^{n(k)} \wedge u) > \frac{1}{k}\right\} \wedge \hat{\alpha}^K.$$

Then

$$\lim_{k \rightarrow \infty} \hat{\mathbb{Q}}[T_k < \hat{\alpha}^K] = 0$$

and so,

$$\begin{aligned} \varepsilon_{n(k)}^{-1} \left\langle Z^{n(k)}, Y \right\rangle_{t \wedge T_k} &\rightarrow \frac{1}{3} \int_0^{t \wedge \hat{\alpha}^K} s_u (\sigma_u^Y)^2 du, \\ \varepsilon_{n(k)}^{-2} \left\langle Z^{n(k)} \right\rangle_{t \wedge T_k} &\rightarrow \frac{1}{6} \int_0^{t \wedge \hat{\alpha}^K} a_u^2 (\sigma_u^Y)^2 du, \end{aligned} \quad (1.A.4)$$

in probability as $k \rightarrow \infty$ for all $t \geq 0$. Let

$$\hat{\tau}_j^k = \tau_j^{n(k)} \wedge T_k$$

for $j \geq 0$. We now give three technical lemmas.

Lemma 1.A.1. *Let $\kappa_u = (\sigma_u^Y / \sigma_u^X)^2$. We have*

$$\frac{1}{3} \varepsilon_{n(k)}^{-1} \sum_{j=0}^{N_{t \wedge T_k}^{n(k)}} \kappa_{\hat{\tau}_j^k \wedge t} \mathbb{E}^{\hat{\mathbb{Q}}}\left[(X_{\hat{\tau}_{j+1}^k \wedge t} - X_{\hat{\tau}_j^k \wedge t})^3 | \mathcal{F}_{\hat{\tau}_j^k \wedge t}\right] - \varepsilon_{n(k)}^{-1} \left\langle Z^{n(k)}, Y \right\rangle_{t \wedge T_k} \rightarrow 0,$$

in probability as $k \rightarrow \infty$ for all $t \geq 0$.

Proof. By Itô's formula,

$$\frac{1}{3} \mathbb{E}^{\hat{\mathbb{Q}}}[(X_{\hat{\tau}_{j+1}^k \wedge t} - X_{\hat{\tau}_j^k \wedge t})^3 | \mathcal{F}_{\hat{\tau}_j^k \wedge t}] = \mathbb{E}^{\hat{\mathbb{Q}}}\left[\int_{\hat{\tau}_j^k \wedge t}^{\hat{\tau}_{j+1}^k \wedge t} (X_u - X_{\hat{\tau}_j^k}) d\langle X \rangle_u | \mathcal{F}_{\hat{\tau}_j^k \wedge t}\right].$$

We now show that

$$\begin{aligned} & \varepsilon_{n(k)}^{-1} \sum_{j=0}^{N_{t \wedge T_k}^{n(k)}} \kappa_{\hat{\tau}_j^k \wedge t} \mathbb{E}^{\hat{\mathbb{Q}}}\left[\int_{\hat{\tau}_j^k \wedge t}^{\hat{\tau}_{j+1}^k \wedge t} (X_u - X_{\hat{\tau}_j^k}) d\langle X \rangle_u | \mathcal{F}_{\hat{\tau}_j^k \wedge t}\right] \\ & - \varepsilon_{n(k)}^{-1} \sum_{j=0}^{N_{t \wedge T_k}^{n(k)}} \kappa_{\hat{\tau}_j^k \wedge t} \int_{\hat{\tau}_j^k \wedge t}^{\hat{\tau}_{j+1}^k \wedge t} (X_u - X_{\hat{\tau}_j^k}) d\langle X \rangle_u \rightarrow 0 \end{aligned} \quad (1.A.5)$$

and

$$\varepsilon_{n(k)}^{-1} \sum_{j=0}^{N_{t \wedge T_k}^{n(k)}} \kappa_{\hat{\tau}_j^k \wedge t} \int_{\hat{\tau}_j^k \wedge t}^{\hat{\tau}_{j+1}^k \wedge t} (X_u - X_{\hat{\tau}_j^k}) d\langle X \rangle_u - \varepsilon_{n(k)}^{-1} \langle Z^{n(k)}, Y \rangle_{t \wedge T_k} \rightarrow 0, \quad (1.A.6)$$

in probability.

By Lenglart inequality for discrete martingales (see e.g., Lemma A.2 of [Fuk11c]), a sufficient condition for (1.A.5) is the fact that

$$\varepsilon_{n(k)}^{-2} \sum_{j=0}^{N_{t \wedge T_k}^{n(k)}} \kappa_{\hat{\tau}_j^k \wedge t}^2 \mathbb{E}^{\hat{\mathbb{Q}}}\left[\left(\int_{\hat{\tau}_j^k \wedge t}^{\hat{\tau}_{j+1}^k \wedge t} (X_u - X_{\hat{\tau}_j^k}) d\langle X \rangle_u\right)^2 | \mathcal{F}_{\hat{\tau}_j^k \wedge t}\right] \rightarrow 0, \quad (1.A.7)$$

in probability. To get this convergence, first use successively Hölder inequality, Itô's formula and Burkholder-Davis-Gundy inequality to obtain that

$$\begin{aligned} & \sum_{j=0}^{N_{t \wedge T_k}^{n(k)}} \kappa_{\hat{\tau}_j^k \wedge t}^2 \mathbb{E}^{\hat{\mathbb{Q}}}\left[\left(\int_{\hat{\tau}_j^k \wedge t}^{\hat{\tau}_{j+1}^k \wedge t} (X_u - X_{\hat{\tau}_j^k}) d\langle X \rangle_u\right)^2 | \mathcal{F}_{\hat{\tau}_j^k \wedge t}\right] \\ & \leq C \sum_{j=0}^{N_{t \wedge T_k}^{n(k)}} \kappa_{\hat{\tau}_j^k \wedge t}^2 \mathbb{E}^{\hat{\mathbb{Q}}}\left[\left(\langle X \rangle_{\hat{\tau}_{j+1}^k \wedge t} - \langle X \rangle_{\hat{\tau}_j^k \wedge t}\right)^{3/2} \left(\int_{\hat{\tau}_j^k \wedge t}^{\hat{\tau}_{j+1}^k \wedge t} (X_u - X_{\hat{\tau}_j^k})^4 d\langle X \rangle_u\right)^{1/2} | \mathcal{F}_{\hat{\tau}_j^k \wedge t}\right] \\ & \leq C \left(\sum_{j=0}^{N_{t \wedge T_k}^{n(k)}} \mathbb{E}^{\hat{\mathbb{Q}}}\left[\left(\langle X \rangle_{\hat{\tau}_{j+1}^k \wedge t} - \langle X \rangle_{\hat{\tau}_j^k \wedge t}\right)^3 | \mathcal{F}_{\hat{\tau}_j^k \wedge t}\right]\right)^{1/2} \left(\sum_{j=0}^{N_{t \wedge T_k}^{n(k)}} \mathbb{E}^{\hat{\mathbb{Q}}}\left[\int_{\hat{\tau}_j^k \wedge t}^{\hat{\tau}_{j+1}^k \wedge t} (X_u - X_{\hat{\tau}_j^k})^4 d\langle X \rangle_u | \mathcal{F}_{\hat{\tau}_j^k \wedge t}\right]\right)^{1/2} \\ & = C \left(\sum_{j=0}^{N_{t \wedge T_k}^{n(k)}} \mathbb{E}^{\hat{\mathbb{Q}}}\left[\left(\langle X \rangle_{\hat{\tau}_{j+1}^k \wedge t} - \langle X \rangle_{\hat{\tau}_j^k \wedge t}\right)^3 | \mathcal{F}_{\hat{\tau}_j^k \wedge t}\right]\right)^{1/2} \left(\sum_{j=0}^{N_{t \wedge T_k}^{n(k)}} \mathbb{E}^{\hat{\mathbb{Q}}}\left[\left(X_{\hat{\tau}_{j+1}^k \wedge t} - X_{\hat{\tau}_j^k \wedge t}\right)^6 | \mathcal{F}_{\hat{\tau}_j^k \wedge t}\right]\right)^{1/2} \\ & \leq C \sum_{j=0}^{N_{t \wedge T_k}^{n(k)}} \mathbb{E}^{\hat{\mathbb{Q}}}\left[\left(\langle X \rangle_{\hat{\tau}_{j+1}^k \wedge t} - \langle X \rangle_{\hat{\tau}_j^k \wedge t}\right)^3 | \mathcal{F}_{\hat{\tau}_j^k \wedge t}\right] \\ & \leq C \sum_{j=0}^{N_{t \wedge T_k}^{n(k)}} \mathbb{E}^{\hat{\mathbb{Q}}}\left[|\hat{\tau}_{j+1}^k \wedge t - \hat{\tau}_j^k \wedge t|^3 | \mathcal{F}_{\hat{\tau}_j^k \wedge t}\right]. \end{aligned}$$

Note also that

$$\{j \leq N_{t \wedge T_k}^{n(k)}\} = \{\tau_j^{n(k)} \leq t \wedge T_k\} \in \mathcal{F}_{\hat{\tau}_j^k \wedge t}.$$

Then (1.A.7) follows since

$$\varepsilon_{n(k)}^{-2} \mathbb{E}^{\hat{\mathbb{Q}}} \left[\sum_{j=0}^{\infty} |\hat{\tau}_{j+1}^k \wedge t - \hat{\tau}_j^k \wedge t|^3 \right] \leq \varepsilon_{n(k)}^{-2} \frac{\varepsilon_{n(k)}^{8/3}}{k^2} \mathbb{E}^{\hat{\mathbb{Q}}} \left[\sum_{j=0}^{\infty} |\hat{\tau}_{j+1}^k \wedge t - \hat{\tau}_j^k \wedge t| \right] \leq \frac{\varepsilon_{n(k)}^{2/3}}{k^2} t \rightarrow 0.$$

We now turn to (1.A.6). Note that

$$\varepsilon_{n(k)}^{-1} \sum_{j=0}^{N_{t \wedge T_k}^{n(k)}} \kappa_{\hat{\tau}_j^k \wedge t} \int_{\hat{\tau}_j^k \wedge t}^{\hat{\tau}_{j+1}^k \wedge t} (X_u - X_{\hat{\tau}_j^k}) d \langle X \rangle_u = \varepsilon_{n(k)}^{-1} \sum_{j=0}^{\infty} \kappa_{\hat{\tau}_j^k \wedge t} \int_{\hat{\tau}_j^k \wedge t}^{\hat{\tau}_{j+1}^k \wedge t} (X_u - X_{\hat{\tau}_j^k}) d \langle X \rangle_u$$

and

$$\varepsilon_{n(k)}^{-1} \langle Z^{n(k)}, Y \rangle_{t \wedge T_k} = \varepsilon_{n(k)}^{-1} \sum_{j=0}^{\infty} \int_{\hat{\tau}_j^k \wedge t}^{\hat{\tau}_{j+1}^k \wedge t} (X_u - X_{\hat{\tau}_j^k}) \kappa_u d \langle X \rangle_u.$$

Therefore, the absolute value of the left hand side of (1.A.6) is dominated by

$$\begin{aligned} & \varepsilon_{n(k)}^{-1} \sum_{j=0}^{\infty} \int_{\hat{\tau}_j^k \wedge t}^{\hat{\tau}_{j+1}^k \wedge t} |X_u - X_{\hat{\tau}_j^k}| |\kappa_u - \kappa_{\hat{\tau}_j^k \wedge t}| d \langle X \rangle_u \\ & \leq \left(\varepsilon_{n(k)}^{-2} \sum_{j=0}^{\infty} \int_{\hat{\tau}_j^k \wedge t}^{\hat{\tau}_{j+1}^k \wedge t} |X_u - X_{\hat{\tau}_j^k}|^2 |\kappa_u - \kappa_{\hat{\tau}_j^k \wedge t}|^2 d \langle X \rangle_u \right)^{1/2} \left(\sum_{j=0}^{\infty} \int_{\hat{\tau}_j^k \wedge t}^{\hat{\tau}_{j+1}^k \wedge t} d \langle X \rangle_u \right)^{1/2} \\ & \leq \sup_{u \geq 0} \frac{1}{\kappa_{u \wedge T_k}} \sup_{u \geq 0, j \geq 0} |\kappa_{\hat{\tau}_{j+1}^k \wedge u} - \kappa_{\hat{\tau}_j^k \wedge u}| \varepsilon_{n(k)}^{-1} \langle Z^{n(k)} \rangle_{t \wedge T_k}^{1/2} \langle X \rangle_{t \wedge T_k}^{1/2}, \end{aligned}$$

which converges to 0 due to (1.A.4) and the uniform continuity of κ . \square

Lemma 1.A.2. *We have*

$$\frac{1}{6} \varepsilon_{n(k)}^{-2} \sum_{j=0}^{N_{t \wedge T_k}^{n(k)}} \kappa_{\hat{\tau}_j^k \wedge t} \mathbb{E}^{\hat{\mathbb{Q}}} [(X_{\hat{\tau}_{j+1}^k \wedge t} - X_{\hat{\tau}_j^k \wedge t})^4 | \mathcal{F}_{\hat{\tau}_j^k \wedge t}] - \varepsilon_{n(k)}^{-2} \langle Z^{n(k)} \rangle_{t \wedge T_k} \rightarrow 0,$$

in probability as $k \rightarrow \infty$, for all $t \geq 0$.

Proof. The proof is very similar to the previous one. By Itô's formula,

$$\frac{1}{6} \mathbb{E}^{\hat{\mathbb{Q}}} [(X_{\hat{\tau}_{j+1}^k \wedge t} - X_{\hat{\tau}_j^k \wedge t})^4 | \mathcal{F}_{\hat{\tau}_j^k \wedge t}] = \mathbb{E}^{\hat{\mathbb{Q}}} \left[\int_{\hat{\tau}_j^k \wedge t}^{\hat{\tau}_{j+1}^k \wedge t} (X_u - X_{\hat{\tau}_j^k})^2 d \langle X \rangle_u | \mathcal{F}_{\hat{\tau}_j^k \wedge t} \right].$$

We now show that

$$\begin{aligned} & \varepsilon_{n(k)}^{-2} \sum_{j=0}^{N_{t \wedge T_k}^{n(k)}} \kappa_{\hat{\tau}_j^k \wedge t} \mathbb{E}^{\hat{\mathbb{Q}}} \left[\int_{\hat{\tau}_j^k \wedge t}^{\hat{\tau}_{j+1}^k \wedge t} (X_u - X_{\hat{\tau}_j^k})^2 d \langle X \rangle_u | \mathcal{F}_{\hat{\tau}_j^k \wedge t} \right] \\ & - \varepsilon_{n(k)}^{-2} \sum_{j=0}^{N_{t \wedge T_k}^{n(k)}} \kappa_{\hat{\tau}_j^k \wedge t} \int_{\hat{\tau}_j^k \wedge t}^{\hat{\tau}_{j+1}^k \wedge t} (X_u - X_{\hat{\tau}_j^k})^2 d \langle X \rangle_u \rightarrow 0 \end{aligned} \tag{1.A.8}$$

and

$$\varepsilon_{n(k)}^{-2} \sum_{j=0}^{N_{t \wedge T_k}^{n(k)}} \kappa_{\hat{\tau}_j^k \wedge t} \int_{\hat{\tau}_j^k \wedge t}^{\hat{\tau}_{j+1}^k \wedge t} (X_u - X_{\hat{\tau}_j^k})^2 d \langle X \rangle_u - \varepsilon_{n(k)}^{-2} \langle Z^{n(k)} \rangle_{t \wedge T_k} \rightarrow 0, \tag{1.A.9}$$

in probability.

By Lenglart inequality for discrete martingales, a sufficient condition for (1.A.8) is

$$\varepsilon_{n(k)}^{-4} \sum_{j=0}^{N_{t \wedge T_k}^{n(k)}} \kappa_{\hat{\tau}_j^k \wedge t}^2 \mathbb{E}^{\hat{\mathbb{Q}}} \left[\left(\int_{\hat{\tau}_j^k \wedge t}^{\hat{\tau}_{j+1}^k \wedge t} (X_u - X_{\hat{\tau}_j^k})^2 d\langle X \rangle_u \right)^2 \middle| \mathcal{F}_{\hat{\tau}_j^k \wedge t} \right] \rightarrow 0, \quad (1.A.10)$$

in probability. To get this convergence, first use successively Hölder inequality, Itô's formula and Burkholder-Davis-Gundy inequality to obtain that

$$\begin{aligned} & \sum_{j=0}^{N_{t \wedge T_k}^{n(k)}} \kappa_{\hat{\tau}_j^k \wedge t}^2 \mathbb{E}^{\hat{\mathbb{Q}}} \left[\left(\int_{\hat{\tau}_j^k \wedge t}^{\hat{\tau}_{j+1}^k \wedge t} (X_u - X_{\hat{\tau}_j^k})^2 d\langle X \rangle_u \right)^2 \middle| \mathcal{F}_{\hat{\tau}_j^k \wedge t} \right] \\ & \leq \sum_{j=0}^{N_{t \wedge T_k}^{n(k)}} \kappa_{\hat{\tau}_j^k \wedge t}^2 \mathbb{E}^{\hat{\mathbb{Q}}} \left[(\langle X \rangle_{\hat{\tau}_{j+1}^k \wedge t} - \langle X \rangle_{\hat{\tau}_j^k \wedge t})^{4/3} \left(\int_{\hat{\tau}_j^k \wedge t}^{\hat{\tau}_{j+1}^k \wedge t} (X_u - X_{\hat{\tau}_j^k})^6 d\langle X \rangle_u \right)^{2/3} \middle| \mathcal{F}_{\hat{\tau}_j^k \wedge t} \right] \\ & \leq C \left(\sum_{j=0}^{N_{t \wedge T_k}^{n(k)}} \mathbb{E}^{\hat{\mathbb{Q}}} \left[(\langle X \rangle_{\hat{\tau}_{j+1}^k \wedge t} - \langle X \rangle_{\hat{\tau}_j^k \wedge t})^4 \middle| \mathcal{F}_{\hat{\tau}_j^k \wedge t} \right] \right)^{1/3} \left(\sum_{j=0}^{N_{t \wedge T_k}^{n(k)}} \mathbb{E}^{\hat{\mathbb{Q}}} \left[\int_{\hat{\tau}_j^k \wedge t}^{\hat{\tau}_{j+1}^k \wedge t} (X_u - X_{\hat{\tau}_j^k})^6 d\langle X \rangle_u \middle| \mathcal{F}_{\hat{\tau}_j^k \wedge t} \right] \right)^{2/3} \\ & = C \left(\sum_{j=0}^{N_{t \wedge T_k}^{n(k)}} \mathbb{E}^{\hat{\mathbb{Q}}} \left[(\langle X \rangle_{\hat{\tau}_{j+1}^k \wedge t} - \langle X \rangle_{\hat{\tau}_j^k \wedge t})^4 \middle| \mathcal{F}_{\hat{\tau}_j^k \wedge t} \right] \right)^{1/3} \left(\sum_{j=0}^{N_{t \wedge T_k}^{n(k)}} \mathbb{E}^{\hat{\mathbb{Q}}} \left[(X_{\hat{\tau}_{j+1}^k \wedge t} - X_{\hat{\tau}_j^k \wedge t})^8 \middle| \mathcal{F}_{\hat{\tau}_j^k \wedge t} \right] \right)^{2/3} \\ & \leq C \sum_{j=0}^{N_{t \wedge T_k}^{n(k)}} \mathbb{E}^{\hat{\mathbb{Q}}} \left[(\langle X \rangle_{\hat{\tau}_{j+1}^k \wedge t} - \langle X \rangle_{\hat{\tau}_j^k \wedge t})^4 \middle| \mathcal{F}_{\hat{\tau}_j^k \wedge t} \right] \\ & \leq C \sum_{j=0}^{N_{t \wedge T_k}^{n(k)}} \mathbb{E}^{\hat{\mathbb{Q}}} \left[|\hat{\tau}_{j+1}^k \wedge t - \hat{\tau}_j^k \wedge t|^4 \middle| \mathcal{F}_{\hat{\tau}_j^k \wedge t} \right]. \end{aligned}$$

Then, observe that

$$\varepsilon_{n(k)}^{-4} \mathbb{E}^{\hat{\mathbb{Q}}} \left[\sum_{j=0}^{\infty} |\hat{\tau}_{j+1}^k \wedge t - \hat{\tau}_j^k \wedge t|^4 \right] \leq \frac{t}{k^3} \rightarrow 0,$$

which gives (1.A.8). The proof for (1.A.9) is obtained in the same way as that for (1.A.6). \square

We finally give the following almost straightforward result, which is easily deduced from simplified versions of the proofs of the previous lemma.

Lemma 1.A.3. *We have*

$$\sum_{j=0}^{N_{t \wedge T_k}^{n(k)}} \kappa_{\hat{\tau}_j^k \wedge t} \mathbb{E}^{\hat{\mathbb{Q}}} \left[(X_{\hat{\tau}_{j+1}^k \wedge t} - X_{\hat{\tau}_j^k \wedge t})^2 \middle| \mathcal{F}_{\hat{\tau}_j^k \wedge t} \right] - \langle Y \rangle_{t \wedge T_k} \rightarrow 0,$$

in probability as $k \rightarrow \infty$ for all $t \geq 0$.

We are now ready to complete the proof of Proposition 1.2.1. From (1.A.4) and Lemmas 1.A.1,

1.A.2 and 1.A.3, we have the following convergences in probability as $k \rightarrow \infty$ for all $0 \leq v \leq t$:

$$\begin{aligned} & \sum_{j=N_{v \wedge T_k}^{n(k)}+1}^{N_{t \wedge T_k}^{n(k)}} \kappa_{\hat{\tau}_j^k \wedge t} \mathbb{E}^{\hat{\mathbb{Q}}}[(X_{\hat{\tau}_{j+1}^k \wedge t} - X_{\hat{\tau}_j^k \wedge t})^2 | \mathcal{F}_{\hat{\tau}_j^k \wedge t}] \rightarrow \int_{v \wedge \hat{\alpha}^K}^{t \wedge \hat{\alpha}^K} (\sigma_y^Y)^2 du, \\ \varepsilon_{n(k)}^{-1} & \sum_{j=N_{v \wedge T_k}^{n(k)}+1}^{N_{t \wedge T_k}^{n(k)}} \kappa_{\hat{\tau}_j^k \wedge t} \mathbb{E}^{\hat{\mathbb{Q}}}[(X_{\hat{\tau}_{j+1}^k \wedge t} - X_{\hat{\tau}_j^k \wedge t})^3 | \mathcal{F}_{\hat{\tau}_j^k \wedge t}] \rightarrow \int_{v \wedge \hat{\alpha}^K}^{t \wedge \hat{\alpha}^K} s_u (\sigma_y^Y)^2 du, \\ \varepsilon_{n(k)}^{-2} & \sum_{j=N_{v \wedge T_k}^{n(k)}+1}^{N_{t \wedge T_k}^{n(k)}} \kappa_{\hat{\tau}_j^k \wedge t} \mathbb{E}^{\hat{\mathbb{Q}}}[(X_{\hat{\tau}_{j+1}^k \wedge t} - X_{\hat{\tau}_j^k \wedge t})^4 | \mathcal{F}_{\hat{\tau}_j^k \wedge t}] \rightarrow \int_{v \wedge \hat{\alpha}^K}^{t \wedge \hat{\alpha}^K} a_u^2 (\sigma_y^Y)^2 du. \end{aligned}$$

Since

$$(\mathbb{E}^{\hat{\mathbb{Q}}}[(X_{\hat{\tau}_{j+1}^k \wedge t} - X_{\hat{\tau}_j^k \wedge t})^3 | \mathcal{F}_{\hat{\tau}_j^k \wedge t}])^2 \leq \mathbb{E}^{\hat{\mathbb{Q}}}[(X_{\hat{\tau}_{j+1}^k \wedge t} - X_{\hat{\tau}_j^k \wedge t})^2 | \mathcal{F}_{\hat{\tau}_j^k \wedge t}] \mathbb{E}^{\hat{\mathbb{Q}}}[(X_{\hat{\tau}_{j+1}^k \wedge t} - X_{\hat{\tau}_j^k \wedge t})^4 | \mathcal{F}_{\hat{\tau}_j^k \wedge t}],$$

we have

$$\begin{aligned} & \left(\varepsilon_{n(k)}^{-1} \sum_{j=N_{v \wedge T_k}^{n(k)}+1}^{N_{t \wedge T_k}^{n(k)}} \kappa_{\hat{\tau}_j^k} \mathbb{E}^{\hat{\mathbb{Q}}}[(X_{\hat{\tau}_{j+1}^k \wedge t} - X_{\hat{\tau}_j^k \wedge t})^3 | \mathcal{F}_{\hat{\tau}_j^k \wedge t}] \right)^2 \\ & \leq \varepsilon_{n(k)}^{-2} \sum_{j=N_{v \wedge T_k}^{n(k)}+1}^{N_{t \wedge T_k}^{n(k)}} \kappa_{\hat{\tau}_j^k} \mathbb{E}^{\hat{\mathbb{Q}}}[(X_{\hat{\tau}_{j+1}^k \wedge t} - X_{\hat{\tau}_j^k \wedge t})^4 | \mathcal{F}_{\hat{\tau}_j^k \wedge t}] \sum_{j=N_{v \wedge T_k}^{n(k)}+1}^{N_{t \wedge T_k}^{n(k)}} \kappa_{\hat{\tau}_j^k} \mathbb{E}^{\hat{\mathbb{Q}}}[(X_{\hat{\tau}_{j+1}^k \wedge t} - X_{\hat{\tau}_j^k \wedge t})^2 | \mathcal{F}_{\hat{\tau}_j^k \wedge t}]. \end{aligned}$$

This implies that for all $0 \leq v \leq t$,

$$\left(\int_{v \wedge \hat{\alpha}^K}^{t \wedge \hat{\alpha}^K} s_u (\sigma_u^Y)^2 du \right)^2 \leq \int_{v \wedge \hat{\alpha}^K}^{t \wedge \hat{\alpha}^K} a_u^2 (\sigma_u^Y)^2 du \int_{v \wedge \hat{\alpha}^K}^{t \wedge \hat{\alpha}^K} (\sigma_u^Y)^2 du.$$

Thus we obtain (1.A.3).

1.A.2 Proof of Proposition 1.2.2

In this proof, using a classical localization procedure together with Girsanov theorem, we can assume that $b^X = 0$ and that σ^X and σ^Y are bounded on $[0, T]$. We start with two technical lemmas and their proof.

Lemma 1.A.4. *For any $p \in [0, 2)$,*

$$\varepsilon_n^{-p} \sup_{j \geq 0} (\langle X \rangle_{\tau_{j+1}^n} - \langle X \rangle_{\tau_j^n}) \rightarrow 0,$$

in probability.

Proof. Let $K > 0$ and

$$\gamma_K^n = \inf\{t > 0; \varepsilon_n^{-1} |X_t - X_t^n| \geq K\} \wedge T. \quad (1.A.11)$$

Using the tightness of the family (1.2.9), we get

$$\lim_{K \rightarrow \infty} \sup_{n \in \mathbb{N}} \mathbb{P}[\gamma_K^n < T] = 0. \quad (1.A.12)$$

Therefore, it is enough to show that for any $K > 0$,

$$\varepsilon_n^{-p} \sup_{j \geq 0} (\langle X \rangle_{\tau_{j+1}^n \wedge \gamma_K^n} - \langle X \rangle_{\tau_j^n \wedge \gamma_K^n}) \rightarrow 0,$$

in probability. Take an integer $m > 2/(2-p)$. Since

$$\sup_{j \geq 0} (\langle X \rangle_{\tau_{j+1}^n \wedge \gamma_K^n} - \langle X \rangle_{\tau_j^n \wedge \gamma_K^n}) \leq \left(\sum_{j=0}^{\infty} (\langle X \rangle_{\tau_{j+1}^n \wedge \gamma_K^n} - \langle X \rangle_{\tau_j^n \wedge \gamma_K^n})^m \right)^{1/m},$$

the statement of the lemma follows from the fact that

$$\begin{aligned} \mathbb{E} \left[\sum_{j=0}^{\infty} (\langle X \rangle_{\tau_{j+1}^n \wedge \gamma_K^n} - \langle X \rangle_{\tau_j^n \wedge \gamma_K^n})^m \right] &\leq C \mathbb{E} \left[\sum_{j=0}^{\infty} \left(\sup_{t \geq 0} |X_{\tau_{j+1}^n \wedge \gamma_K^n \wedge t} - X_{\tau_j^n \wedge \gamma_K^n \wedge t}| \right)^{2m} \right] \\ &\leq C \varepsilon_n^{2m-2} \mathbb{E} \left[\sum_{j=0}^{\infty} \left(\sup_{t \geq 0} |X_{\tau_{j+1}^n \wedge \gamma_K^n \wedge t} - X_{\tau_j^n \wedge \gamma_K^n \wedge t}| \right)^2 \right] \\ &\leq C \varepsilon_n^{2m-2} \mathbb{E} \left[\sum_{j=0}^{\infty} (\langle X \rangle_{\tau_{j+1}^n \wedge \gamma_K^n} - \langle X \rangle_{\tau_j^n \wedge \gamma_K^n}) \right] \\ &\leq C \varepsilon_n^{2m-2}. \end{aligned}$$

Here we have used that $\mathbb{E}[\langle X \rangle_T] < \infty$. The result follows using Hölder inequality. \square

Lemma 1.A.5. *For any $p \in [0, 2)$ and $T_0 \in [0, T)$,*

$$\varepsilon_n^{-p} \sup_{j \geq 0} (\tau_{j+1}^n \wedge T_0 - \tau_j^n \wedge T_0) \rightarrow 0,$$

in probability. In particular, the convergence in probability (1.2.8) holds for all $T_0 \in [0, T)$.

Proof. Let $T_0 \in [0, T)$, $K > 0$ and

$$\hat{\gamma}_K = \inf\{t > 0; \sigma_t^X \leq 1/K\} \wedge T_0. \quad (1.A.13)$$

Using the continuity and the positivity of σ^X on $[0, T)$, we get

$$\lim_{K \rightarrow \infty} \mathbb{P}[\hat{\gamma}_K < T_0] = 0.$$

Therefore, it is enough to show that for any $K > 0$,

$$\varepsilon_n^{-p} \sup_{j \geq 0} (\tau_{j+1}^n \wedge \hat{\gamma}_K - \tau_j^n \wedge \hat{\gamma}_K) \rightarrow 0,$$

in probability. This follows from Lemma 1.A.4 since

$$\sup_{j \geq 0} (\tau_{j+1}^n \wedge \hat{\gamma}_K - \tau_j^n \wedge \hat{\gamma}_K) \leq C \sup_{j \geq 0} (\langle X \rangle_{\tau_{j+1}^n} - \langle X \rangle_{\tau_j^n}).$$

\square

We now give the end of the proof of Proposition 1.2.2. Define γ_K^n by (1.A.11). It suffices to show that for any $K > 0$

$$\begin{aligned} \sup_{t \geq 0} \left| \varepsilon_n^{-2} \langle Z^n \rangle_{t \wedge \gamma_K^n} - \frac{1}{6} \int_0^{t \wedge \gamma_K^n} a_u^2(\sigma_u^Y)^2 du \right| &\rightarrow 0, \\ \sup_{t \geq 0} \left| \varepsilon_n^{-1} \langle Z^n, Y \rangle_{t \wedge \gamma_K^n} - \frac{1}{3} \int_0^{t \wedge \gamma_K^n} s_u(\sigma_u^Y)^2 du \right| &\rightarrow 0, \end{aligned}$$

in probability. Since

$$\varepsilon_n^{-1} \sup_{t \geq 0} |X_{t \wedge \gamma_K^n} - X_{t \wedge \gamma_K^n}^n| \leq K, \quad (1.A.14)$$

the families $\varepsilon_n^{-2} \langle Z^n \rangle_{\cdot \wedge \gamma_K^n}$ and $\varepsilon_n^{-1} \langle Z^n, Y \rangle_{\cdot \wedge \gamma_K^n}$ are equicontinuous. So we just need to prove that for any $t \in [0, T)$,

$$\begin{aligned} \varepsilon_n^{-2} \langle Z^n \rangle_{t \wedge \gamma_K^n} - \frac{1}{6} \int_0^{t \wedge \gamma_K^n} a_u^2(\sigma_u^Y)^2 du &\rightarrow 0, \\ \varepsilon_n^{-1} \langle Z^n, Y \rangle_{t \wedge \gamma_K^n} - \frac{1}{3} \int_0^{t \wedge \gamma_K^n} s_u(\sigma_u^Y)^2 du &\rightarrow 0, \end{aligned}$$

in probability. Let

$$\beta_M^n = \inf\{u > 0; (1/\sigma_u^X) \geq M\} \wedge t \wedge \gamma_K^n$$

for $M > 0$. Since $t < T$, and σ^X is a.s. positive on $[0, T)$,

$$\lim_{M \rightarrow \infty} \sup_{n \in \mathbb{N}} \mathbb{P}[\beta_M^n < t \wedge \gamma_K^n] = 0.$$

Therefore, it is enough to show that for any $M > 0$,

$$\begin{aligned} \varepsilon_n^{-2} \langle Z^n \rangle_{\beta_M^n} - \frac{1}{6} \int_0^{\beta_M^n} a_u^2(\sigma_u^Y)^2 du &\rightarrow 0, \\ \varepsilon_n^{-1} \langle Z^n, Y \rangle_{\beta_M^n} - \frac{1}{3} \int_0^{\beta_M^n} s_u(\sigma_u^Y)^2 du &\rightarrow 0, \end{aligned}$$

in probability. From the assumptions of Proposition 1.2.2, we have

$$\begin{aligned} \varepsilon_n^{-1} \sum_{j=0}^{N_{\beta_M^n}^n} \kappa_{\tau_j^n} \mathbb{E}[\Delta_{j,n}^3 | \mathcal{F}_{\tau_j^n}] + \int_0^{\beta_M^n} s_u(\sigma_u^Y)^2 du &\rightarrow 0, \\ \varepsilon_n^{-2} \sum_{j=0}^{N_{\beta_M^n}^n} \kappa_{\tau_j^n} \mathbb{E}[\Delta_{j,n}^4 | \mathcal{F}_{\tau_j^n}] - \int_0^{\beta_M^n} a_u^2(\sigma_u^Y)^2 du &\rightarrow 0, \end{aligned}$$

in probability. Moreover, by Itô's formula,

$$\begin{aligned} \frac{1}{3} \varepsilon_n^{-1} \sum_{j=0}^{N_{\beta_M^n}^n} \kappa_{\tau_j^n} \mathbb{E}[\Delta_{j,n}^3 | \mathcal{F}_{\tau_j^n}] &= \varepsilon_n^{-1} \sum_{j=0}^{N_{\beta_M^n}^n} \kappa_{\tau_j^n} \mathbb{E}\left[\int_{\tau_j^n}^{\tau_{j+1}^n} (X_u - X_{\tau_j^n}^n) d\langle X \rangle_u \mid \mathcal{F}_{\tau_j^n}\right], \\ \frac{1}{6} \varepsilon_n^{-2} \sum_{j=0}^{N_{\beta_M^n}^n} \kappa_{\tau_j^n} \mathbb{E}[\Delta_{j,n}^4 | \mathcal{F}_{\tau_j^n}] &= \varepsilon_n^{-2} \sum_{j=0}^{N_{t \wedge \beta_M^n}^n} \kappa_{\tau_j^n} \mathbb{E}\left[\int_{\tau_j^n}^{\tau_{j+1}^n} (X_u - X_{\tau_j^n}^n)^2 d\langle X \rangle_u \mid \mathcal{F}_{\tau_j^n}\right]. \end{aligned}$$

Now remark that the following convergences in probability hold:

$$\begin{aligned} \sup_{t \geq 0} \left| \varepsilon_n^{-1} \sum_{j=0}^{N_{\beta_M^n}^n} \kappa_{\tau_j^n} \int_{\tau_j^n}^{\tau_{j+1}^n} (X_u - X_{\tau_j^n}^n) d\langle X \rangle_u - \varepsilon_n^{-1} \sum_{j=0}^{N_{\beta_M^n}^n} \kappa_{\tau_j^n} \mathbb{E}\left[\int_{\tau_j^n}^{\tau_{j+1}^n} (X_u - X_{\tau_j^n}^n) d\langle X \rangle_u \mid \mathcal{F}_{\tau_j^n}\right] \right| &\rightarrow 0, \\ \sup_{t \geq 0} \left| \varepsilon_n^{-2} \sum_{j=0}^{N_{\beta_M^n}^n} \kappa_{\tau_j^n} \int_{\tau_j^n}^{\tau_{j+1}^n} (X_u - X_{\tau_j^n}^n)^2 d\langle X \rangle_u - \varepsilon_n^{-2} \sum_{j=0}^{N_{\beta_M^n}^n} \kappa_{\tau_j^n} \mathbb{E}\left[\int_{\tau_j^n}^{\tau_{j+1}^n} (X_u - X_{\tau_j^n}^n)^2 d\langle X \rangle_u \mid \mathcal{F}_{\tau_j^n}\right] \right| &\rightarrow 0. \end{aligned} \quad (1.A.15)$$

Indeed, as seen in the proofs of Lemmas 1.A.1 and 1.A.2, the convergences in probability in (1.A.15) are deduced from the following ones:

$$\begin{aligned} \varepsilon_n^{-2} \sum_{j=0}^{N_{\beta_M^n}^n} \kappa_{\tau_j^n}^2 \mathbb{E} \left[\left(\int_{\tau_j^n}^{\tau_{j+1}^n} (X_u - X_{\tau_j^n}^n) d\langle X \rangle_u \right)^2 \middle| \mathcal{F}_{\tau_j^n} \right] &\rightarrow 0, \\ \varepsilon_n^{-4} \sum_{j=0}^{N_{\beta_M^n}^n} \kappa_{\tau_j^n}^2 \mathbb{E} \left[\left(\int_{\tau_j^n}^{\tau_{j+1}^n} (X_u - X_{\tau_j^n}^n)^2 d\langle X \rangle_u \right)^2 \middle| \mathcal{F}_{\tau_j^n} \right] &\rightarrow 0. \end{aligned} \quad (1.A.16)$$

Since $Q_n = \varepsilon_n^{-4} \sup_{t \in [0, T]} |X_t^n - X_t|^4$ is uniformly integrable and

$$\sum_{j=0}^{\infty} (\langle X \rangle_{\tau_{j+1}^n} - \langle X \rangle_{\tau_j^n})^2$$

is bounded and converges to 0 in probability by Lemma 1.A.4, we have

$$\mathbb{E} \left[\sum_{j=0}^{N_{\beta_M^n}^n} \kappa_{\tau_j^n}^2 \left(\int_{\tau_j^n}^{\tau_{j+1}^n} \varepsilon_n^{-k} (X_u - X_{\tau_j^n}^n)^k d\langle X \rangle_u \right)^2 \right] \leq C \mathbb{E} \left[Q_n^{k/2} \sum_{j=0}^{\infty} (\langle X \rangle_{\tau_{j+1}^n} - \langle X \rangle_{\tau_j^n})^2 \right] \rightarrow 0$$

for $k = 1, 2$, which gives (1.A.16).

We also have

$$\begin{aligned} \sup_{t \geq 0} \left| \varepsilon_n^{-1} \sum_{j=0}^{N_{\beta_M^n}^n} \kappa_{\tau_j^n} \int_{\tau_j^n}^{\tau_{j+1}^n} (X_u - X_{\tau_j^n}^n) d\langle X \rangle_u - \varepsilon_n^{-1} \sum_{j=0}^{\infty} \kappa_{\tau_j^n} \int_{\tau_j^n \wedge \beta_M^n}^{\tau_{j+1}^n \wedge \beta_M^n} (X_u - X_{\tau_j^n}^n) d\langle X \rangle_u \right| &\rightarrow 0, \\ \sup_{t \geq 0} \left| \varepsilon_n^{-2} \sum_{j=0}^{N_{\beta_M^n}^n} \kappa_{\tau_j^n} \int_{\tau_j^n}^{\tau_{j+1}^n} (X_u - X_{\tau_j^n}^n)^2 d\langle X \rangle_u - \varepsilon_n^{-2} \sum_{j=0}^{\infty} \kappa_{\tau_j^n} \int_{\tau_j^n \wedge \beta_M^n}^{\tau_{j+1}^n \wedge \beta_M^n} (X_u - X_{\tau_j^n}^n)^2 d\langle X \rangle_u \right| &\rightarrow 0, \end{aligned} \quad (1.A.17)$$

in probability. These two convergences follow using that

$$\sup_{j \geq 0, t \in [0, T]} \int_{\tau_j^n \wedge t}^{\tau_{j+1}^n \wedge t} \varepsilon_n^{-i} |X_u - X_{\tau_j^n}^n|^i d\langle X \rangle_u \rightarrow 0$$

in probability for $i = 1, 2$, which is deduced from (1.A.12) and the fact that

$$\sup_{j \geq 0, t \geq 0} \int_{\tau_j^n \wedge \gamma_K^n \wedge t}^{\tau_{j+1}^n \wedge \gamma_K^n \wedge t} \varepsilon_n^{-i} |X_u - X_{\tau_j^n}^n|^i d\langle X \rangle_u \leq K^i \sup_{j \geq 0} (\langle X \rangle_{\tau_{j+1}^n} - \langle X \rangle_{\tau_j^n}) \rightarrow 0,$$

in probability, by Lemma 1.A.4.

Finally, remark that the uniform continuity of κ and (1.A.14) imply

$$\begin{aligned} \sup_{t \geq 0} \left| \varepsilon_n^{-1} \sum_{j=0}^{\infty} \kappa_{\tau_j^n} \int_{\tau_j^n \wedge \beta_M^n}^{\tau_{j+1}^n \wedge \beta_M^n} (X_u - X_{\tau_j^n}^n) d\langle X \rangle_u + \varepsilon_n^{-1} \langle Z^n, Y \rangle_{\beta_M^n} \right| &\rightarrow 0, \\ \sup_{t \geq 0} \left| \varepsilon_n^{-2} \sum_{j=0}^{\infty} \kappa_{\tau_j^n} \int_{\tau_j^n \wedge \beta_M^n}^{\tau_{j+1}^n \wedge \beta_M^n} (X_u - X_{\tau_j^n}^n)^2 d\langle X \rangle_u - \varepsilon_n^{-2} \langle Z^n \rangle_{\beta_M^n} \right| &\rightarrow 0, \end{aligned} \quad (1.A.18)$$

in probability. Then Proposition 1.2.2 is eventually obtained from (1.A.15) together with (1.A.17) and (1.A.18).

1.A.3 Proof of Proposition 1.2.3

Proof of the convergence in law to (1.2.3)

We start with the stable convergence in law of the renormalized hedging error. Such convergence being stable against localization procedures, we can assume without loss of generality that $|b^X|$, σ^X , $|b^Y|$, σ^Y , $1/\sigma^Y$, \bar{l} , $1/\bar{l}$, \underline{l} and $1/\underline{l}$ are bounded by a constant $M > 0$. Then in particular we have $\varepsilon_n^{-1} \sup_{t \in [0, T]} |X_t^n - X_t| \leq M$.

Fix $T_0 \in [0, T]$ and define $\hat{\alpha}^K$ by (1.A.1). Then we have (1.A.2) and so, we can assume without loss of generality that $1/\sigma^X \leq K$ in order to show the convergences (1.2.5) and (1.2.6) on $[0, T_0]$. Also, thanks to the Girsanov-Maruyama transformation, we can assume $b^X = 0$. In order to apply Proposition 1.2.2, it remains to show (1.2.10).

Part 1: Technical lemma

We give here a first technical lemma.

Lemma 1.A.6. *The sequence $\varepsilon_n^2 N_{T_0}^n$ is tight.*

Proof. Since for $j < N_{T_0}^n$,

$$|X_{\tau_{j+1}^n \wedge T_0} - X_{\tau_j^n \wedge T_0}|^2 \geq \frac{\varepsilon_n^2}{K^2},$$

we have

$$\varepsilon_n^2 N_{T_0}^n \leq K^2 \sum_{j=0}^{N_{T_0}^n - 1} (X_{\tau_{j+1}^n \wedge T_0} - X_{\tau_j^n \wedge T_0})^2 \rightarrow K^2 \langle X \rangle_{T_0},$$

in probability by Lemma 1.A.5. □

Part 2: Approximation lemma

Let $\tilde{\tau}_{j+1}^n$ be the exit time of fixed barriers defined by

$$\tilde{\tau}_{j+1}^n = \inf \{t > \tau_j^n : X_t \notin (X_{\tau_j^n} - \varepsilon_n \underline{l}_{\tau_j^n}, X_{\tau_j^n} + \varepsilon_n \bar{l}_{\tau_j^n})\} \wedge T_0. \quad (1.A.19)$$

We have the following lemma.

Lemma 1.A.7. *We have*

$$\sum_{j=0}^{N_{T_0}^n} \mathbb{E}[|\tilde{\tau}_{j+1}^n - \tau_{j+1}^n| | \mathcal{F}_{\tau_j^n}] \rightarrow 0,$$

in probability.

Proof. Define for $\delta > 0$ and $t \in [0, T_0]$

$$w_t(\delta) = \sup\{|\bar{l}_u - \bar{l}_v| + |\underline{l}_u - \underline{l}_v|; 0 \leq u \leq t, 0 \leq v \leq t, |u - v| \leq \delta\}.$$

Since \bar{l} and \underline{l} are continuous and bounded, we have

$$\mathbb{E}[w_{T_0}(\delta)] \rightarrow 0$$

as $\delta \rightarrow 0$. For $m \in \mathbb{N}$, let

$$T_m^n = \inf\{t > 0; w_t(\varepsilon_n) \geq m \mathbb{E}[w_{T_0}(\varepsilon_n)]\} \wedge T_0.$$

For the proof of, the lemma, we can estimate the probability as follows:

$$\begin{aligned} & \mathbb{P}\left[\sum_{j=0}^{N_{T_0}^n} \mathbb{E}[|\tilde{\tau}_{j+1}^n - \tau_{j+1}^n| \mid \mathcal{F}_{\tau_j^n}] \geq \omega\right] \\ & \leq \mathbb{P}\left[\sum_{j=0}^{N_{T_0}^n} \mathbb{E}[|\tilde{\tau}_{j+1}^n - \tau_{j+1}^n| \mid \mathcal{F}_{\tau_j^n}] \geq \omega, T_m^n = T_0\right] + \mathbb{P}[T_m^n < T_0]. \end{aligned}$$

Since $w_t(\varepsilon_n)$ is increasing in t ,

$$\sup_{n \in \mathbb{N}} \mathbb{P}[T_m^n < T_0] \leq \sup_{n \in \mathbb{N}} \mathbb{P}[w_{T_0}(\varepsilon_n) \geq m \mathbb{E}[w_{T_0}(\varepsilon_n)]] \leq \frac{1}{m} \rightarrow 0$$

as $m \rightarrow \infty$. On the set $\{T_m^n = T_0\}$, we can replace \bar{l} and \underline{l} by $\bar{l}_{\cdot \wedge T_m^n}$ and $\underline{l}_{\cdot \wedge T_m^n}$, respectively, which means that without loss of generality it is enough to prove the lemma under the assumption that $w_{T_0}(\varepsilon_n) \leq m \mathbb{E}[w_{T_0}(\varepsilon_n)]$.

Since the sequence $\varepsilon_n^2 N_{T_0}^n$ is tight, it is enough to show that

$$\frac{1}{\varepsilon_n^2} \sup_{j \leq N_{T_0}^n} \mathbb{E}[|\tilde{\tau}_{j+1}^n - \tau_{j+1}^n| \mid \mathcal{F}_{\tau_j^n}] \rightarrow 0.$$

We write

$$\frac{1}{\varepsilon_n^2} \sup_{j \leq N_{T_0}^n} \mathbb{E}[|\tilde{\tau}_{j+1}^n - \tau_{j+1}^n| \mid \mathcal{F}_{\tau_j^n}] = R_1 + R_2,$$

with

$$\begin{aligned} R_1 &= \frac{1}{\varepsilon_n^2} \sup_{j \leq N_{T_0}^n} \mathbb{E}[|\tilde{\tau}_{j+1}^n - \tau_{j+1}^n| \mathbb{1}_{\{\tilde{\tau}_{j+1}^n \vee \tau_{j+1}^n \geq \tau_j^n + \varepsilon_n\}} \mid \mathcal{F}_{\tau_j^n}], \\ R_2 &= \frac{1}{\varepsilon_n^2} \sup_{j \leq N_{T_0}^n} \mathbb{E}[|\tilde{\tau}_{j+1}^n - \tau_{j+1}^n| \mathbb{1}_{\{\tilde{\tau}_{j+1}^n \vee \tau_{j+1}^n < \tau_j^n + \varepsilon_n\}} \mid \mathcal{F}_{\tau_j^n}]. \end{aligned}$$

We first treat R_1 . We have

$$R_1 \leq \frac{T}{\varepsilon_n^2} \sup_{j \leq N_{T_0}^n} \mathbb{P}[\tilde{\tau}_{j+1}^n \vee \tau_{j+1}^n \geq \tau_j^n + \varepsilon_n \mid \mathcal{F}_{\tau_j^n}].$$

Since \bar{l} , \underline{l} and bounded by M and σ^X is bounded from below by $1/K$, using the Dambis, Dubins-Schwartz theorem we get

$$\mathbb{P}[\tilde{\tau}_{j+1}^n \vee \tau_{j+1}^n \geq \tau_j^n + \varepsilon_n \mid \mathcal{F}_{\tau_j^n}] \leq \mathbb{P}[\rho^n \geq \varepsilon_n / K^2],$$

with ρ^n being the first exit time of $[-\varepsilon_n M, \varepsilon_n M]$ by a Brownian motion starting from zero. Since ρ^n admits an exponential moment (see e.g., Exercice 3.10 in [RY99]), from the scaling property of the Brownian motion we have that $\mathbb{E}[(\rho^n)^k] \leq C \varepsilon_n^{2k}$ for $k \in \mathbb{N}$. Markov's inequality then gives the convergence to zero of R_1 .

We now turn to R_2 . Recall that $w_{T_0}(\varepsilon_n) \leq m \mathbb{E}[w_{T_0}(\varepsilon_n)] =: \delta_n \rightarrow 0$. Then, we have

$$|\tilde{\tau}_{j+1}^n - \tau_{j+1}^n| \mathbb{1}_{\{\tilde{\tau}_{j+1}^n \vee \tau_{j+1}^n < \tau_j^n + \varepsilon_n\}} \leq \hat{J}_{j+1}^n - \check{J}_{j+1}^n,$$

with

$$\begin{aligned} \hat{J}_{j+1}^n &= \inf \left\{ t \geq \tau_j^n; X_t - X_{\tau_j^n} \notin (-\varepsilon_n(\underline{l}_{\tau_j^n} + \delta_n), \varepsilon_n(\bar{l}_{\tau_j^n} + \delta_n)) \right\} \wedge T_0, \\ \check{J}_{j+1}^n &= \inf \left\{ t \geq \tau_j^n; X_t - X_{\tau_j^n} \notin (-\varepsilon_n(\underline{l}_{\tau_j^n} - \delta_n), \varepsilon_n(\bar{l}_{\tau_j^n} - \delta_n)) \right\} \wedge T_0. \end{aligned}$$

By the Dambis, Dubins-Schwarz theorem, there exists a Brownian motion W^* such that

$$X_{\tau_j^n+t} - X_{\tau_j^n} = W_{\int_{\tau_j^n}^{\tau_j^n+t} (\sigma_s^X)^2 ds}^*.$$

Therefore, since σ^X is bounded from below by $\frac{1}{K}$

$$\begin{aligned} \hat{J}_{j+1}^n - \check{J}_{j+1}^n &\leq (\hat{J}_{j+1}^n - \check{J}_{j+1}^n) \mathbf{1}_{\check{J}_{j+1}^n < T_0} \\ &\leq \mathbf{1}_{\check{J}_{j+1}^n < T_0, X_{\tau_j^n} - X_{\tau_j^n} = -\varepsilon_n(\bar{l}_{\tau_j^n} - \underline{l}_{\tau_j^n})} \inf\{t \geq 0 : W_{t/K^2}^* \notin (-2\varepsilon_n\delta_n, \varepsilon_n(\bar{l}_{\tau_j^n} + \underline{l}_{\tau_j^n}))\} \\ &\quad + \mathbf{1}_{\check{J}_{j+1}^n < T_0, X_{\tau_j^n} - X_{\tau_j^n} = \varepsilon_n(\bar{l}_{\tau_j^n} - \underline{l}_{\tau_j^n})} \inf\{t \geq 0 : W_{t/K^2}^* \notin (-\varepsilon_n(\bar{l}_{\tau_j^n} + \underline{l}_{\tau_j^n}), 2\varepsilon_n\delta_n)\}. \end{aligned}$$

An explicit computation with the Brownian exit time yields:

$$\mathbb{E}\left[\mathbb{E}[\hat{J}_{j+1}^n - \check{J}_{j+1}^n | \mathcal{F}_{\tau_j^n}]\right] \leq C\varepsilon_n^2\delta_n.$$

Consequently,

$$\mathbb{E}[R_2] \leq C\delta_n,$$

which gives the result. \square

Part 3: Proof of (1.2.10)

Here we prove (1.2.10), which completes the proof of the convergence in law of $\varepsilon_n^{-1}Z_T^n$ with the help of Proposition 1.2.1 and Proposition 1.2.2. As already seen, by Itô's formula, we have

$$\begin{aligned} \mathbb{E}[\Delta_{j,n}^4 | \mathcal{F}_{\tau_j^n}] &= 6\mathbb{E}\left[\int_{\tau_j^n}^{\tau_j^n+1} (X_t - X_{\tau_j^n})^2 d\langle X \rangle_t | \mathcal{F}_{\tau_j^n}\right] = :A_j, \\ \mathbb{E}[\Delta_{j,n}^3 | \mathcal{F}_{\tau_j^n}] &= 3\mathbb{E}\left[\int_{\tau_j^n}^{\tau_j^n+1} (X_t - X_{\tau_j^n}) d\langle X \rangle_t | \mathcal{F}_{\tau_j^n}\right] = :B_j. \end{aligned}$$

Therefore, we obtain

$$\begin{aligned} \varepsilon_n^{-2} \sum_{j=0}^{N_t^n} \kappa_{\tau_j^n} A_j &= \varepsilon_n^{-2} \sum_{j=0}^{N_t^n} \kappa_{\tau_j^n} \mathbb{E}[(X_{\tilde{\tau}_{j+1}^n} - X_{\tau_j^n})^4 | \mathcal{F}_{\tau_j^n}] + R_t, \\ \varepsilon_n^{-1} \sum_{j=0}^{N_t^n} \kappa_{\tau_j^n} B_j &= \varepsilon_n^{-1} \sum_{j=0}^{N_t^n} \kappa_{\tau_j^n} \mathbb{E}[(X_{\tilde{\tau}_{j+1}^n} - X_{\tau_j^n})^3 | \mathcal{F}_{\tau_j^n}] + R'_t, \end{aligned}$$

where

$$\begin{aligned} R_t &= 6\varepsilon_n^{-2} \sum_{j=0}^{N_t^n} \kappa_{\tau_j^n} \mathbb{E}\left[\int_{\tilde{\tau}_{j+1}^n}^{\tau_j^n+1} (X_u - X_u^n)^2 (\sigma_u^X)^2 du | \mathcal{F}_{\tau_j^n}\right], \\ R'_t &= 3\varepsilon_n^{-1} \sum_{j=0}^{N_t^n} \kappa_{\tau_j^n} \mathbb{E}\left[\int_{\tilde{\tau}_{j+1}^n}^{\tau_j^n+1} (X_u - X_u^n) (\sigma_u^X)^2 du | \mathcal{F}_{\tau_j^n}\right]. \end{aligned}$$

Note that the integrals above may be signed integrals. Since $\varepsilon_n^{-1} \sup_t |X_t - X_t^n| \leq M$ and $\sigma^X \leq M$, R and R' converge to 0 uniformly in probability on $[0, T_0]$ by Lemma 1.A.7.

Let $\bar{p} = \mathbb{P}[X_{\tilde{\tau}_{j+1}^n} - X_{\tau_j^n} = \varepsilon_n \bar{l} | \mathcal{F}_{\tau_j^n}]$, $\underline{p} = \mathbb{P}[X_{\tilde{\tau}_{j+1}^n} - X_{\tau_j^n} = -\varepsilon_n \underline{l} | \mathcal{F}_{\tau_j^n}]$ and $p_0 = \mathbb{P}[\tilde{\tau}_{j+1}^n = T_0 | \mathcal{F}_{\tau_j^n}]$. By the optional sampling theorem,

$$\frac{\underline{l}}{\underline{l} + \bar{l}} - p_0 \leq \bar{p} \leq \frac{\underline{l}}{\underline{l} + \bar{l}} + p_0 \quad \text{and} \quad \frac{\bar{l}}{\underline{l} + \bar{l}} - p_0 \leq \underline{p} \leq \frac{\bar{l}}{\underline{l} + \bar{l}} + p_0.$$

For $q \in \mathbb{N}$, we then get,

$$\frac{\bar{l}^q \underline{l} + (-1)^q \bar{l}^q \underline{l}}{\underline{l} + \bar{l}} - Cp_0 \leq \varepsilon_n^{-q} \mathbb{E}[(X_{\tilde{\tau}_{j+1}^n} - X_{\tau_j^n})^q | \mathcal{F}_{\tau_j^n}] \leq \frac{\bar{l}^q \underline{l} + (-1)^q \bar{l}^q \underline{l}}{\underline{l} + \bar{l}} + Cp_0,$$

and in particular

$$\bar{l} - Cp_0 \leq \varepsilon_n^{-q} \mathbb{E}[(X_{\tilde{\tau}_{j+1}^n} - X_{\tau_j^n})^2 | \mathcal{F}_{\tau_j^n}] \leq \bar{l} + Cp_0. \quad (1.A.20)$$

Therefore

$$\left| \varepsilon_n^{-2} \sum_{j=0}^{N_t^n} \kappa_{\tau_j^n} \mathbb{E}[(X_{\tilde{\tau}_{j+1}^n} - X_{\tau_j^n})^4 | \mathcal{F}_{\tau_j^n}] - \varepsilon_n^2 \sum_{j=0}^{N_t^n} \kappa_{\tau_j^n} \bar{l}_{\tau_j^n} \underline{l}_{\tau_j^n} a_{\tau_j^n}^2 \right| \leq C \varepsilon_n^2 \sum_{j=0}^{N_t^n} \kappa_{\tau_j^n} \mathbb{E}[\mathbf{1}_{\tilde{\tau}_{j+1}^n = T_0} | \mathcal{F}_{\tau_j^n}]$$

$$\left| \varepsilon_n^{-1} \sum_{j=0}^{N_t^n} \kappa_{\tau_j^n} \mathbb{E}[(X_{\tilde{\tau}_{j+1}^n} - X_{\tau_j^n})^3 | \mathcal{F}_{\tau_j^n}] + \varepsilon_n^2 \sum_{j=0}^{N_t^n} \kappa_{\tau_j^n} \bar{l}_{\tau_j^n} \underline{l}_{\tau_j^n} s_{\tau_j^n} \right| \leq C \varepsilon_n^2 \sum_{j=0}^{N_t^n} \kappa_{\tau_j^n} \mathbb{E}[\mathbf{1}_{\tilde{\tau}_{j+1}^n = T_0} | \mathcal{F}_{\tau_j^n}],$$

where

$$a^2 = \bar{l}^2 + \underline{l}^2 - \bar{l}\underline{l}, \quad s = \underline{l} - \bar{l}.$$

The terms in the right-hand side are positive and converge to zero in expectation, thus also in probability. Using (1.A.20), we see that to complete the proof it suffices to show that the convergences

$$\sum_{j=0}^{N_t^n} \kappa_{\tau_j^n} a_{\tau_j^n}^2 \mathbb{E}[(X_{\tilde{\tau}_{j+1}^n} - X_{\tau_j^n})^2 | \mathcal{F}_{\tau_j^n}] \rightarrow \int_0^{\cdot} a_u^2 (\sigma_u^Y)^2 du,$$

$$\sum_{j=0}^{N_t^n} \kappa_{\tau_j^n} s_{\tau_j^n} \mathbb{E}[(X_{\tilde{\tau}_{j+1}^n} - X_{\tau_j^n})^2 | \mathcal{F}_{\tau_j^n}] \rightarrow \int_0^{\cdot} s_u (\sigma_u^Y)^2 du$$

hold uniformly in probability on $[0, T_0]$. These follow from

$$\mathbb{E}[(X_{\tilde{\tau}_{j+1}^n} - X_{\tau_j^n})^2 | \mathcal{F}_{\tau_j^n}] = \mathbb{E}\left[\int_{\tau_j^n}^{\tilde{\tau}_{j+1}^n} (\sigma_u^X)^2 du \middle| \mathcal{F}_{\tau_j^n}\right] = \mathbb{E}\left[\int_{\tau_j^n}^{\tilde{\tau}_{j+1}^n} \frac{1}{\kappa_u} (\sigma_u^Y)^2 du \middle| \mathcal{F}_{\tau_j^n}\right]$$

and Lemmas 1.A.4 and 1.A.7 with the aid of Lenglart inequality for discrete martingales.

Proof of (1.2.14)

Before completing the proof for the admissibility, here we prove (1.2.14) that shows the efficiency of the hitting times strategy. The same localization procedure applies here as in 1.A.3. From (1.A.20) and the following estimates, we have

$$\varepsilon_n^2 N_t^n = \sum_{j=0}^{N_t^n - 1} \frac{1}{\bar{l}_{\tau_j^n} \underline{l}_{\tau_j^n}} \mathbb{E}[(X_{\tilde{\tau}_{j+1}^n} - X_{\tau_j^n})^2 | \mathcal{F}_{\tau_j^n}] + o_p(1) = \sum_{j=0}^{N_t^n - 1} \frac{1}{\bar{l}_{\tau_j^n} \underline{l}_{\tau_j^n}} \mathbb{E}\left[\int_{\tau_j^n}^{\tilde{\tau}_{j+1}^n} (\sigma_u^X)^2 du \middle| \mathcal{F}_{\tau_j^n}\right] + o_p(1),$$

which implies (1.2.14) by Lenglart inequality for discrete martingales.

Proof of (1.2.4)

Here, we prove a moment convergence result. Thus, the localization procedure does not apply here. We set

$$A_n = \varepsilon_n^{-1} \int_0^T (X_t^n - X_t) b_t^Y dt,$$

$$B_n = \varepsilon_n^{-1} \int_0^T (X_t^n - X_t) \sigma_t^Y dW_t^Y.$$

We have

$$(\varepsilon_n^{-1} Z_T^n)^2 = (A_n + B_n)^2 \leq 2(A_n^2 + B_n^2).$$

Thus it is enough to prove the uniform integrability of (A_n^2) and (B_n^2) to obtain the result. For (A_n^2) , we have

$$\sup_n (A_n)^2 \leq \left(\int_0^T (\bar{l}_t \vee l_t) |b_t^Y| dt \right)^2 \leq T \int_0^T (\bar{l}_t \vee l_t)^2 (\rho_t)^2 (\sigma_t^Y)^2 dt.$$

The right hand side of the last inequality being integrable, this gives the result for $(A_n)^2$. We now turn to (B_n^2) . The sequence (B_n^2) is non-negative, integrable, and converges in law towards an integrable limit. Thus, the uniform integrability is equivalent to the convergence in expectation, see for example [Bil09]. Since

$$\varepsilon_n^{-2} \langle Z^n \rangle_T \rightarrow \frac{1}{6} \int_0^T a_t^2 (\sigma_t^Y)^2 dt$$

and

$$\varepsilon_n^{-2} \langle Z^n \rangle_T \leq \int_0^T (\bar{l}_t \vee l_t)^2 (\sigma_t^Y)^2 dt,$$

we readily obtain

$$\mathbb{E}[B_n^2] = \mathbb{E}[\varepsilon_n^{-2} \langle Z^n \rangle_T] \rightarrow \frac{1}{6} \mathbb{E} \left[\int_0^T a_t^2 (\sigma_t^Y)^2 dt \right],$$

which concludes the proof.

1.A.4 Proof of Theorem 1.4.1

We start with the first part of Theorem 1.4.1. Let (m^*, v^*) be an efficient couple. Suppose that there exists a process s'_t such that the associated expectation, say $m' = \mathbb{E}[Z_{T,s'}]$, is larger than m^* and the expected error, say $v' = \mathbb{E}[(Z_{T,s'})^2]$, is strictly smaller than v^* . From Lemma 1.2.1, for any η we can find an admissible strategy with limiting variable $Z_{s'+\eta, s'}$. Clearly, we can find η small enough, such that $\mathbb{E}[Z_{s'+\eta, s'}] = m'$ and

$$v' \leq \mathbb{E}[(Z_{s'+\eta, s'}^*)^2] < v^*,$$

which is a contradiction. On the other hand, taking Z_{a^*, s^*} to be the limiting random variable associated to (m^*, v^*) , the process Z_{T, s^*} satisfies the required conditions.

For the second part, assume that s^* is an adapted continuous process satisfying the conditions of the theorem. Then, since $a_t^2 \geq s_t^2$, there is no admissible discretization rule whose associated limiting random variable $Z_{a, s}^*$ satisfies

$$\mathbb{E}[Z_{a, s}^*] \geq m^* \quad \text{and} \quad \mathbb{E}[(Z_{a, s}^*)^2] < v^*.$$

It remains to show that the proposed discretization rules indeed lead to nearly efficient couples. The fact that they are admissible is clear from Proposition 1.2.3. Recall now that for the suggested rule

$$a_t^2 = (s_t^*)^2 + \frac{6\delta}{(\sigma_t^Y)^2}.$$

This equality gives that the limiting variable $Z_{a, s}^*$ associated to this discretization rule satisfies

$$\mathbb{E}[Z_{a, s}^*] = \frac{1}{3} \mathbb{E} \left[\int_0^T s_t^* dY_t \right]$$

and

$$\begin{aligned}\mathbb{E}[(Z_{a,s}^*)^2] &= \frac{1}{9}\mathbb{E}\left[\left(\int_0^T s_t^* dY_t\right)^2\right] + \frac{1}{6}\mathbb{E}\left[\int_0^T \left(a_t^2 - \frac{2}{3}(s_t^*)^2\right)(\sigma_t^Y)^2 dt\right] \\ &= \frac{1}{9}\mathbb{E}\left[\left(\int_0^T s_t^* dY_t\right)^2\right] + \frac{1}{18}\mathbb{E}\left[\int_0^T \left((s_t^*)^2(\sigma_t^Y)^2\right) dt\right] + \delta T.\end{aligned}$$

The couple

$$\left(\frac{1}{3}\mathbb{E}\left[\int_0^T s_t^* dY_t\right], \frac{1}{9}\mathbb{E}\left[\left(\int_0^T s_t^* dY_t\right)^2\right] + \frac{1}{18}\mathbb{E}\left[\int_0^T \left(\frac{1}{2}(s_t^*)^2(\sigma_t^Y)^2\right) dt\right]\right)$$

being an efficient couple with respect to \mathcal{Z}_T , we obtain the result.

Appendix 1.B Linear-quadratic optimal control

We give here a summary of useful formulas from [ZL00]. Consider a controlled system governed by the following linear SDE:

$$\begin{cases} dX_t = (A_t X_t + B_t u_t + f_t) dt + \sum_{j=1}^m D_t^j u_t dW_t^j, \\ X_0 = x \in \mathbb{R}^n, \end{cases} \quad (1.B.1)$$

where x is the initial state and $W = (W^1, \dots, W^m)$ is a m -dimensional Brownian motion on a given filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t \geq 0})$ and $u \in L_{\mathcal{F}}^2([0, T], \mathbb{R}^m)$ is a control. For each control u , the associated cost is

$$J(u) = \mathbb{E} \left[\int_0^T \frac{1}{2} (X_t' Q_t X_t + u_t' R_t u_t) dt + \frac{1}{2} X_T' H X_T \right]. \quad (1.B.2)$$

We suppose that all the parameters are deterministic and continuous on $[0, T]$ and H belongs to S_+^n the set of $n \times n$ symmetric positive matrices. We introduce the following matrix Riccati equation

$$\begin{cases} \dot{P}_t = -P_t A_t - A_t' P_t - Q_t + P_t B_t K_t^{-1} B_t' P_t, \\ P_T = H, \\ K_t = R_t + \sum_{j=1}^m D_t^j P_t D_t^j > 0, \quad \forall t \in [0, T], \end{cases} \quad (1.B.3)$$

along with an equation

$$\begin{cases} \dot{g}_t = -A_t' g_t + P_t B_t K_t^{-1} B_t' g_t - P_t f_t, \\ g_T = 0. \end{cases} \quad (1.B.4)$$

Then following result is given in [ZL00].

Theorem 1.B.1. *If (1.B.3) and (5.4.7) admit solutions $P \in C([0, T], S_+^n)$ and $g \in C([0, T], \mathbb{R}^n)$, respectively, then the stochastic linear-quadratic control problem (1.B.1)-(1.B.2) has an optimal feedback control*

$$u^*(t, x) = -K_t^{-1} B_t' (P_t X_t + g_t).$$

Moreover, the optimal cost value is

$$J^* = \frac{1}{2} \int_0^T \left(2f_t' g_t - g_t B_t K_t^{-1} B_t' g_t \right) dt + \frac{1}{2} x' P_0 x + x g_0.$$

Chapter 2

Asymptotic replication with modified volatility under small proportional costs¹

Contents

2.1	Introduction	55
2.2	Framework	60
2.2.1	Benchmark strategy	61
2.2.2	Continuous control strategies	62
2.2.3	Hedging error	63
2.3	Main results	63
2.3.1	Limit theorem of hedging error	63
2.3.2	Minimum conditional variance	66
2.3.3	Asymptotically optimal sequence	67
2.3.4	Implementation and numerical experiments	69
2.4	Heuristic derivation	71
2.4.1	Local probability model	73
2.4.2	Relation with Leland's strategy	75
2.4.3	Relation with Whalley and Wilmott's strategy	76
Appendix 2.A	Proof of Theorem 2.3.1	78
Appendix 2.B	Proof of Theorem 2.3.2	81

2.1 Introduction

We consider in this chapter dynamic hedging of an European option under a general local volatility model with proportional transaction costs. The standard delta hedging strategy is no longer feasible since a continuous rebalance according to the strategy incurs infinite transaction cost. Moreover, as shown by [SSC95, LS97], the super replication of the option payoff is as expensive as a trivial static hedging strategy. Therefore, the classical framework of perfect (super) replication does not give a reasonable price of the option. One should give up hedging in the almost-sure sense and try to reduce hedging error in a distributional sense by a dynamic strategy with a reasonable amount of initial capital. The aim of this study is to give such a strategy which is as explicit as possible and optimal in a suitable sense.

1. This is a joint work with Prof. Masaaki Fukasawa.

There have been two approaches for this problem in the literature. The first approach is based on the utility indifference principle: [HN89, DPZ93]. In general, optimal strategies are not explicit. Denote by κ the proportionality coefficient of transaction cost. Considering κ to be small as it is in liquid markets, [WW97] performed a formal asymptotic analysis to give an explicit strategy which is asymptotically optimal. The results have been rigorously proved by [Bic12, KL13]. The asymptotically optimal hedging strategy is a singular control; the number of shares of the underlying assets is singularly controlled so that it keeps staying around a benchmark level which is optimal under no transaction cost. [BS98] studied another scaling limit of the utility indifference price, where not only κ tends to zero but also the parameter of risk aversion tends to infinity. Their optimal strategy is again a singular control, while their benchmark strategy is not the same as in the Whalley-Wilmott strategy and given in terms of a nonlinear Black-Scholes equation.

The second approach is to replicate the option payoff asymptotically. [Lel85] claimed that a carefully modified and discretized delta hedging strategy asymptotically replicates the option payoff as $\kappa \rightarrow 0$. A rigorous treatment was given by Lott; see [KS09]. An associated central limit theorem was shown by [DK10]. As we explain later in more detail, Leland's idea has two ingredients. The first is to find a good benchmark strategy which yields certain surplus in the absence of transaction cost. The second is to construct a good approximation to the benchmark by a strategy of finite transaction cost so that the incurred costs are compensated by the surplus. Thanks to this compensation, Leland's strategy enjoys a better rate of convergence than the Whalley-Wilmott strategy does. The rate is the same as for the Barles-Soner strategy as we show later. For the approximation part, Leland used an equidistant discretization of the benchmark strategy. Other discretization rules have been examined in the literature: [Tof96, GS96, ADG98, DK10, Fuk11b, Lan13]. They result in different asymptotic variances of hedging error. None of them uniformly outperforms others. To find an optimal method has been an open problem.

To discuss the optimality among discretization rules is however not relevant because other than discretized processes (simple predictable processes), any predictable process can work as an approximation of finite cost as long as it is staying around the benchmark and of finite variation. Moreover, the utility-based approach explained above suggests the superiority of singular control strategies that form a particular class of such processes. In this chapter, we consider a reasonable class of continuous control strategies including singular ones, and give a condition under which the option payoff is asymptotically replicated. More precisely, we consider a class of trading strategies of finite transaction cost and study the deviations of the associated wealth processes from a given benchmark portfolio value process. We prove a central limit theorem for the tracking error processes as $\kappa \rightarrow 0$. The limit distribution serves as a reasonable approximation to the distribution of the tracking (or hedging) error when κ is small as it is in liquid markets. Further, we give a sequence of explicit strategies which is asymptotically optimal in the sense that the asymptotic variance of hedging error is minimized. The minimization is done by fixing a benchmark; to optimize the benchmark strategy becomes a relevant problem only after the tracking error is minimized and so, is beyond the scope of this already lengthy chapter. We find in particular that our optimal strategy enjoys a better rate of convergence than the Whalley-Wilmott strategy does, and that the Barles-Soner strategy is not always optimal according to our criterion.

Our optimal strategy uniformly outperforms Leland's strategy. To explain it, here we focus on the case that the payoff function is convex and the benchmark strategy is given by a constant enlargement of volatility, although our framework is much more general as described in Section 2.2.

Let us start with a revisit to Leland's idea. The first trick is to consider the delta hedging strategy with enlarged volatility. Let p^α be a solution of the partial differential equation (PDE)

$$\partial_t p^\alpha(s, t) + \frac{1}{2} \left(1 + \frac{2}{\alpha}\right) \sigma(s, t)^2 \partial_s^2 p^\alpha(s, t) = 0, \quad p^\alpha(s, T) = f(s), \quad (2.1.1)$$

where f is a convex payoff function, T is the maturity of the option, σ is the local volatility function of the underlying asset S , and α is an arbitrary positive constant that controls the enlargement of volatility. By the convexity of f , we have $\partial_s^2 p^\alpha \geq 0$ under a reasonable condition on σ ; see [EJPS98]. By Itô's formula,

$$f(S_T) = \Pi_0^\alpha + \int_0^T X_u^\alpha dS_u - \frac{1}{\alpha} \int_0^T \Gamma_u^\alpha d\langle S \rangle_u,$$

where

$$\Pi_t^\alpha = p^\alpha(S_t, t), \quad X_t^\alpha = \partial_s p^\alpha(S_t, t), \quad \Gamma_t^\alpha = \partial_s^2 p^\alpha(S_t, t). \quad (2.1.2)$$

This means that, without transaction costs and assuming zero interest rates, the self-financing strategy X^α with initial capital Π_0^α super-hedges the payoff $f(S_T)$ with surplus

$$\frac{1}{\alpha} \int_0^T \Gamma_t^\alpha d\langle S \rangle_t \geq 0. \quad (2.1.3)$$

The second trick of Leland is to approximate X^α by a strategy of finite transaction cost, and exploit the surplus (2.1.3) to compensate the incurred costs. Under the Black-Scholes model ($\sigma(s, t) = vs$, $v > 0$), Leland considered an equidistant discretization of X^α ; define $X^{\alpha, \kappa}$ by

$$X_t^{\alpha, \kappa} = X_{jh}^\alpha, \quad t \in (jh, (j+1)h], \quad j = 0, 1, 2, \dots, \quad (2.1.4)$$

with $h > 0$ the interval of rebalancing. Set the initial capital $\Pi_{0-}^{\alpha, \kappa}$, that is, the price of the option to be

$$\Pi_{0-}^{\alpha, \kappa} = \Pi_0^\alpha + \kappa S_0 |\Delta X_0^\alpha|.$$

The second term is to compensate the transaction cost at the inception. The associated wealth process $\Pi^{\alpha, \kappa}$ is then

$$\Pi_t^{\alpha, \kappa} = \Pi_0^\alpha + \int_0^t X_u^{\alpha, \kappa} dS_u - \kappa \sum_{0 < u \leq t} S_u |\Delta X_u^{\alpha, \kappa}|. \quad (2.1.5)$$

The magic is that, by choosing

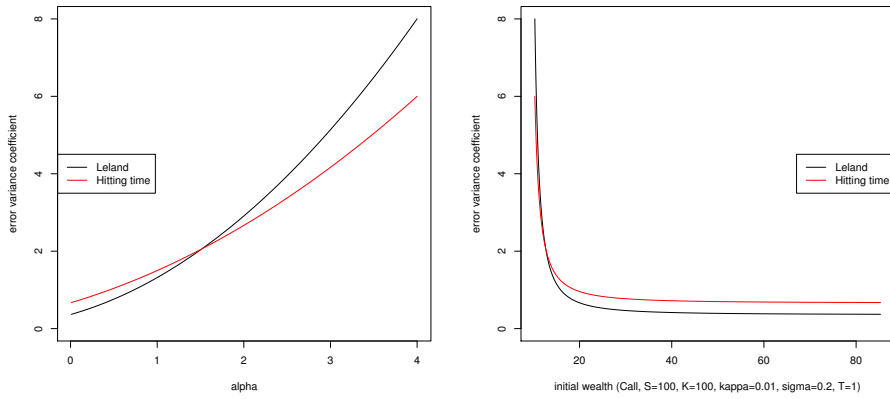
$$h = \frac{2}{\pi} \frac{\kappa^2 \alpha^2}{v^2}, \quad (2.1.6)$$

we have

$$\begin{aligned} \int_0^T X_u^{\alpha, \kappa} dS_u &\rightarrow \int_0^T X_u^\alpha dS_u, \\ \kappa \sum_{0 < u \leq T} S_u |\Delta X_u^{\alpha, \kappa}| &\rightarrow \frac{1}{\alpha} \int_0^T \Gamma_u^\alpha d\langle S \rangle_u \end{aligned} \quad (2.1.7)$$

as $\kappa \rightarrow 0$ with rate κ .

Consequently, the terminal wealth $\Pi_T^{\alpha, \kappa}$ is close to $f(S_T)$ when κ is small as it is in liquid markets. In this sense, the self-financing strategy $X^{\alpha, \kappa}$ is an asymptotic replication strategy. The


 Figure 2.1 – Comparison between $\eta_L(\alpha)$ and $\eta_F(\alpha)$.

way how to discretize X^α is essential. The first convergence of (2.1.7) holds in general as transactions are more and more frequent. On the other hand, if they are too frequent, then the total amount of transaction costs exceeds the surplus (2.1.3) and the second convergence of (2.1.7) fails to hold. Therefore the frequency (2.1.6) results from a delicate balance. Given κ , the value of h can be very small if α is very small, which is the case that the pricing volatility is much enlarged to make the option price close to the super replication price.

Naturally we expect that a strategy with smaller α (requiring more initial capital) results in a smaller error. In fact, we have (cf. [DK10, KS09])

$$\kappa^{-1}(\Pi^\alpha - \Pi^{\alpha,\kappa}) \rightarrow W_Q, \quad Q = \eta_L(\alpha) \int_0^\cdot |\Gamma_u^\alpha S_u|^2 d\langle S \rangle_u \quad (2.1.8)$$

stably in law on $D[0, T]$ as $\kappa \rightarrow 0$, with an independent standard Brownian motion W and

$$\eta_L(\alpha) = \frac{1}{\pi}\alpha^2 + \frac{2}{\pi}\alpha + 1 - \frac{2}{\pi}.$$

The function η_L is increasing as shown in the left panel of Fig. 2.1 (black curve). In particular, the distribution of normalized hedging error $(f(S_T) - \Pi_T^{\alpha,\kappa})/\kappa$ is asymptotically mixed normal with mean zero and variance

$$Q_T = \eta_L(\alpha) \int_0^T |\Gamma_u^\alpha S_u|^2 d\langle S \rangle_u.$$

The conditional variance Q represents in a certain sense the quality of replication. The smaller the conditional variance is, the better hedging is achieved. Therefore, this asymptotic result serves as an important element when considering the trade-off between cost (initial capital) and risk (hedging error) by controlling α ; see the right of Fig. 2.1, where $(\Pi_0^{\alpha,\kappa}, \eta_L(\alpha))$ is plotted for $\alpha \in (0, 4)$ and $f(s) = (s - 100)_+$, $T = 1$, $S_0 = 100$, $\kappa = 0.01$ under the Black-Scholes model with volatility 0.2.

Fig. 2.1 also represents another function

$$\eta_F(\alpha) = \frac{(\alpha + 2)^2}{6}$$

together (in red), which appears instead of η_L when we use a discretization rule with respect to a specific sequence of random times

$$\tau_0^\kappa = 0, \quad \tau_{j+1}^\kappa = \inf\{t > \tau_j^\kappa; |X_t^\alpha - X_{\tau_j^\kappa}^\alpha| \geq \kappa b_{\tau_j^\kappa}\}, \quad b_t = \alpha S_t \Gamma_t^\alpha. \quad (2.1.9)$$

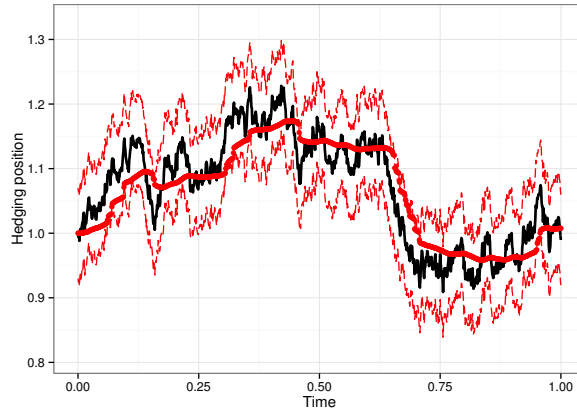


Figure 2.2 – Candidate strategies. Black: X_t^α , red: $X_t^{b,c,\kappa}$, blue: barriers $X_t^\alpha \pm b(S_t, t)$.

As Fig. 2.1 shows, neither η_L nor η_F is uniformly smaller; see [Fuk11b] for more details.

While the approach of Leland is appreciated since it is easy to implement, the solution is still far from being satisfactory. The strategies (2.1.4) and (2.1.9) are by no means the only choices as an approximation to X^α . For example, it is not necessary to match the benchmark strategy X^α after each rebalancing. In particular, it is possible to trade frequently but with a small trading volume each time. Indeed, as already mentioned, several results on related problems under the framework of utility maximization suggest that, under proportional transaction costs, the optimal strategy is to trade a minimal amount in continuous-time to keep the deviation from the benchmark inside a non-transaction zone; see [WW97, Bic14, BS98, ST13, KL13]. The contribution of the current chapter in this context is therefore two-fold. First, we consider a reasonable class of continuous trading strategies of finite transaction cost, and provide a limit theorem for the corresponding replication error. In particular, we identify the condition for those strategies to (asymptotically) replicate or super-replicate the option in question. Second, we minimize the asymptotic variance of replication error among those asymptotic replication strategies.

Under our framework, a candidate strategy $X^{b,c,\kappa}$, see Fig. 2.2, is indexed by two nonnegative functions $b(s, t)$ and $c(z, s, t)$. Let $Z^\kappa = (X^\alpha - X^{b,c,\kappa})/\kappa$ be the normalized deviation of $X^{b,c,\kappa}$ from the benchmark position X^α . We treat $X^{b,c,\kappa}$ of the form

$$dX_t^{b,c,\kappa} = \frac{1}{\kappa} \text{sgn}(Z^\kappa) c(|Z_t^\kappa|, S_t, t) d\langle X^\alpha \rangle_t - \kappa dL_t^\kappa + \kappa dR_t^\kappa, \quad X_{0+}^{b,c,\kappa} = X_0^\alpha,$$

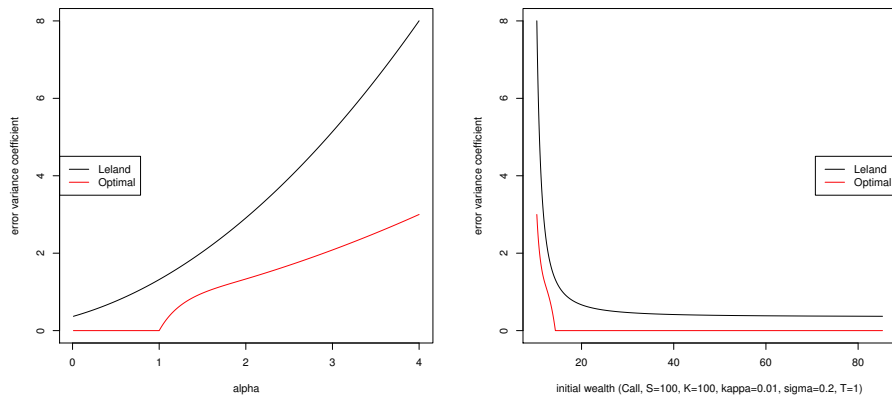
where L^κ and R^κ are non-decreasing processes such that

$$L_t^\kappa = \int_0^t 1_{\{Z_u^\kappa = -b(S_u, u)\}} dL_u^\kappa, \quad R_t^\kappa = \int_0^t 1_{\{Z_u^\kappa = b(S_u, u)\}} dR_u^\kappa, \quad |Z_t^\kappa| \leq b(S_t, t).$$

Intuitively, the regular control part $\kappa^{-1} \text{sgn}(Z^\kappa) c(|Z_t^\kappa|, S_t, t) d\langle X^\alpha \rangle_t$ pushes $X^{b,c,\kappa}$ toward X^α and is active when X^α moves. The singular control part $-\kappa dL_t^\kappa + \kappa dR_t^\kappa$ keeps Z^κ within the stochastic interval $[-b(S_t, t), b(S_t, t)]$, and is active only when Z^κ touches the boundary.

Denoting by $\Pi^{b,c,\kappa}$ the wealth process associated with $X^{b,c,\kappa}$, we show in Theorem 2.3.1 that

$$\kappa^{-1} \left(\Pi^\alpha - \Pi^{b,c,\kappa} - \int_0^\cdot \delta^{b,c}(S_t, t) dt \right) \rightarrow W_{Q^{b,c}},$$

Figure 2.3 – Comparison between $\eta_L(\alpha)$ and $\eta_{\dagger}(\alpha)$.

stably in law on $C[0, T]$ as $\kappa \rightarrow 0$, where W is a standard Brownian motion independent of S , and $Q^{b,c}$ and $\delta^{b,c}$ are explicitly determined by b and c . Fixing the local limit of hedging error $\delta^{b,c} = \delta$, we provide in Theorem 2.3.2 an explicit expression for the infimum

$$Q_*^\delta := \operatorname{ess\,inf}_{(b,c) \text{ s.t. } \delta^{b,c} \equiv \delta} Q^{b,c},$$

among the candidate strategies, together with a sequence of explicit strategies attaining asymptotically the infimum Q_*^δ . In particular, if b and c are chosen in such a way that $\delta^{b,c} \equiv 0$ (i.e. asymptotic replication), then we have

$$Q_*^0 = \eta_{\dagger}(\alpha) \int_0^\cdot |\Gamma_u^\alpha S_u|^2 d\langle S \rangle_u.$$

The coefficient $\eta_{\dagger}(\alpha)$ is represented in Fig. 2.3(red), where we observe a significant improvement in terms of the asymptotic variance of hedging error compared to Leland's strategy.

The rest of the chapter is structured as follows. In Section 2.2, we define a class of continuous trading strategies which is admissible in the presence of proportional transaction costs. The main results are stated in Section 5.4. We also provide a detailed description for the implementation in practice and several numerical experiments. In Section 2.4, a heuristic derivation of main results and a detailed comment on the relation with existing literature can be found. We give the rigorous proofs in Section 2.A and 2.B.

2.2 Framework

In this section, we give a rigorous formulation of the problem and describe our class of continuous trading strategies. Let $(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t; t \geq 0\})$ be a filtered probability space satisfying the usual assumptions. Let $T > 0$ be a constant which stands for the maturity of an European option. Let f be a Borel function on $(0, \infty)$ which stands for the payoff function of the option. We suppose the underlying asset price process S of the option to be positive and continuous on $[0, T]$ and to follow

$$dS_t = \Theta_t dt + \sigma(S_t, t) dB_t$$

on $[0, T]$, where Θ is an $\{\mathcal{F}_t\}$ -adapted locally bounded process, B is an $\{\mathcal{F}_t\}$ -standard Brownian motion and σ is a positive $C^{3,1}$ functions on $(0, \infty) \times [0, T]$. We assume interest rates to be zero for brevity.

A trading strategy is given by an $\{\mathcal{F}_t\}$ -adapted caglad process X and the associated wealth process Π is defined by

$$\Pi_t = \Pi_{0-} + \int_0^t X_u dS_u - \kappa \int_{[0,t]} \lambda(S_u, u) d\|X\|_u, \quad (2.2.1)$$

where $\|\cdot\|$ is the total variation representing the trading volume, and λ is a nonnegative $C^{3,1}$ function on $(0, \infty) \times [0, T]$ representing the structure of transaction cost. Note that (2.1.5) is the particular case where $\lambda(s, t) = s$. The constant $\kappa > 0$ appeared in (2.1.5) and (2.2.1) represents the proportionality coefficient of transaction cost, which is supposed to be small. We will study the asymptotic behavior of hedging as $\kappa \rightarrow 0$, which serves as a valid approximation to the hedging behavior when κ is sufficiently small.

2.2.1 Benchmark strategy

Denote by \mathcal{A} the set of $C^{3,1}$ functions φ on $(0, \infty) \times [0, T]$ such that for each $i \in \{0, 1, 2, 3\}$ and $j \in \{0, 1\}$, $\partial_s^i \partial_t^j \varphi(S_t, t)$ converges almost surely as $t \rightarrow T$.

Definition 2.2.1. Let q be a continuous function on $(0, \infty) \times [0, T]$ such that

1. for each $i \in \{0, 1, 2\}$ and $j \in \{0, 1\}$, $\partial_s^i \partial_t^j q$ exists and belongs to \mathcal{A} ,
2. $\lim_{t \rightarrow T} q(S_t, t) = f(S_T)$ almost surely, and
3. $\partial_s^2 q$ is nondegenerate in the sense that

$$\int_0^T 1_{\{\partial_s^2 q(S_t, t) = 0\}} dt = 0 \quad (2.2.2)$$

almost surely.

Define Π^q, X^q and Γ^q by

$$\Pi_t^q = q(S_t, t), \quad X_t^q = \partial_s q(S_t, t), \quad \Gamma_t^q = \partial_s^2 q(S_t, t).$$

Then we call Π^q the benchmark portfolio value and X^q the benchmark strategy.

By Itô's formula, the benchmark portfolio value verifies

$$\Pi_t^q = \Pi_0^q + \int_0^t X_u^q dS_u + \int_0^t \mathcal{L}q(S_u, u) du, \quad (2.2.3)$$

where

$$\mathcal{L}q(s, t) = \partial_t q(s, t) + \frac{1}{2} \sigma(s, t)^2 \partial_s^2 q(s, t).$$

Example 2.2.1. When f is a piecewise C^2 convex non-affine function and S_T admits a density, an example of q is given by the Black-Scholes pricing function

$$q(s, t) = \int f(s \exp\{-w^2(T-t)/2 + w\sqrt{T-t}z\}) \phi(z) dz, \quad (2.2.4)$$

where $w > 0$ is an arbitrary constant and ϕ is the standard normal density. In fact, $\partial_s^2 q > 0$ since f is a piecewise C^2 convex non-affine function (consider its decomposition to a smooth payoff and a linear combination of call and put payoff functions). The derivatives of q belong to \mathcal{A} because q is singular at only finitely many points. Since q is the solution of

$$\partial_t q(s, t) + \frac{1}{2} w^2 s^2 \partial_s^2 q(s, t) = 0, \quad q(s, T) = f(s),$$

we have

$$\mathcal{L}q(s, t) = -\frac{1}{2} (w^2 s^2 - \sigma(s, t)^2) \partial_s^2 q(s, t).$$

Under the Black-Scholes model with $\sigma(s, t) = vs$, the solution p^α of (2.1.1) with convex f is of the form (2.2.4) with $w = v\sqrt{1 + 2/\alpha}$. We have then

$$\mathcal{L}p^\alpha(s, t) = -\frac{v^2}{\alpha} s^2 \partial_s^2 p^\alpha(s, t). \quad (2.2.5)$$

2.2.2 Continuous control strategies

As already mention in the Introduction, we are interested in a trading strategy X of finite cost which approximates the benchmark strategy X^q . In order to make the tracking error small, a reasonable control would be based on the deviation between X^q and X . Let $Z^\kappa = (X^q - X)/\kappa$. We will consider $X = X^{b,c,\kappa}$ of the form

$$dX_t = \frac{1}{\kappa} \operatorname{sgn}(Z_t^\kappa) c(|Z_t^\kappa|, S_t, t) d\langle X^q \rangle_t - \kappa dL_t^\kappa + \kappa dR_t^\kappa, \quad X_{0+} = X_0^q \quad (2.2.6)$$

where c is a nonnegative Borel function, and L^κ and R^κ are nondecreasing processes such that

$$L_t^\kappa = \int_0^t 1_{\{Z_u^\kappa = -b(S_u, u)\}} dL_u^\kappa, \quad R_t^\kappa = \int_0^t 1_{\{Z_u^\kappa = b(S_u, u)\}} dR_u^\kappa, \quad |Z_t^\kappa| \leq b(S_t, t) \quad (2.2.7)$$

on $[0, T]$ for a positive Borel function b . The idea is to introduce a regular control part which pushes X up or down if Z^κ is positive or negative respectively at some rate determined by c , and a singular control part which keeps Z^κ within a stochastic interval determined by b .

The existence of the triplet $(X^{b,c,\kappa}, L^\kappa, R^\kappa)$ follows from that of a solution of a Skorokhod-type equation. Denote by \mathcal{B} the set of the positive functions b on $(0, \infty) \times [0, T]$ such that both b and $1/b$ belong to \mathcal{A} . For $b \in \mathcal{B}$, denote by \mathcal{C}_b the set of nonnegative and piecewise $C^{0,3,1}$ functions c on

$$\mathcal{D}_b := \{(x, s, t) \in \mathbb{R} \times (0, \infty) \times [0, T] ; |x| \leq b(s, t)\}$$

such that

1. for all (s, t) , $c(\cdot, s, t)$ are even:

$$c(x, s, t) = c(-x, s, t),$$

2. for all x , $c(x, \cdot)$ are $C^{3,1}$ and

$$\sup\{|\partial_s^i \partial_t^j c(x, S_t, t)|; t \in [0, T], x \in [-b(S_t, t), b(S_t, t)]\} < \infty \quad (2.2.8)$$

almost surely for each $i \in \{0, 1, 2, 3\}$ and $j \in \{0, 1\}$, and

3. for any compact set $A \subset \mathcal{D}_b$, there exists $K > 0$ such that

$$(x - y)(-\operatorname{sgn}(x)c(x, s, t) + \operatorname{sgn}(y)c(y, s, t)) \leq K|x - y|^2 \quad (2.2.9)$$

for all $(x, s, t), (y, s, t) \in A$.

For $b \in \mathcal{B}$ and $c \in \mathcal{C}_b$, by a fixed point argument thanks to the one-sided Lipschitz condition (2.2.9) (see e.g., [Tan79]), we can show that there exists a unique solution $(Z^\kappa, L^\kappa, R^\kappa)$ of a Skorokhod-type equation

$$dZ_t^\kappa = \frac{1}{\kappa} dX_t^q - \frac{1}{\kappa^2} \operatorname{sgn}(Z_t^\kappa) c(Z_t^\kappa, S_t, t) \nu(S_t, t)^2 dt + dL_t^\kappa - dR_t^\kappa, \quad Z_0^\kappa = 0 \quad (2.2.10)$$

with (2.2.7) on $[0, T]$. The strategy (2.2.6) is therefore well-defined by

$$X^{b,c,\kappa} = X^q - \kappa Z^\kappa$$

for each $b \in \mathcal{B}$ and $c \in \mathcal{C}_b$.

2.2.3 Hedging error

Denote by $\Pi^{b,c,\kappa}$ the associated wealth process of the strategy $X^{b,c,\kappa}$ with initial capital

$$\Pi_{0-} = \Pi_0^q + \kappa\lambda(S_0, 0)|\Delta X_0^q|. \quad (2.2.11)$$

The total variation of $X^{b,c,\kappa}$ is given by

$$d\|X^{b,c,\kappa}\|_t = \frac{1}{\kappa}c(Z_t^\kappa, S_t, t)d\langle X^q \rangle_t + \kappa[dL_t^\kappa + dR_t^\kappa].$$

Then, from (2.2.1) and (2.2.3), the associated tracking error $\mathcal{E}^{b,c,\kappa}$ is given by

$$\begin{aligned} \mathcal{E}_t^{b,c,\kappa} &= \Pi_t^q - \Pi_t^{b,c,\kappa} \\ &= \kappa \int_0^t Z_u^\kappa dS_u + \int_0^t \lambda(S_u, u)c(Z_u^\kappa, S_u, u)d\langle X^q \rangle_u \\ &\quad + \kappa^2 \int_0^t \lambda(S_u, u)[dL_u^\kappa + dR_u^\kappa] + \int_0^t \mathcal{L}q(S_u, u)du. \end{aligned} \quad (2.2.12)$$

Since $b \in \mathcal{B} \subset \mathcal{A}$, $\sup\{|Z_t|; t \in [0, T]\}$ is finite almost surely. Therefore, the process $\mathcal{E}^{b,c,\kappa}$ is well-defined as a $C[0, T]$ -valued random variable. The terminal value $\mathcal{E}_T^{b,c,\kappa} = f(S_T) - \Pi_T^{b,c,\kappa}$ represents the hedging error. The question is whether the tracking error converges to 0 as $\kappa \rightarrow 0$ with a good rate. If the answer is positive for a certain class of b and c , then the next question is which combination of b and c is optimal. We will answer these questions in the following sections.

2.3 Main results

2.3.1 Limit theorem of hedging error

Let $b \in \mathcal{B}$ and $c \in \mathcal{C}_b$. Define

$$\begin{aligned} g(x, s, t) &= \exp \left\{ -2 \int_0^{|x|} c(y, s, t) dy \right\}, \\ a(s, t) &= 2 \int_0^{b(s,t)} g(x, s, t) dx, \\ h(x, s, t) &= \frac{2 \operatorname{sgn}(x)}{g(x, s, t)} \int_0^{|x|} \left(c(z, s, t) - \frac{1}{a(s, t)} \right) g(z, s, t) dz. \end{aligned} \quad (2.3.1)$$

For notational simplicity, the dependence on b and c is omitted. See Lemma 2.4.1 for an interpretation of these quantities. We have the first main result.

Theorem 2.3.1. *Let $b \in \mathcal{B}$, $c \in \mathcal{C}_b$ and $\mathcal{E}^{b,c,\kappa}$ be the process of tracking error associated with the strategy $X^{b,c,\kappa}$ defined by (2.2.6) and (2.2.12). Then*

$$\kappa^{-1} \left(\mathcal{E}^{b,c,\kappa} - \int_0^\cdot \delta^{b,c}(S_t, t) dt \right) \rightarrow W_{Q^{b,c}}$$

stably in law on $C[0, T]$ as $\kappa \rightarrow 0$, where W is a standard Brownian motion independent of \mathcal{F} and

$$\begin{aligned} \delta^{b,c}(s, t) &= \frac{1}{a(s, t)} \lambda(s, t) |\partial_s^2 q(s, t)|^2 \sigma(s, t)^2 + \mathcal{L}q(s, t), \\ Q^{b,c} &= \int_0^\cdot \eta^{b,c}(S_t, t) d\langle S \rangle_t, \\ \eta^{b,c}(s, t) &= \frac{2}{a(s, t)} \int_0^{b(s,t)} (x - \lambda(s, t) \partial_s^2 q(s, t) h(x, s, t))^2 g(x, s, t) dx. \end{aligned} \quad (2.3.2)$$

In particular, for any q with $\mathcal{L}q < 0$ on $(0, \infty) \times [0, T]$, by taking $b \in \mathcal{B}$ and $c \in \mathcal{C}_b$ such that

$$a \geq \frac{\lambda |\partial_s^2 q|^2 \sigma^2}{|\mathcal{L}q|},$$

we have $\delta^{b,c} \leq 0$ and therefore an asymptotic super replication strategy:

$$\Pi_T^{b,c,\kappa} \rightarrow f(S_T) - \int_0^T \delta^{b,c}(S_t, t) dt \geq f(S_T)$$

as $\kappa \rightarrow 0$. This is always possible; say, let

$$b(s, t) = \frac{1}{2} \frac{\lambda(s, t) |\partial_s^2 q(s, t)|^2 \sigma(s, t)^2}{|\mathcal{L}q(s, t)|} + \epsilon, \quad c(x, s, t) = 0 \quad (2.3.3)$$

with $\epsilon > 0$. Further if $1/\partial_s^2 q, 1/\mathcal{L}q \in \mathcal{A}$, one can take $\epsilon = 0$ to have $\delta^{b,c} = 0$.

Example 2.3.1 (Black-Scholes model with convex payoff). Let $\sigma(s, t) = vs$, $v > 0$, $\lambda(s, t) = s$, f be a piecewise C^2 convex non-affine function and $\alpha > 0$ be a constant. Let q be given by (2.2.4) with $w = v\sqrt{1 + 2/\alpha} > v$. As explained in Introduction, the original Leland strategy uses the equidistant discretization of X^q with (2.1.6). The renormalized tracking error at time $t \in [0, T]$ converges in law to the mixed normal distribution with mean 0 and variance

$$\eta_L(\alpha) \int_0^t |S_u \Gamma_u^q|^2 d\langle S \rangle_u.$$

The use of the hitting times (2.1.9) changes the asymptotic variance to

$$\eta_F(\alpha) \int_0^t |S_u \Gamma_u^q|^2 d\langle S \rangle_u$$

without changing the initial capital $q(S_0, 0) + \kappa S_0 |\Delta X_0^q|$. Now, let us consider the simplest control strategy in our framework for the same function q . We have

$$\mathcal{L}q(s, t) = -\frac{1}{\alpha} v^2 s^2 \partial_s^2 q(s, t) < 0, \quad (s, t) \in (0, \infty) \times [0, T]$$

Let $b \in \mathcal{B}$ and $c = 0 \in \mathcal{C}_b$. Then, we have

$$g(z, s, t) = 1, \quad a(s, t) = 2b(s, t), \quad h(x, s, t) = -x/b(s, t)$$

by definition and so,

$$\delta^{b,c}(s, t) = v^2 s^2 \left\{ \frac{s |\partial_s^2 q(s, t)|^2}{2b} - \frac{|\partial_s^2 q(s, t)|}{\alpha} \right\}, \quad \eta^{b,c}(s, t) = \frac{1}{3} \left(b(s, t) + s \partial_s^2 q(s, t) \right)^2.$$

With the same initial capital $q(S_0, 0) + \kappa S_0 |\Delta X_0^q|$ as before and taking b and c as (2.3.3), that is,

$$b(s, t) = \frac{\alpha}{2} s |\partial_s^2 q(s, t)| + \epsilon, \quad c(x, s, t) = 0, \quad (2.3.4)$$

the tracking error has the asymptotic variance

$$\frac{(\alpha + 2)^2}{12} \int_0^t |S_u \Gamma_u^q|^2 d\langle S \rangle_u + \frac{\epsilon}{3} (\alpha + 2) \int_0^t S_u \Gamma_u^q d\langle S \rangle_u + \frac{\epsilon^2}{3} \langle S \rangle_T$$

with the asymptotic positive surplus $-\delta^{b,c} > 0$. It is easy to see

$$\frac{(\alpha + 2)^2}{12} < \min\{\eta_L(\alpha), \eta_F(\alpha)\}$$

for all $\alpha > 0$ and so, the simplest singular control with sufficiently small $\epsilon > 0$ is already more preferable than the existing ones. This superiority continues to hold even if we consider a time-varying α as in [DK10]. We will minimize $\eta^{b,c}$ in the next section.

Example 2.3.2 (Local volatility model with convex payoff). The Black-Scholes pricing function q given by (2.2.4) is useful under a more general local volatility model. Suppose a volatility bound $\sigma(s, t) < vs$ is given for a constant $v > 0$. Let

$$\alpha(s, t) = \frac{2\sigma(s, t)^2}{v^2s^2 - \sigma(s, t)^2}.$$

Then, the function q given by (2.2.4) solves (2.1.1). When f is piecewise C^2 , non-affine and convex, we have

$$\mathcal{L}q(s, t) = -\frac{1}{\alpha(s, t)}\sigma(s, t)^2\partial_s^2q(s, t) < 0.$$

If we take

$$b(s, t) = \frac{1}{2}\alpha(s, t)s|\partial_s^2q(s, t)| + \epsilon, \quad c(x, s, t) = 0,$$

then, we have

$$\int_0^\cdot \delta^{b,c}(S_t, t)dt = -\epsilon \int_0^\cdot \frac{2}{\alpha(S_t, t)} \frac{|\Gamma_t^q|^2}{\alpha(S_t, t)S_t|\Gamma_t^q| + 2\epsilon} d\langle S \rangle_t < 0 \quad (2.3.5)$$

and

$$\int_0^\cdot \eta^{b,c}(S_t, t)d\langle S \rangle_t = \frac{1}{12} \int_0^\cdot ((\alpha(S_t, t) + 2)S_t\Gamma_t^q + 2\epsilon)^2 d\langle S \rangle_t.$$

Example 2.3.3 (Local volatility model with concave payoff). The preceding two examples are for a convex payoff f . It is not difficult to treat a concave payoff in a similar way. Under the Black-Scholes model of Example 3.1, consider the Black-Scholes function q given by (2.2.4) with $w = v\sqrt{1 - 2/\alpha}$ for $\alpha > 2$. Then,

$$\mathcal{L}q(s, t) = \frac{1}{\alpha}v^2s^2\partial_s^2q(s, t) < 0.$$

if f is piecewise C^2 , non-affine and concave. The strategy (2.3.4) results in the asymptotic positive surplus $-\delta^{b,c} > 0$ with asymptotic variance

$$\frac{(\alpha - 2)^2}{12} \int_0^t |S_u\Gamma_u^q|^2 d\langle S \rangle_u + \frac{\epsilon}{3}(\alpha - 2) \int_0^t S_u\Gamma_u^q d\langle S \rangle_u + \frac{\epsilon^2}{3}\langle S \rangle_T.$$

Remark that when $\alpha = 2$, we have $q(s, 0) = f(s)$, which is the capital required by a trivial static hedging strategy for the concave payoff f . Under a general local volatility model with a lower bound $\sigma(s, t) > v^2s^2$, $v > 0$, the strategy of Example 3.2 works with

$$\alpha(s, t) = \frac{2\sigma(s, t)^2}{\sigma(s, t)^2 - v^2s^2}$$

and (2.3.5), while the asymptotic variance changes to

$$\int_0^\cdot \eta^{b,c}(S_t, t)d\langle S \rangle_t = \frac{1}{12} \int_0^\cdot ((\alpha(S_t, t) - 2)S_t|\Gamma_t^q| + 2\epsilon)^2 d\langle S \rangle_t.$$

Example 2.3.4 (The Barles-Soner strategy). [BS98] studied a scaling limit of the utility indifference price under $\sigma(s, t) = vs$, $v > 0$ and $\lambda(s, t) = s$. As the asymptotically optimal strategy, they derived a continuous control strategy (2.2.6) (2.2.7) with

$$b(s, t) = \frac{g(\alpha^2s^2\partial_s^2q(s, t))}{\alpha^2s}, \quad c(z, s, t) = 0,$$

where $\alpha > 0$ is a parameter of the scaling limit and

$$g(x) = \sqrt{xh(x)} - x, \quad x \in \mathbb{R}.$$

The function h is the unique solution of the ordinary differential equation

$$h'(x) = \frac{1 + h(x)}{2\sqrt{xh(x)} - x}, \quad h(0) = 0. \quad (2.3.6)$$

The benchmark strategy $X_t^q = \partial_s q(S_t, t)$ is determined by the solution q of a nonlinear PDE

$$\partial_t q(s, t) + \frac{1}{2}v^2 s^2 \partial_s^2 q(s, t)(1 + h(\alpha^2 s^2 \partial_s^2 q)) = 0, \quad q(s, T) = f(s). \quad (2.3.7)$$

If q is enough regular that our assumptions are satisfied, then we can apply our result with

$$\delta^{b,c}(s, t) = \frac{1}{2}v^2 \frac{A}{\alpha^2} \left\{ \frac{A}{g(A)} - h(A) \right\}, \quad \eta^{b,c}(s, t) = \frac{1}{3} \frac{(g(A) + A)^2}{\alpha^4 s^2} = \frac{1}{3} \frac{Ah(A)}{\alpha^4 s^2},$$

where $A = \alpha^2 s^2 \partial_s^2 q(s, t)$. Remark that the PDE (2.3.7) has an explicit smooth solution

$$q(s, t) = -k \log s + \frac{1}{2}v^2 k(1 + h(\alpha^2 k))(T - t) \quad (2.3.8)$$

when $f(s) = -k \log s$, $k \in \mathbb{R}$. It turns out that $\delta^{b,c} \neq 0$, which means that the strategy is not asymptotically replicating; for an asymptotic replication of order κ , the function h has to satisfy $h(x) = x/g(x)$, which leads to an equation similar to but different from (2.3.6):

$$h'(x) = \frac{1 + h(x)}{3\sqrt{xh(x)} - 2x}, \quad h(0) = 0.$$

2.3.2 Minimum conditional variance

By Theorem 2.3.1, the law of the hedging error at the maturity T associated with the strategy (2.2.6) is approximated by the mixed normal distribution with conditional mean

$$\int_0^T \delta^{b,c}(S_u, u) du$$

and conditional variance

$$\kappa^2 Q_T^{b,c} = \kappa^2 \int_0^T \eta^{b,c}(S_u, u) d\langle S \rangle_u.$$

If the bias $\delta^{b,c}$ is zero, then the strategy is asymptotically replicating with rate κ . If the bias is negative, the strategy is asymptotically super-replicating. A natural problem is to minimize the variance $Q^{b,c}$ while fixing the bias $\delta^{b,c} \equiv \delta$. In the case $\delta = 0$, the strategy minimizing $Q^{b,c}$, if exists, minimizes also

$$\lim_{\kappa \rightarrow \infty} E[l(\kappa^{-1} \mathcal{E}_T^{b,c,\kappa})] = \int E[l(\sqrt{Q_T^{b,c}} z)] \phi(z) dz$$

for any nonnegative bounded continuous quasi-convex function l with $l(0) = 0$. Recall that q is fixed. Since there is a one-to-one correspondence between $\delta^{b,c}$ and a by (2.3.2), it suffices to minimize $Q^{b,c}$ among b and c with a fixed. Therefore, for given $a \in \mathcal{B}$, we consider

$$\mathcal{S}_a = \left\{ (b, c) \in \mathcal{B} \times \mathcal{C}_b \mid a(s, t) = 2 \int_0^{b(s,t)} g(x, s, t) dx, \right. \\ \left. g(x, s, t) = \exp \left\{ -2 \int_0^{|x|} c(y, s, t) dy \right\} \right\}. \quad (2.3.9)$$

Note that

$$Q^{b,c} \geq \int_0^T \eta_*^a(S_u, u) d\langle S \rangle_u, \quad \eta_*^a(s, t) = \inf_{(b,c) \in \mathcal{S}_a} \eta^{b,c}(s, t)$$

for any $(b, c) \in \mathcal{S}_a$.

Definition 2.3.1. Let $a \in \mathcal{B}$. We say that a sequence of strategies $(b_n, c_n) \in \mathcal{S}_a$, $n \in \mathbb{N}$ is asymptotically optimal if

$$\limsup_{n \rightarrow \infty} \eta^{b_n, c_n}(s, t) = \eta_*^a(s, t) \quad (2.3.10)$$

for all $(s, t) \in (0, \infty) \times [0, T]$ and

$$\int_0^t \eta^{b_n, c_n}(S_u, u) d\langle S \rangle_u \rightarrow \int_0^t \eta_*^a(S_u, u) d\langle S \rangle_u$$

in probability for all $t \in [0, T]$.

Below is the second main result of this chapter.

Theorem 2.3.2. (i) Let

$$\eta_{\dagger}(x) = \begin{cases} 0 & \text{if } -2 < x \leq 1, \\ \eta_1(x) & \text{if } 1 < x < 2, \\ \eta_2(x) & \text{if } |x| \geq 2, \end{cases}$$

where

$$\eta_1(x) = \frac{4(x+2)^2(x-1)}{3x^3(4-x)}, \quad \eta_2(x) = \frac{1}{12}(x+2)^2,$$

and denote

$$\gamma(s, t) = \lambda(s, t) \partial_s^2 q(s, t). \quad (2.3.11)$$

Then for all $(s, t) \in (0, \infty) \times [0, T]$,

$$\eta_*^a(s, t) = |\gamma(s, t)|^2 \eta_{\dagger}\left(\frac{a(s, t)}{\gamma(s, t)}\right)$$

if $\gamma(s, t) \neq 0$, and

$$\eta_*^a(s, t) = \lim_{\epsilon \rightarrow 0} \epsilon^2 \eta_{\dagger}\left(\frac{a(s, t)}{\epsilon}\right) = \frac{a(s, t)^2}{12}$$

otherwise.

(ii) There exists an asymptotically optimal sequence (b_n, c_n) .

The function $\alpha \rightarrow \eta_{\dagger}(\alpha)$ is represented in Fig. 2.3 in Introduction.

2.3.3 Asymptotically optimal sequence

Here, we construct a sequence of continuous trading strategies (b_n^*, c_n^*) which can be considered essentially asymptotically optimal. This is not necessarily asymptotically optimal in the precise sense of Definition 2.3.1 because each (b_n^*, c_n^*) is not necessarily satisfying the technical conditions in Section 2.2.2. It means that the strategy associated with (b_n^*, c_n^*) is not well-defined in general. However in the proof of Theorem 2.3.2 in Section 4.5, we show that there exists a sequence $(b_n, c_n) \in \mathcal{S}_a$ which is an approximation to (b_n^*, c_n^*) in a suitable sense and in fact asymptotically optimal. Therefore, from practical point of view, one may use (b_n^*, c_n^*) defined below as an optimal strategy.

Define $l = l(s, t)$ as follows:

$$l = \begin{cases} 0 & \text{if } \gamma \geq a \\ 2\gamma(a - \gamma)/(4\gamma - a) & \text{if } a > \gamma \geq a/2 \\ \gamma & \text{if } a/2 > \gamma \geq 0 \\ 0 & \text{if } 0 > \gamma, \end{cases} \quad (2.3.12)$$

where $a = a(s, t)$ and $\gamma = \gamma(s, t)$ (see (2.3.11)). Let $\psi_n = 1 - 2/n$ and

$$r_n = l + e^{n^2}(\gamma - a/2)_+^5 + |\gamma|(1 - \exp\{-n(\gamma + a/2)_-\}^5).$$

Consider the following sequence of (b_n^*, c_n^*) :

$$b_n^*(s, t) = r_n(s, t) + \frac{\gamma(s, t) + \psi_n r_n(s, t)}{\gamma(s, t) + \psi_n l(s, t)} \left(\frac{a(s, t)}{2} - l(s, t) \right), \quad (2.3.13)$$

$$c_n^*(x, s, t) = \frac{\psi_n a(s, t) + 2\gamma(s, t)}{2(a(s, t) - 2l(s, t))} \frac{1}{\gamma(s, t) + \psi_n |x|} \mathbf{1}_{[l(s, t), r_n(s, t)]}(|x|). \quad (2.3.14)$$

Remark 2.3.1. With c_n^* given as in (2.3.14), b_n^* is derived from

$$a(s, t) = 2 \int_0^{b_n^*(s, t)} g_n^*(x, s, t) dx. \quad (2.3.15)$$

where g_n^* is defined by (2.3.1).

In terms of γ and a , this strategy can be divided into four cases.

1. If $\gamma(s, t) < -a(s, t)/2$, then $l(s, t) = 0$ and

$$b_n^*(s, t) = \left(1 + \frac{\psi_n a(s, t)}{2\gamma(s, t)} \right) r_n(s, t) + \frac{a(s, t)}{2},$$

$$c_n^*(x, s, t) = \frac{\psi_n a(s, t) + 2\gamma(s, t)}{2a(s, t)} \frac{1}{\gamma(s, t) + \psi_n |x|} \mathbf{1}_{\{|x| < r_n(s, t)\}},$$

We have $r_n < \min\{b_n^*, |\gamma|\}$, hence $r_n, b_n^* \rightarrow |\gamma|$ and $b_n^* - r_n \rightarrow 0$. The strategy is essentially a regular control with singularity at $x = \pm\gamma$; see Fig. 2.4a.

2. if $-a(s, t)/2 \leq \gamma(s, t) \leq a(s, t)/2$, then $l(s, t) = r_n(s, t)$ and so,

$$b_n^*(s, t) = \frac{a(s, t)}{2}, \quad c_n^*(x, s, t) = 0,$$

The strategy has only a singular part and is given by Fig. 2.4b. Note that the strategy does not depend on n .

3. If $a(s, t)/2 < \gamma(s, t) < a(s, t)$, then for sufficiently large n ,

$$b_n^*(s, t) = r_n(s, t) + \frac{(2\gamma(s, t) - a(s, t))(2\gamma(s, t) + a(s, t))}{2(4\gamma(s, t) - a(s, t))} \frac{\gamma(s, t) + r_n(s, t)}{\gamma(s, t) + l(s, t)},$$

$$c_n^*(x, s, t) = \frac{\psi_n a(s, t) + 2\gamma(s, t)}{a(s, t) + 2\gamma(s, t)} \frac{4\gamma(s, t) - a(s, t)}{2(2\gamma(s, t) - a(s, t))} \frac{1}{\gamma(s, t) + |x|} \mathbf{1}_{\{l(s, t) \leq |x| < r_n(s, t)\}},$$

and we have $r_n, b_n^*, b_n^* - r_n \rightarrow \infty$. The strategy is essentially a regular control with non-intervention zone near zero; see Fig. 2.4c.

4. If $\gamma(s, t) \geq a(s, t)$, then $l(s, t) = 0$ and

$$b_n^*(s, t) = r_n(s, t) + \frac{\psi_n a(s, t)}{2\gamma(s, t)} (\gamma(s, t) + r_n(s, t)),$$

$$c_n^*(x, s, t) = \frac{\psi_n a(s, t) + 2\gamma(s, t)}{2a(s, t)} \frac{1}{\gamma(s, t) + \psi_n |x|} \mathbf{1}_{\{|x| < r_n(s, t)\}},$$

We have $r_n, b_n^*, b_n^* - r_n \rightarrow \infty$. The strategy is again, essentially a regular control; see Fig. 2.4d.

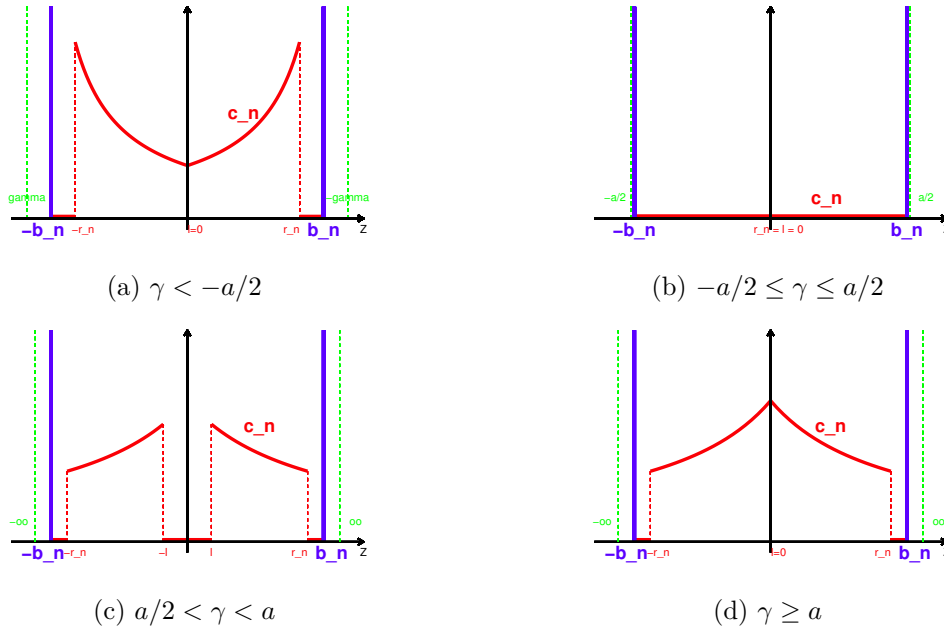


Figure 2.4 – Qualitative behavior of asymptotically optimal strategies

Remark 2.3.2. The Barles-Soner strategy is not always asymptotically optimal in the sense of Definition 2.3.1. Here, we use the notation in Example 2.3.4. The strategy is a purely singular control and therefore, asymptotically optimal only if

$$s|\partial_s^2 q(s, t)| = |\gamma(s, t)| \leq \frac{a(s, t)}{2} = b(s, t) = \frac{g(\alpha^2 s^2 \partial_s^2 q(s, t))}{\alpha^2 s}$$

for all (s, t) ; see case 2 above. This is equivalent to say

$$\frac{g(A)}{|A|} \geq 2, \quad A = \alpha^2 s^2 \partial_s^2 q(s, t). \quad (2.3.16)$$

According to [BS98], $\lim_{A \rightarrow \infty} h(A)/A = 1$, which implies that $\lim_{A \rightarrow \infty} g(A)/A = 0$. Therefore, the condition (2.3.16) is violated, say, for a log contract (2.3.8) with $\alpha^2 k$ sufficiently large.

2.3.4 Implementation and numerical experiments

Here we describe the implementation of our new results in the specific case that the payoff is convex and the benchmark is given by a constant enlargement of volatility. Let f be convex and $q = p^\alpha$ be the solution of (2.1.1). Consider the case $\delta \equiv 0$. By (2.2.5) and (2.3.2), we have $a = \alpha \lambda \partial_s^2 p^\alpha$. Hence

$$\gamma(s, t) = \frac{1}{\alpha} a(s, t).$$

The situation can be divided into three cases: $0 < \alpha \leq 1$, $1 < \alpha < 2$ and $\alpha \geq 2$.

The case where $0 < \alpha \leq 1$ corresponds to Fig. 2.4d. The optimal strategy X is essentially given by

$$dX_t = \text{sgn}(X_t^\alpha - X_t) \frac{\alpha + 2}{2\alpha} \frac{|\Gamma_t^\alpha|^2}{\kappa \lambda(S_t, t) \Gamma_t^\alpha + |X_t^\alpha - X_t|} \sigma(S_t, t)^2 dt. \quad (2.3.17)$$

Denoting by Π the wealth process associated with X , we have

$$\kappa^{-1}(\Pi^\alpha - \Pi) \rightarrow 0$$

uniformly on $[0, T]$ in probability as $\kappa \rightarrow 0$. This means that the hedging error under (2.3.17) is of a smaller order than κ while, as explained above, Leland's strategy has a nondegenerate limit of order κ . The strategy (2.3.17) can be implemented as, say, $X_t = Y_j$ for $t \in (j\delta, (j+1)\delta]$,

$$Y_j = \begin{cases} \hat{Y}_j & \text{if } (\hat{Y}_j - Y_{j-1})(X_{j\delta}^\alpha - \hat{Y}_j) \geq 0, \\ X_{j\delta}^\alpha & \text{otherwise,} \end{cases}$$

where

$$\hat{Y}_j = Y_{j-1} + \text{sgn}(X_{j\delta}^\alpha - Y_{j-1}) \frac{\alpha + 2}{2\alpha} \frac{|\Gamma_{j\delta}^\alpha|^2}{\kappa \lambda(S_{j\delta}, j\delta) \Gamma_{j\delta}^\alpha + |X_{j\delta}^\alpha - Y_{j-1}|} \sigma(S_{j\delta}, j\delta)^2 \delta$$

with a sufficiently small $\delta > 0$. This is a discretized process of a process of finite variation and so, taking δ very small is not costly under the proportional transaction costs and makes the discretization error negligible.

The case where $1 < \alpha < 2$ corresponds to Fig. 2.4c and the optimal strategy X is essentially given by

$$dX_t = \text{sgn}(X_t^\alpha - X_t) \frac{4 - \alpha}{2(2 - \alpha)} \frac{|\Gamma_t^\alpha|^2}{\kappa \lambda(S_t, t) \Gamma_t^\alpha + |X_t^\alpha - X_t|} 1_{\{|X_t^\alpha - X_t| \geq \kappa \hat{\alpha} \lambda(S_t, t) \Gamma_t^\alpha\}} \sigma(S_t, t)^2 dt,$$

where $\hat{\alpha} = 2(\alpha - 1)/(4 - \alpha)$. The associated central limit theorem for $\Pi = \Pi^{\alpha, \kappa}$ is (2.1.8) with $\eta_L(\alpha)$ replaced by

$$\eta_{\dagger}(\alpha) = \frac{4}{3} \frac{(\alpha + 2)^2 (\alpha - 1)}{\alpha^3 (4 - \alpha)}.$$

See the left panel of Fig. 2.3 on the region $1 < \alpha < 2$. The implementation of the strategy can be done in a way similar to (2.3.17): $X_t = Y_j$ for $t \in (j\delta, (j+1)\delta]$,

$$Y_j = \begin{cases} \hat{Y}_j & \text{if } |X_{j\delta}^\alpha - \hat{Y}_j| \geq \kappa \hat{\alpha} \Lambda_{j\delta} \Gamma_{j\delta}^\alpha \text{ or } \hat{Y}_j = Y_{j-1}, \\ X_{j\delta}^\alpha + \text{sgn}(Y_{j-1} - X_{j\delta}^\alpha) \kappa \hat{\alpha} \Lambda_{j\delta} \Gamma_{j\delta}^\alpha & \text{otherwise,} \end{cases}$$

where $\Lambda_{j\delta} = \lambda(S_{j\delta}, j\delta)$ and

$$\begin{aligned} \hat{Y}_j &= Y_{j-1} + \text{sgn}(X_{j\delta}^\alpha - Y_{j-1}) \frac{4 - \alpha}{2(2 - \alpha)} \frac{|\Gamma_{j\delta}^\alpha|^2}{\kappa \Lambda_{j\delta} \Gamma_{j\delta}^\alpha + |X_{j\delta}^\alpha - Y_{j-1}|} \sigma(S_{j\delta}, j\delta)^2 \delta \\ &\quad \times 1_{\{|X_{j\delta}^\alpha - Y_{j-1}| \geq \kappa \hat{\alpha} \Lambda_{j\delta} \Gamma_{j\delta}^\alpha\}} \end{aligned}$$

with a sufficiently small $\delta > 0$.

The case where $\alpha \geq 2$ corresponds to Fig. 2.4b and the optimal strategy is

$$X_t = X_0^\alpha - \kappa L_t^\kappa + \kappa R_t^\kappa \tag{2.3.18}$$

where L^κ and R^κ are nondecreasing processes such that

$$L_t^\kappa = \int_0^t 1_{\{X_u^\alpha - X_u = -\kappa b(S_u, u)\}} dL_u^\kappa, \quad R_t^\kappa = \int_0^t 1_{\{X_u^\alpha - X_u = \kappa b(S_u, u)\}} dR_u^\kappa$$

and

$$|X_u^\alpha - X_u| \leq \kappa b(S_t, t), \quad b(S_t, t) = \frac{\alpha}{2} \lambda(S_t, t) \Gamma_t^\alpha.$$

This is a singular control strategy which keeps the distance between X_t and X_t^α less than $\kappa b(S_t, t)$. The associated central limit theorem for $\Pi = \Pi^{\alpha, \kappa}$ is (2.1.8) with $\eta_L(\alpha)$ replaced by

$$\eta_{\dagger}(\alpha) = \frac{(\alpha + 2)^2}{12}.$$

See Fig. 2.3 on the region $\alpha \geq 2$. The function η_{\dagger} is continuously concatenated at $\alpha = 2$. It is interesting to note that the hitting barrier b_t in (2.1.9) coincides with $2b(S_t, t)$ and $\eta_F(\alpha) = 2\eta_{\dagger}(\alpha)$ for $\alpha \geq 2$. The strategy (2.3.18) can be implemented as

$$X_{j\delta} = \begin{cases} X_{j\delta}^{\alpha} + \kappa b(S_{j\delta}, j\delta) & \text{if } X_{j\delta}^{\alpha} - X_{(j-1)\delta} < -\kappa b(S_{j\delta}, j\delta), \\ X_{j\delta}^{\alpha} - \kappa b(S_{j\delta}, j\delta) & \text{if } X_{j\delta}^{\alpha} - X_{(j-1)\delta} > \kappa b(S_{j\delta}, j\delta), \\ X_{(j-1)\delta} & \text{otherwise} \end{cases}$$

with a sufficiently small $\delta > 0$.

Now we examine numerically the validity of the discretization methods of the optimal strategies explained in Section 1.3. We take

$$dS_t = S_t(0.1dt + 0.2dB_t), \quad S_0 = 100$$

with small proportional transaction costs : $\kappa = 0.01$ and $\lambda(s, t) = s$, and a call payoff $f(s) = (s - 110)_+$ with maturity $T = 1$. Here we consider

$$\alpha \in \{0.5, 1.5, 2.5\}.$$

For each case, we calculate the hedging error at the maturity from 10000 paths simulated from the model. We choose the discretization step size $\delta = 10^{-5}$, which corresponds to about one minute in terms of trading time. The hedging error of Leland's strategy explained in the introduction is calculated for comparison. Then we draw the histogram of normalized hedging error

$$\frac{f(S_T) - \Pi_T}{\kappa \sqrt{\int_0^T S_t^2 |\Gamma_t^\alpha|^2 d\langle S \rangle_t}}.$$

Fig. 2.5a is for the case $\alpha = 0.5$. The left and right are respectively for the discretized optimal strategy and for Leland's strategy. The green curves represent the normal density with mean 0 and variance $\eta_L(\alpha)$ that corresponds to the limit distribution for Leland's strategy. First we notice that the limit distribution for Leland's strategy approximates quite well to the corresponding hedging error distribution with a realistic size of transaction coefficient $\kappa = 0.01$. Next, observe that the normalized hedging error of the discretized optimal strategy is more concentrated to its mean value, which is consistent to our theoretical result that the limit is degenerate. We however observe a negative bias which has not been explained by the theory. Remark that this negative bias means the strategy tends to be super-hedging for the seller of the option.

Figures 2.5b and 2.5c are respectively for the cases $\alpha = 1.5$ and $\alpha = 2.5$. The red curves represent the normal density with mean 0 and variance $\eta(\alpha)$ that corresponds to the limit distribution for the optimal strategy in each case. The normalized hedging error distributions are quite well approximated by their theoretical limits. This suggests that the discretization error with $\delta = 10^{-5}$ is negligible and the proposed strategies in fact improve Leland's strategy significantly.

2.4 Heuristic derivation

In this section, we provide a probabilistic interpretation for Theorem 2.3.1, and explain the relation with the equidistant strategy (2.1.4) of [Lel85] and the random time strategy (2.1.9) of [Fuk11b]. We will also compare the difference between Leland's approach and the indifference pricing approach, the latter being based on utility maximization (cf. [WHW94, Bic12, ST13, KL13]).

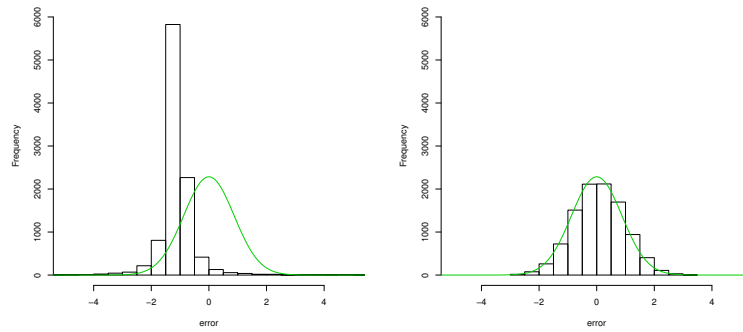
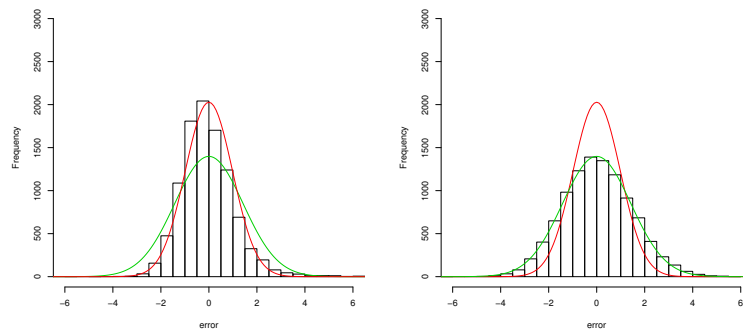
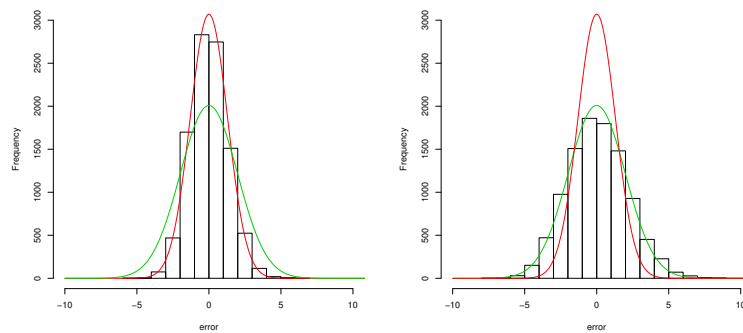
(a) Histograms for the optimal (left) and Leland's strategy (right) : $\alpha = 0.5$.(b) Histograms for the optimal (left) and Leland's strategy (right) : $\alpha = 1.5$.(c) Histograms for the optimal (left) and Leland's strategy (right) : $\alpha = 2.5$.

Figure 2.5 – Histograms for optimal and Leland's strategies.

2.4.1 Local probability model

Let $X^{b,c,\kappa}$ be the continuous replication strategy based on b and c under proportional cost parameter κ . Recall that the hedging error is given by

$$\begin{aligned} \mathcal{E}_t^{b,c,\kappa} &= \kappa \int_0^t Z_u^\kappa dS_u + \int_0^t \lambda(S_u, u) c(Z_u^\kappa, S_u, u) d\langle X^q \rangle_u \\ &\quad + \kappa^2 \int_0^t \lambda(S_u, u) (dL_u^\kappa + dR_u^\kappa) + \int_0^t \mathcal{L}q(S_u, u) du. \end{aligned}$$

Consider a decomposition of the horizon $[0, T)$ into small intervals $[u_j, u_{j+1})$ where $u_j = j\kappa, j = 0, 1, \dots, N_T^\kappa = \lfloor T/\kappa \rfloor$. Then we have

$$\mathcal{E}_t^{b,c,\kappa} = \sum_j \frac{1}{\kappa} \left(\mathcal{E}_{(j+1)\kappa}^{b,c,\kappa} - \mathcal{E}_{j\kappa}^{b,c,\kappa} \right) (u_{j+1} - u_j).$$

Therefore the total hedging error can be seen as the Riemann sum of local hedging error over the horizon $[u_j, u_{j+1})$:

$$\frac{1}{\kappa} \left(\mathcal{E}_{(j+1)\kappa}^{b,c,\kappa} - \mathcal{E}_{j\kappa}^{b,c,\kappa} \right). \quad (2.4.1)$$

Now let's take a closer look at (2.4.1). Fix j and perform the following time scaling

$$\tilde{Z}_t^\kappa = Z_{u_j + \kappa^2 t}^\kappa, \quad \tilde{L}_t^\kappa = L_{u_j + \kappa^2 t}^\kappa, \quad \tilde{R}_t^\kappa = R_{u_j + \kappa^2 t}^\kappa,$$

and

$$\tilde{B}_t^\kappa = \frac{1}{\kappa} B_{u_j + \kappa^2 t}$$

where W is the driving Brownian motion of S and $t \in [0, T^\kappa = \frac{1}{\kappa})$. We obtain

$$\begin{aligned} &\frac{1}{\kappa} \left(\mathcal{E}_{(j+1)\kappa}^{b,c,\kappa} - \mathcal{E}_{j\kappa}^{b,c,\kappa} \right) \\ &= \int_{j\kappa}^{(j+1)\kappa} Z_u^\kappa dS_u + \frac{1}{\kappa} \int_{j\kappa}^{(j+1)\kappa} \lambda(S_u, u) c(Z_u^\kappa, S_u, u) d\langle X^q \rangle_u \\ &\quad + \kappa \int_{j\kappa}^{(j+1)\kappa} \lambda(S_u, u) (dL_u^\kappa + dR_u^\kappa) + \frac{1}{\kappa} \int_{j\kappa}^{(j+1)\kappa} \mathcal{L}q(S_u, u) du \\ &\simeq \frac{1}{T^\kappa} A_{j\kappa} \int_0^{T^\kappa} \tilde{Z}_t^\kappa d\tilde{B}_t^\kappa + \frac{1}{T^\kappa} K_{j\kappa} \zeta_{j\kappa}^2 \int_0^{T^\kappa} c(\tilde{Z}_t^\kappa, S_{j\kappa}, j\kappa) dt \\ &\quad + \frac{1}{T^\kappa} K_{j\kappa} \int_0^{T^\kappa} (d\tilde{L}_t^\kappa + d\tilde{R}_t^\kappa) + \mathcal{L}q(S_{j\kappa}, j\kappa), \end{aligned}$$

where

$$A_t = \sigma(S_t, t), \quad K_t = \lambda(S_t, t), \quad \zeta_t = \sigma(S_t, t) \Gamma_t^q \quad (2.4.2)$$

and the dynamic of \tilde{Z}^κ is given by

$$\begin{aligned} d\tilde{Z}_t^\kappa &= \sigma(S_{u_j + \kappa^2 t}, u_j + \kappa^2 t) \Gamma_{u_j + \kappa^2 t}^q d\tilde{B}_t^\kappa \\ &\quad - \text{sgn}(\tilde{Z}_t^\kappa) c(\tilde{Z}_t^\kappa, S_{u_j + \kappa^2 t}, u_j + \kappa^2 t) \zeta_{u_j + \kappa^2 t}^2 dt + d\tilde{L}_t^\kappa - d\tilde{R}_t^\kappa, \\ &\simeq \zeta_{j\kappa} d\tilde{B}_t^\kappa - \text{sgn}(\tilde{Z}_t^\kappa) c(\tilde{Z}_t^\kappa, S_{j\kappa}, j\kappa) \zeta_{j\kappa}^2 dt + d\tilde{L}_t^\kappa - d\tilde{R}_t^\kappa, \end{aligned}$$

for $t \in [0, T^\kappa)$. As the time scale of \tilde{Z}^κ is faster than S , we are led to consider the following model of controlled Brownian motion with constant coefficients.

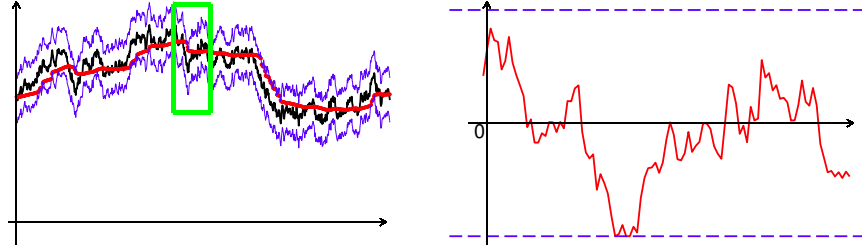


Figure 2.6 – Rescaling of tracking error Z^κ on $[j\kappa, (j+1)\kappa]$ (green zone).

Lemma 2.4.1 (Local probability model). *Let $b \in \mathbb{R}^+$ and $c : \mathbb{R} \rightarrow \mathbb{R}^+$ an even function. Let Z_t be a controlled Brownian motion with drift $-\text{sgn}(Z_t)c(Z_t)\zeta^2$ and reflection at $\pm b$, i.e.*

$$dZ_t = \zeta dB_t - \text{sgn}(Z_t)c(Z_t)\zeta^2 dt + dL_t - dR_t, \quad (2.4.3)$$

see right panel of Fig.2.6. Then we have

$$\begin{aligned} (LLN) \quad & \frac{1}{T} \left(K\zeta^2 \int_0^T c(Z_t) dt + K(L_T + R_T) \right) \rightarrow m, \\ (CLT) \quad & \frac{1}{\sqrt{T}} \left(A \int_0^T Z_t dB_t + K\zeta^2 \int_0^T c(Z_t) dt + K(L_T + R_T) - mT \right) \Rightarrow \mathcal{N}(0, v), \end{aligned}$$

where

$$m = K \frac{\zeta^2}{a}, \quad v = \frac{2}{a} \int_0^b (Az - K\zeta h(z))^2 g(z) dz.$$

and

$$g(z) = \exp\left(-2 \int_0^{|z|} c(x) dx\right), \quad a = 2 \int_0^b g(z) dz, \quad h(z) = \frac{2\text{sgn}(z)}{g(z)} \int_0^{|z|} (c(x) - \frac{1}{a}) g(x) dx.$$

In particular, if $c \equiv 0$, we have

$$m = K \frac{\zeta^2}{2b}, \quad v = \frac{1}{3} (Ab + K\zeta)^2.$$

Proof. See e.g. [GW13]. □

As a consequence of Lemma 2.4.1, we have

$$\frac{1}{\kappa} \left(\mathcal{E}_{(j+1)\kappa}^{b,c,\kappa} - \mathcal{E}_{j\kappa}^{b,c,\kappa} \right) \simeq \delta_{j\kappa} + \sqrt{\kappa} \mathcal{N}_j(0, v_{j\kappa}),$$

where \mathcal{N}_j are normal distributions and

$$\begin{aligned} \delta_{j\kappa} &= m(A_{j\kappa}, K_{j\kappa}, \zeta_{j\kappa}, c(\cdot, S_{j\kappa}, j\kappa), b(S_{j\kappa}, j\kappa)) + \mathcal{L}q(S_{j\kappa}, j\kappa) \\ v_{j\kappa} &= v(A_{j\kappa}, K_{j\kappa}, \zeta_{j\kappa}, c(\cdot, S_{j\kappa}, j\kappa), b(S_{j\kappa}, j\kappa)) \end{aligned}$$

Intuitively, \mathcal{N}_j can be considered as independent. By (2.4.1), we have

$$\mathcal{E}_{\cdot}^{b,c,\kappa} \simeq \int_0^\cdot \delta_u du + \kappa \mathcal{N}\left(0, \int_0^\cdot v_u du\right),$$

which is the main content of Theorem 2.3.1.

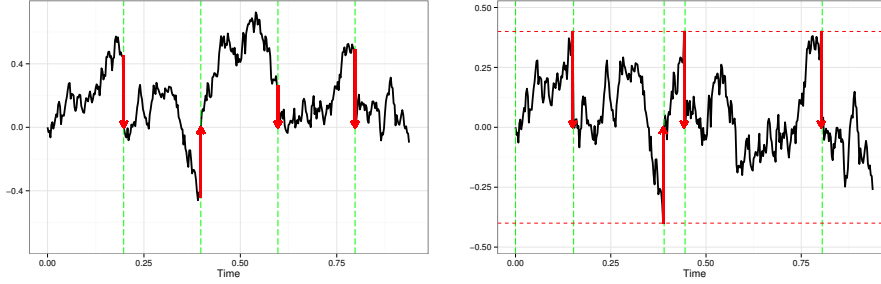


Figure 2.7 – Local models for [Lel85] (left) and [Fuk11b] (right).

2.4.2 Relation with Leland's strategy

Using a similar argument as in the previous section, we can easily identify the underlying probability models for [Lel85], [DK10] and [Fuk11b].

Lemma 2.4.2 (Local probability model for [Lel85],[DK10]). *Let $h \in \mathbb{R}^+$ and*

$$Z_t = \zeta B_t + \sum_{0 < \tau_j \leq t} \xi_j$$

with $\tau_j = jh$, $j = 1, 2, \dots$, and $\xi_j = -X_{\tau_j-}$, see Fig. 2.7. We have

$$\begin{aligned} (LLN) \quad & \frac{1}{T} \sum_{0 < \tau_j \leq T} K|\xi_j| \rightarrow m_L, \\ (CLT) \quad & \frac{1}{\sqrt{T}} \left(A \int_0^T Z_t dB_t + \sum_{0 < \tau_j \leq T} K|\xi_j| - m_L T \right) \Rightarrow \mathcal{N}(0, v_L), \end{aligned}$$

where

$$m_L = Kh^{-1/2} \zeta \sqrt{\frac{2}{\pi}}, \quad v_L = A^2 \zeta^2 \frac{h}{2} + AK \sqrt{\frac{2}{\pi}} \zeta^2 \sqrt{h} + K^2 \zeta^2 \left(1 - \frac{2}{\pi}\right). \quad (2.4.4)$$

Lemma 2.4.3 (Local probability model for [Fuk11b]). *Let $b \in \mathbb{R}^+$ and*

$$Z_t = \zeta B_t + \sum_{0 < \tau_j \leq t} \xi_j$$

with $\tau_j = \inf\{t > \tau_{j-1}, |X_t| \geq b\}$ for $j \geq 1$, $\tau_0 = 0$, and $\xi_j = -X_{\tau_j-}$, see Fig. 2.7. We have

$$\begin{aligned} (LLN) \quad & \frac{1}{T} \sum_{0 < \tau_j \leq T} K|\xi_j| \rightarrow m_F, \\ (CLT) \quad & \frac{1}{\sqrt{T}} \left(A \int_0^T Z_t dB_t + \sum_{0 < \tau_j \leq T} K|\xi_j| - m_F T \right) \Rightarrow \mathcal{N}(0, v_F), \end{aligned}$$

where

$$m_F = K \frac{\zeta^2}{b}, \quad v_F = \frac{1}{6} (Ab + 2K\zeta)^2. \quad (2.4.5)$$

Consider the case of convex payoff and $q = p^\alpha$ under Black-Scholes model $\sigma(s, t) = vs$. Taking h in (2.1.6) and b in (2.1.9) and the parameters in (2.4.2), we recover from (2.4.4) and (2.4.5) that

$$m_L = m_F = \frac{1}{\alpha} v^2 S^2 \Gamma^\alpha,$$

Since $\mathcal{L}p^\alpha = -\frac{1}{\alpha}v^2S^2\Gamma^\alpha$ (see (2.2.5)), we obtain

$$\delta_L = \delta_F = 0,$$

which means that with the choice of h in (2.1.6) and b in (2.1.9), Leland's strategy replicates asymptotically the payoff. Moreover, we recover

$$\begin{aligned} v_L &= \eta_L(\alpha)(S\Gamma^\alpha)^2(vS)^2, \\ v_F &= \eta_F(\alpha)(S\Gamma^\alpha)^2(vS)^2, \end{aligned}$$

with η_L and η_F the same as those in [DK10] and [Fuk11b].

2.4.3 Relation with Whalley and Wilmott's strategy

Under the framework of indifference pricing, [WHW94] obtain a singular control as their asymptotically optimal strategy, which is given by

$$dX_t = -\kappa^{1/3}dL_t^\kappa + \kappa^{1/3}dR_t^\kappa$$

when $\lambda(s, t) = s$ and $dS_t = vS_t dW_t$ with $v > 0$, where L^κ and R^κ are nondecreasing processes such that

$$L_t^\kappa = \int_0^t 1_{\{X_u^q - X_u = -\kappa^{1/3}b(S_u, u)\}} dL_u^\kappa, \quad R_t^\kappa = \int_0^t 1_{\{X_u^q - X_u = \kappa^{1/3}b(S_u, u)\}} dR_u^\kappa$$

and

$$|X^q - X_u| \leq \kappa^{1/3}b(S_t, t), \quad b(S_t, t) = \left| \frac{3}{2\alpha} S_t \partial_s^2 q(S_t, t) \right|^{1/3}. \quad (2.4.6)$$

The constant $\alpha > 0$ is the parameter of risk-aversion. The function q is given by (2.2.4) with $w = v$, which satisfies $\mathcal{L}q = 0$. This is the result of a delicate balancing between the utility loss due to deviation from frictionless optimal strategy and transaction costs. In particular, as $\kappa \rightarrow 0$, both the loss due to deviation and the accumulated transaction costs are of order $O(\kappa^{2/3})$. We can actually formally recover their result by our approach. Just replacing κ by $\kappa^{1/3}$ in our candidate strategy, the associated tracking error becomes

$$\begin{aligned} \hat{\mathcal{E}}_t^{b,c,\kappa} &= \Pi_t^q - \Pi_t^{b,c,\kappa} \\ &= \kappa^{1/3} \int_0^t \hat{Z}_u^\kappa dS_u + \kappa^{2/3} \int_0^t \lambda(S_u, u) c(Z_u^\kappa, S_u, u) d\langle X^q \rangle_u \\ &\quad + \kappa^{4/3} \int_0^t \lambda(S_u, u) [dL_u^\kappa + dR_u^\kappa], \end{aligned}$$

where $\hat{Z} = (X^q - X)/\kappa^{1/3}$. By a similar argument to the proof of Theorem 2.3.1 below, we can prove that

$$\hat{\mathcal{E}}_t^{b,c,\kappa} = \kappa^{1/3} \int_0^t \hat{Z}_t^\kappa dS_t + \kappa^{2/3} \int_0^t \delta^{b,c}(S_t, t) dt + O_p(\kappa)$$

and

$$\int_0^t \hat{Z}_t^\kappa dS_t \rightarrow \hat{W}_{\hat{Q}}, \quad \hat{Q} = \int_0^t \hat{\eta}^{b,c}(S_t, t) d\langle S \rangle_t,$$

where \hat{W} is an independent Brownian motion and

$$\hat{\eta}^{b,c}(s, t) = \frac{2}{a(s, t)} \int_0^{b(s, t)} x^2 g(x, s, t) dx.$$

This leads to the maximization problem of the asymptotic utility

$$\begin{aligned} & -E \left[\exp \left\{ -\alpha \left(-\kappa^{2/3} \int_0^T \delta^{b,c}(S_t, t) dt + \kappa^{1/3} \mathcal{N} \left(0, \int_0^T \hat{\eta}^{b,c}(S_t, t) d\langle S \rangle_t \right) \right) \right\} \right] \\ & = -E \left[\exp \left\{ \kappa^{2/3} \left(\alpha \int_0^T \delta^{b,c}(S_t, t) dt + \frac{\alpha^2}{2} \int_0^T \hat{\eta}^{b,c}(S_t, t) d\langle S \rangle_t \right) \right\} \right]. \end{aligned}$$

First by fixing $\delta^{b,c}$, or equivalently, fixing a , the pointwise minimization of $\hat{\eta}^{b,c}$ results in

$$b(s, t) = \frac{a(s, t)}{2}, \quad c(z, s, t) = 0, \quad \hat{\eta}^{b,c}(s, t) = \frac{a(s, t)^2}{12}.$$

Next, the pointwise minimization of

$$\delta^{b,c} + \frac{\alpha}{2} \frac{a^2}{12} \sigma^2 = \frac{\lambda}{a} |\partial_s^2 q \sigma|^2 + \frac{\alpha a^2 \sigma^2}{24}$$

results in

$$a = \left| \frac{12}{\alpha} \lambda |\partial_s^2 q|^2 \right|^{1/3},$$

which coincides with the Whalley-Wilmott strategy (2.4.6).

As the above argument shows, the rate of convergence in the Whalley-Wilmott strategy is $\kappa^{1/3}$ with transaction costs of $O(\kappa^{2/3})$. On the other hand, we attain a convergence with a faster rate κ with transaction costs of $O(1)$. The transaction costs can be compensated at the beginning by a modification of volatility. In terms of utility, it amounts to considering a scaling limit where the parameter of risk aversion tends to infinity as in [BS98].

It is easier to understand if we look at the corresponding local probability models. It is known that, in the limit of small proportional transaction costs, the first order correction of utility maximization corresponds to an ergodic control of Brownian motion. That is,

$$I = \inf_{(L,R)} \limsup_{T \rightarrow \infty} \frac{1}{T} \left(\int_0^T \Lambda Z_t^2 dt + \kappa(L_T + R_T) \right),$$

where Λ is related to the risk aversion of the investor and

$$dZ_t = \zeta dB_t + dL_t - dR_t,$$

with L_t and R_t non-decreasing processes. See Part II for more details. In contrary, Theorem 2.3.2 corresponds to the following optimization problem

$$\inf_{(b,c)} v(b, c; A, K, \zeta),$$

with (2.4.3) under the constraint that

$$m(b, c; A, K, \zeta) = \hat{m} \tag{2.4.7}$$

for a given constant \hat{m} . Here $v = v(b, c; A, K, \zeta)$ and $m = m(b, c; A, K, \zeta)$ are defined in Lemma 2.4.1. It turns out that the optimal strategy depends on the constraint (2.4.7), see Fig.2.4, which is the main content of Theorem 2.3.2.

Appendix 2.A Proof of Theorem 2.3.1

We denote in this section

$$\nu(s, t) = \partial_s^2 g(s, t) \sigma(s, t).$$

The proof of Theorem 2.3.1 utilizes a homogenization technique of two-scale stochastic differential equations and the theory of scale function and speed measure for one-dimensional ergodic diffusions. In fact the function g defined in (2.3.2) is the speed measure density for the corresponding diffusion. See [GW13] for more general results. We start with the following lemma.

Lemma 2.A.1. *Let ψ be a piecewise $C^{0,2,1}$ function on \mathcal{D}_b such that*

1. *for each (s, t) ,*

$$\int_{-b(s,t)}^{b(s,t)} \psi(x, s, t) g(x, s, t) dx = 0$$

and

2. *for each x , $\psi(x, \cdot)$ is $C^{2,1}$ and*

$$\sup\{|\partial_s^i \partial_t^j \psi(x, S_t, t)|; t \in [0, T), x \in [-b(S_t, t), b(S_t, t)]\} < \infty \quad (2.A.1)$$

almost surely for each $i \in \{0, 1, 2\}$ and $j \in \{0, 1\}$.

Then,

$$\sup_{t \in [0, T]} \left| \int_0^t \psi(Z_u^\kappa, S_u, u) \nu(S_u, u)^2 du \right| \rightarrow 0$$

in probability as $\kappa \rightarrow 0$.

Proof. Let

$$\begin{aligned} \Psi(x, s, t) &= \int_0^x \psi_1(z, s, t) dz, \\ \psi_1(z, s, t) &= \frac{2}{g(z, s, t)} \int_{-b(s,t)}^z \psi(x, s, t) g(x, s, t) dx. \end{aligned}$$

Then, Ψ is a $C^{1,2,1}$ and piecewise $C^{2,2,1}$ function and

$$\psi_1(b(s, t), s, t) = \psi_1(-b(s, t), s, t) = 0 \quad (2.A.2)$$

by the assumption. Note also that

$$-\operatorname{sgn}(z) c(z, s, t) \psi_1(z, s, t) + \frac{1}{2} \partial_z \psi_1(z, s, t) = \psi(z, s, t). \quad (2.A.3)$$

By a generalized Itô formula of [Pes07],

$$\begin{aligned} \Psi(Z_t^\kappa, S_t, t) &= \Psi(Z_0^\kappa, S_0, 0) + \int_0^t \psi_1(Z_u^\kappa, S_u, u) dZ_u^\kappa + \int_0^t \partial_s \Psi(Z_u^\kappa, S_u, u) dS_u \\ &\quad + \int_0^t \partial_t \Psi(Z_u^\kappa, S_u, u) du + \frac{1}{2} \int_0^t \partial_s^2 \Psi(Z_u^\kappa, S_u, u) d\langle S \rangle_u \\ &\quad + \int_0^t \partial_s \psi_1(Z_u^\kappa, S_u, u) d\langle Z^\kappa, S \rangle_u + \frac{1}{2} \int_0^t \partial_z \psi_1(Z_u^\kappa, S_u, u) d\langle Z^\kappa \rangle_u \\ &= \Psi(Z_0^\kappa, S_0, 0) + \frac{1}{\kappa^2} \int_0^t \psi(Z_u^\kappa, S_u, u) \nu(S_u, u)^2 du + M_t^\kappa \end{aligned}$$

for $t \in [0, T)$, where

$$\begin{aligned} M_t^\kappa &= \frac{1}{\kappa} \int_0^t \psi_1(Z_u^\kappa, S_u, u) dX_u^q + \int_0^t \partial_s \Psi(Z_u^\kappa, S_u, u) dS_u \\ &\quad + \int_0^t \partial_t \Psi(Z_u^\kappa, S_u, u) du + \frac{1}{2} \int_0^t \partial_s^2 \Psi(Z_u^\kappa, S_u, u) d\langle S \rangle_u \\ &\quad + \frac{1}{\kappa} \int_0^t \partial_s \psi_1(Z_u^\kappa, S_u, u) d\langle X^q, S \rangle_u. \end{aligned}$$

Here we have used (2.2.10), (2.A.2) and (2.A.3). Since

$$|Z_t^\kappa| \leq b(S_t, t), \quad (2.A.4)$$

the condition (2.A.1) implies that

$$\lim_{t \rightarrow T} \Psi(Z_t^\kappa, S_t, t) = \Psi(Z_0^\kappa, S_0, 0) + \frac{1}{\kappa^2} \int_0^T \psi(Z_u^\kappa, S_u, u) \nu(S_u, u)^2 du + M_T^\kappa.$$

Therefore,

$$\begin{aligned} &\sup_{t \in [0, T]} \left| \int_0^t \psi(Z_u^\kappa, S_u, u) \nu(S_u, u)^2 du \right| \\ &\leq \kappa^2 \sup_{t \in [0, T]} |\Psi(Z_t^\kappa, S_t, t) - \Psi(Z_0^\kappa, S_0, 0)| + \kappa^2 \sup_{t \in [0, T]} |M_t^\kappa| \rightarrow 0 \end{aligned}$$

in probability as $\kappa \rightarrow 0$. □

Lemma 2.A.2. *Let ψ be a function satisfying the conditions of Lemma 2.A.1. Then*

$$\sup_{t \in [0, T]} \left| \int_0^t \psi(Z_u^\kappa, S_u, u) du \right| \rightarrow 0$$

in probability as $\kappa \rightarrow \infty$.

Proof. We use Lemma 2.A.1 and (2.2.2). For $n \geq 1$, let ν_n be a $C^{3,1}$ function on $(0, \infty) \times [0, T]$ such that $\nu_n(s, t) = |\nu(s, t)|$ when $|\nu(s, t)| \geq 2/n$ and $|\nu_n(s, t)| \geq 1/n$ and $|\nu(s, t)^2 - \nu_n(s, t)^2| \leq 4n^{-2}$ for all (s, t) . Then,

$$\psi_n(x, s, t) := \frac{\psi(x, s, t)}{\nu_n(s, t)^2}$$

meets the conditions of Lemma 2.A.1 and so,

$$\sup_{t \in [0, T]} \left| \int_0^t \psi_n(Z_u^\kappa, S_u, u) \nu(S_u, u)^2 du \right| \rightarrow 0$$

in probability as $\kappa \rightarrow 0$ for each n . Therefore,

$$\begin{aligned} \sup_{t \in [0, T]} \left| \int_0^t \psi(Z_u^\kappa, S_u, u) du \right| &\leq \sup_{t \in [0, T]} \left| \int_0^t \psi_n(Z_u^\kappa, S_u, u) \nu(S_u, u)^2 du \right| \\ &\quad + 4 \left| \int_0^T \psi(Z_u^\kappa, S_u, u)^2 du \right|^{1/2} \int_0^T 1_{\{|\nu(S_u, u)| < 2n^{-1}\}} du, \end{aligned}$$

which converges to 0 as $n \rightarrow \infty$ and then, $\kappa \rightarrow 0$ by (2.2.2) and (2.A.4). □

Proof of Theorem 2.3.1. Since a stable convergence is preserved under the localization and the Girsanov-Maruyama transformation, we can and do assume $\Theta = 0$ without loss of generality. Note that h defined by (2.3.2) satisfies

$$-\operatorname{sgn}(z)c(z, s, t)h(z, s, t) + \frac{1}{2}\partial_z h(z, s, t) = c(z, s, t) - \frac{1}{a(s, t)} \quad (2.A.5)$$

and

$$h(b(s, t), s, t) = -1, \quad h(-b(s, t), s, t) = 1. \quad (2.A.6)$$

Let

$$H(x, s, t) = \lambda(s, t) \int_0^x h(z, s, t) dz.$$

Then by the generalized Itô formula, using (2.2.10), (2.A.5) and (2.A.6),

$$\begin{aligned} & H(Z_t^\kappa, S_t, t) - H(Z_0^\kappa, S_0, 0) \\ &= \int_0^t \lambda(S_u, u) h(Z_u^\kappa, S_u, u) dZ_u^\kappa + \int_0^t \partial_s H(Z_u^\kappa, S_u, u) dS_u \\ &\quad + \int_0^t \partial_t H(Z_u^\kappa, S_u, u) du + \frac{1}{2} \int_0^t \partial_s^2 H(Z_u^\kappa, S_u, u) d\langle S \rangle_u \\ &\quad + \int_0^t \partial_s(\lambda h)(Z_u^\kappa, S_u, u) d\langle Z^\kappa, S \rangle_u + \frac{1}{2} \int_0^t \lambda(S_u, u) \partial_z h(Z_u^\kappa, S_u, u) d\langle Z^\kappa \rangle_u \\ &= \frac{1}{\kappa} \int_0^t \lambda(S_u, u) h(Z_u^\kappa, S_u, u) dX_u^q + \int_0^t \lambda(S_u, u) d[L^\kappa + R^\kappa]_u \\ &\quad + \frac{1}{\kappa^2} \int_0^t \lambda(S_u, u) \left(c(Z_u^\kappa, S_u, u) - \frac{1}{a(S_u, u)} \right) \nu(S_u, u)^2 du + N_t^\kappa, \end{aligned}$$

where

$$\begin{aligned} N_t^\kappa &= \int_0^t \partial_s H(Z_u^\kappa, S_u, u) dS_u + \int_0^t \partial_t H(Z_u^\kappa, S_u, u) du \\ &\quad + \frac{1}{2} \int_0^t \partial_s^2 H(Z_u^\kappa, S_u, u) d\langle S \rangle_u + \frac{1}{\kappa} \int_0^t \partial_s(\lambda h)(Z_u^\kappa, S_u, u) d\langle X^q, S \rangle_u. \end{aligned}$$

Therefore,

$$\begin{aligned} & \kappa^{-1} \left(\mathcal{E}_t^{b, c, \kappa} - \int_0^t \delta(S_u, u) d\langle S \rangle_u \right) \\ &= \kappa(H(Z_t^\kappa, S_t, t) - H(Z_0^\kappa, S_0, 0)) - \int_0^t h(Z_u^\kappa, S_u, u) \varphi(S_u, u) du - \kappa N_t^\kappa \\ &\quad + \int_0^t (Z_u^\kappa - \lambda(S_u, u)) \partial_s^2 q(S_u, u) h(Z_u^\kappa, S_u, u) dS_u, \end{aligned}$$

where φ is a certain $C^{3,1}$ function. The first term converges to 0 uniformly on $[0, T]$ in probability due to (2.A.4). Since

$$z \mapsto h(z, s, t), \quad z \mapsto \partial_s(\lambda h)(z, s, t)$$

are odd functions and $z \mapsto g(z, s, t)$ is an even function, the second and third terms also converge to 0 in probability by Lemma 2.A.2.

It remains to show that

$$\int_0^t (Z_u^\kappa - \lambda(S_u, u)) \partial_s^2 q(S_u, u) h(Z_u^\kappa, S_u, u) dS_u$$

converges stably to $W_{Q^{b,c}}$ in law on $C[0, T]$. By [JS13, Theorem IX.7.3], it suffices to see that

$$\begin{aligned} & \int_0^t (Z_u^\kappa - \lambda(S_u, u)) \partial_s^2 q(S_u, u) h(Z_u^\kappa, S_u, u))^2 \sigma(S_u, u)^2 du \rightarrow Q_t^{b,c}, \\ & \int_0^t (Z_u^\kappa - \lambda(S_u, u)) \partial_s^2 q(S_u, u) h(Z_u^\kappa, S_u, u)) \sigma(S_u, u)^2 du \rightarrow 0 \end{aligned}$$

in probability for all $t \in [0, T]$, both of which follow from Lemma 2.A.2. \square

Appendix 2.B Proof of Theorem 2.3.2

The aim of this section is to prove Theorem 2.3.2. For notational convenience, fix for a while $(s, t) \in (0, \infty) \times [0, T)$ and write

$$\begin{aligned} a &= a(s, t), \quad b = b(s, t), \quad c(x) = c(x, s, t), \quad g(x) = g(x, s, t), \quad h(x) = h(x, s, t), \\ \gamma &= \gamma(s, t), \quad \eta^{b,c} = \eta^{b,c}(s, t). \end{aligned}$$

Lemma 2.B.1. *If c is continuous, then h' is continuous and for all $x \in (0, b)$,*

$$h(x) > -1, \quad h'(x) \geq -\frac{2}{a}, \quad c(x) = \frac{h'(x) + 2/a}{2(1 + h(x))}. \quad (2.B.1)$$

Proof. Recall that

$$h(0) = 0, \quad h(b) = -1. \quad (2.B.2)$$

By (2.A.5), we have

$$2(1 + h(x))c(x) = h'(x) + \frac{2}{a}. \quad (2.B.3)$$

Therefore h' is continuous on $[0, b]$. Further, if there exists $x \in (0, b)$ such that $h(x) \leq -1$, then $h'(x) \leq -2/a < 0$ since $c \geq 0$. As a result, $h(\hat{x}) < -1$ for all $\hat{x} > x$. This contradicts (2.B.2). Thus we obtain $h(x) > -1$. From this and (2.B.3) again, we conclude (2.B.1). \square

Lemma 2.B.2. *If c is continuous, then*

$$\eta^{b,c} = \frac{2}{a} \int_0^\infty (x(t) + \gamma(1 - x'(t)))^2 e^{-2t/a} dt,$$

where x is the solution of the ordinary differential equation

$$x' = 1 + h(x), \quad x(0) = 0. \quad (2.B.4)$$

Further, the solution x is C^2 and satisfies

$$x(\infty) = b, \quad x'(0) = 1, \quad x'(\infty) = 0, \quad x' > 0, \quad x'' \geq -\frac{2}{a}x'. \quad (2.B.5)$$

Proof. By (2.B.1),

$$\int_0^b g(x) dx = \int_0^b \frac{dx}{1 + h(x)} \exp \left\{ -\frac{2}{a} \int_0^x \frac{dy}{1 + h(y)} \right\} = \int_0^{\hat{b}} e^{-2t/a} dt, \quad (2.B.6)$$

where

$$\hat{b} = \int_0^b \frac{dx}{1 + h(x)}.$$

Since (2.B.6) is equal to $a/2$ by definition, we conclude $\hat{b} = \infty$. As a result,

$$\begin{aligned} \eta^{b,c} &= \frac{2}{a} \int_0^b \frac{(x - \gamma h(x))^2}{1 + h(x)} \exp \left\{ -\frac{2}{a} \int_0^x \frac{dy}{1 + h(y)} \right\} dx \\ &= \frac{2}{a} \int_0^\infty (x(t) + \gamma(1 - x'(t)))^2 e^{-2t/a} dt, \end{aligned}$$

where $t \mapsto x(t)$ is the inverse function of

$$x \mapsto t(x) = \int_0^x \frac{dy}{1 + h(y)}.$$

The rest follows from (2.B.1) and (2.B.2). \square

Lemma 2.B.3. *If c is continuous, then there corresponds an increasing convex C^2 function y on $[0, 1]$ such that*

$$y(0) = 0, \quad y'(0) = \frac{a}{2} \quad (2.B.7)$$

and

$$\eta^{b,c} = \eta_a[y] := \int_0^1 \left(y(u) + \gamma + \frac{2\gamma}{a}(u-1)y'(u) \right)^2 du.$$

Proof. Let x be the solution of (2.B.4) and

$$y(u) = x \left(-\frac{a}{2} \log(1-u) \right) = x(v^{-1}(u)), \quad v(t) = 1 - e^{-2t/a}.$$

Since $x(t) = y(v(t))$, we have

$$x'(t) = \frac{2}{a} y'(v(t))(1-v(t)), \quad x''(t) = \frac{4}{a^2} y''(v(t))(1-v(t))^2 - \frac{2}{a} x'(t).$$

From (2.B.5), we obtain $y'' \geq 0$ and $y' > 0$. Changing variable $u = v(t)$, the result follows from Lemma 2.B.2. \square

The next lemma is a key and the most difficult part to prove.

Lemma 2.B.4. *Denote by \mathcal{Y}_a the set of increasing convex functions y on $[0, 1]$ with (2.B.7). Then*

$$\inf_{y \in \mathcal{Y}_a} \eta_a[y] = \lim_{x \rightarrow \gamma} x^2 \eta_{\dagger}(a/x).$$

Proof. It is easy to see that when $\gamma = 0$, the minimum of η_a is attained by $y(u) = au/2$ and so,

$$\inf_{y \in \mathcal{Y}_a} \eta_a[y] = \frac{a^2}{12} = \lim_{x \rightarrow \gamma} x^2 \eta_{\dagger}(a/x).$$

Now, suppose $\gamma \neq 0$. Then we have five cases: (1) $-2 < a/\gamma < 1$, (2) $a = \gamma$, (3) $1 < a/\gamma < 2$, (4) $a/\gamma \geq 2$ and (5) $a/\gamma \leq -2$.

Case 1). Assume $-2 < a/\gamma < 1$. For $\epsilon \in (0, 1)$, define y_ϵ as

$$y_\epsilon(u) = \begin{cases} \gamma(1-u)^{-a/2\gamma} - \gamma & \text{if } 0 \leq u \leq 1 - \epsilon \\ y_{\epsilon,0} + y'_{\epsilon,0}(u - u_{\epsilon,0}) & \text{if } 1 - \epsilon < u \leq 1, \end{cases}$$

where

$$u_{\epsilon,0} = 1 - \epsilon, \quad y_{\epsilon,0} = \gamma(1 - u_{\epsilon,0})^{-a/2\gamma} - \gamma, \quad y'_{\epsilon,0} = \frac{a}{2}(1 - u_{\epsilon,0})^{-a/2\gamma - 1}.$$

Then, $y_\epsilon \in \mathcal{Y}_a$ for all ϵ . Note that

$$y_\epsilon(u) + \gamma + \frac{2\gamma}{a}(u-1)y'_\epsilon(u) = \gamma(1-u)^{-a/2\gamma} + \frac{2\gamma}{a}(u-1)\frac{a}{2}(1-u)^{-a/2\gamma-1} = 0$$

for $u \in (0, 1 - \epsilon)$. Therefore,

$$\eta_a[y_\epsilon] = \eta(u_{\epsilon,0}, y_{\epsilon,0}, y'_{\epsilon,0}, 2\gamma/a),$$

where

$$\eta(v, w, z, \theta) = \int_v^1 (w + z(u-v) + \gamma + \theta(u-1)z)^2 du.$$

By a straightforward calculation,

$$\begin{aligned} \eta(v, w, z, \theta) &= \frac{\theta^2 - \theta + 1}{3} (1-v)^3 \left(z + \frac{3}{2} \frac{1-\theta}{\theta^2 - \theta + 1} \frac{\gamma + w}{1-v} \right)^2 \\ &\quad + \frac{(\theta + 1)^2}{4(\theta^2 - \theta + 1)} (\gamma + w)^2 (1-v). \end{aligned} \quad (2.B.8)$$

It follows that

$$\eta(u_{\epsilon,0}, y_{\epsilon,0}, y'_{\epsilon,0}, 2\gamma/a) = O(\epsilon^{1-a/\gamma}) \rightarrow 0 \quad (2.B.9)$$

as $\epsilon \rightarrow 0$, which means that

$$\inf_{y \in \mathcal{Y}_a} \eta_a[y] = 0 = |\gamma|^2 \eta_{\dagger}(a/\gamma).$$

Note that if $a/\gamma < -2$ then y_{ϵ} is not convex and if $a/\gamma \geq 1$ then (2.B.9) doesn't hold.

Case 2). Assume $a = \gamma$. For $\epsilon \in (0, 1)$, define y_{ϵ} as

$$y_{\epsilon}(u) = \begin{cases} \frac{\gamma}{1-\epsilon} \left\{ (1-u)^{-(1-\epsilon)/2} - 1 \right\} & \text{if } 0 \leq u \leq 1 - e^{-1/\epsilon^2} \\ y_{\epsilon,0} + y'_{\epsilon,0}(u - u_{\epsilon,0}) & \text{if } 1 - e^{-1/\epsilon^2} < u \leq 1, \end{cases}$$

where

$$u_{\epsilon,0} = 1 - e^{-1/\epsilon^2}, \quad y_{\epsilon,0} = \frac{\gamma}{1-\epsilon} \left\{ (1 - u_{\epsilon,0})^{-(1-\epsilon)/2} - 1 \right\}, \quad y'_{\epsilon,0} = \frac{\gamma}{2} (1 - u_{\epsilon,0})^{-(1-\epsilon)/2-1}.$$

Then, $y_{\epsilon} \in \mathcal{Y}_a$ for all ϵ . Note that

$$\gamma + y_{\epsilon}(u) + 2(u-1)y'_{\epsilon}(u) = -\frac{\gamma\epsilon}{1-\epsilon} + \frac{\gamma\epsilon}{1-\epsilon} (1-u)^{-(1-\epsilon)/2}$$

for $u \in (0, u_{\epsilon,0})$. Therefore,

$$\eta_a[y_{\epsilon}] = \frac{\gamma^2 \epsilon^2}{(1-\epsilon)^2} \int_0^{u_{\epsilon,0}} \left\{ (1-u)^{-(1-\epsilon)/2} - 1 \right\}^2 du + \eta(u_{\epsilon,0}, y_{\epsilon,0}, y'_{\epsilon,0}, 2),$$

which converges to $0 = \eta_{\dagger}(1)$ as $\epsilon \rightarrow 0$ by (2.B.8).

Case 3). Assume $1 < a/\gamma < 2$. For $y \in \mathcal{Y}_a$, let

$$\varphi(u) = (1-u)^{a/2\gamma} (y(u) + \gamma).$$

Then, $\varphi(0) = \gamma$, $\varphi(1) = 0$ and

$$y + \gamma + \frac{2\gamma}{a} (u-1)y'(u) = -\frac{2\gamma}{a} (1-u)^{1-a/2\gamma} \varphi'(u).$$

By the Cauchy-Schwarz inequality, for $u_0 \in [0, 1]$,

$$\varphi(u_0)^2 = \left| \int_{u_0}^1 \varphi'(u) du \right|^2 \leq \int_{u_0}^1 (1-u)^{2-a/\gamma} |\varphi'(u)|^2 du \int_{u_0}^1 (1-u)^{a/\gamma-2} du. \quad (2.B.10)$$

Therefore,

$$\begin{aligned} \int_{u_0}^1 \left(y + \gamma + \frac{2\gamma}{a} (u-1)y'(u) \right)^2 du &= \frac{4\gamma^2}{a^2} \int_{u_0}^1 (1-u)^{2-a/\gamma} |\varphi'(u)|^2 du \\ &\geq \frac{4\gamma^2}{a^2} \left(\frac{a}{\gamma} - 1 \right) (1-u_0)^{1-a/\gamma} \varphi(u_0)^2. \end{aligned} \quad (2.B.11)$$

Under the constraint $\varphi(1) = 0$, (2.B.11) attains equality if and only if there exists $r \in \mathbb{R}$ such that

$$\varphi(u) = r(1-u)^{a/\gamma-1},$$

which corresponds to

$$y(u) = r(1-u)^{a/2\gamma-1} - \gamma. \quad (2.B.12)$$

For $\epsilon \in (0, 1-u_0)$, define y_ϵ as

$$y_\epsilon(u) = \begin{cases} au/2 & \text{if } 0 \leq u \leq u_0 \\ r(1-u)^{a/2\gamma-1} - \gamma & \text{if } u_0 < u \leq u_{\epsilon,0} \\ y_{\epsilon,0} + y'_{\epsilon,0}(u - u_{\epsilon,0}) & \text{if } u_{\epsilon,0} < u \leq 1, \end{cases} \quad (2.B.13)$$

where (u_0, r) is the solution of

$$\frac{a}{2}u_0 = r(1-u_0)^{a/2\gamma-1} - \gamma, \quad \frac{a}{2} = r \left(1 - \frac{a}{2\gamma}\right) (1-u_0)^{a/2\gamma-2} \quad (2.B.14)$$

and

$$u_{\epsilon,0} = 1 - \epsilon, \quad y_{\epsilon,0} = r(1-u_{\epsilon,0})^{a/2\gamma-1} - \gamma, \quad y'_{\epsilon,0} = r \left(1 - \frac{a}{2\gamma}\right) (1-u_{\epsilon,0})^{a/2\gamma-2}.$$

The solution of (2.B.14) uniquely exists and $u_0 \in (0, 1)$ and $r > 0$; in fact

$$u_0 = \frac{4\gamma(a-\gamma)}{a(4\gamma-a)} > 0, \quad 1-u_0 = \frac{4\gamma^2-a^2}{a(4\gamma-a)} > 0, \quad r = \frac{a\gamma}{2\gamma-a}(1-u_0)^{2-a/2\gamma} > 0.$$

Then, $y_\epsilon \in \mathcal{Y}_a$ for all ϵ . By a straightforward calculation,

$$\begin{aligned} \eta_a[y_\epsilon] &= \left(\gamma + \frac{a}{2}\right)^2 \frac{u_0^3}{3} + \frac{4\gamma}{a^2}(a-\gamma)r^2 \left((1-u_0)^{a/\gamma-1} - (1-u_{\epsilon,0})^{a/\gamma-1}\right) \\ &\quad + \eta(u_{\epsilon,0}, y_{\epsilon,0}, y'_{\epsilon,0}, 2\gamma/a) \\ &\rightarrow \left(\gamma + \frac{a}{2}\right)^2 \frac{u_0^3}{3} + \frac{4\gamma}{a^2}(a-\gamma)r^2(1-u_0)^{a/\gamma-1} \\ &= \gamma^2 \eta_\dagger(a/\gamma) \end{aligned}$$

as $\epsilon \rightarrow 0$. Here we have used (2.B.8) to observe $\eta(u_{\epsilon,0}, y_{\epsilon,0}, y'_{\epsilon,0}, 2\gamma/a) = O(\epsilon^{a/\gamma-1})$. Now, let us show this is the infimum. Suppose there exists $y \in \mathcal{Y}_a$ such that $\eta_a[y] < \gamma^2 \eta_\dagger(a/\gamma)$. Since a convex function is approximated by piecewise linear convex functions arbitrarily close, we can and do assume y itself is piecewise linear without loss of generality. Let

$$0 < u_1 < u_2 < \dots < u_n < 1$$

be the discontinuity points of y' . Let $u_0 = 0$ and $u_{n+1} = 1$. Denote $y_i = y(u_i)$, $y'_{i-} = \lim_{u \uparrow u_i} y'(u)$ and $y'_{i+} = \lim_{u \downarrow u_i} y'(u)$. Note that

$$\frac{a}{2} = y'_{0+} = y'_{1-} < y'_{1+} = y'_{2-} < \dots < y'_{n+}$$

and $y'_{i-}u_i \geq y_i$ for each i by convexity. Let (v_i, r_i) be the solution of

$$y_i + y'_{i+}(v_i - u_i) = r_i(1-v_i)^{a/2\gamma-1} - \gamma, \quad y'_{i+} = r_i \left(1 - \frac{a}{2\gamma}\right) (1-v_i)^{a/2\gamma-2}.$$

The solution uniquely exists and $v_i \in (0, 1)$ and $r_i > 0$; in fact

$$\begin{aligned} v_i &= \frac{\gamma(2y'_{i+} + a - 2\gamma) + (2\gamma - a)(y'_{i+}u_i - y_i)}{y'_{i+}(4\gamma - a)} > 0, \\ 1 - v_i &= \frac{(2\gamma - a)(\gamma + y_i + (1 - u_i)y'_{i+})}{y'_{i+}(4\gamma - a)} > 0. \end{aligned} \quad (2.B.15)$$

Further, it holds $v_i < v_{i+1}$ because

$$\begin{aligned}
v_{i+1} - v_i &= \frac{2\gamma + (2\gamma - a)u_{i+1}}{4\gamma - a} - \frac{(2\gamma - a)(\gamma + y_{i+1})}{y'_{i+1+}(4\gamma - a)} \\
&\quad - \frac{2\gamma + (2\gamma - a)u_i}{4\gamma - a} + \frac{(2\gamma - a)(\gamma + y_i)}{y'_{i+}(4\gamma - a)} \\
&> \frac{2\gamma + (2\gamma - a)u_{i+1}}{4\gamma - a} - \frac{(2\gamma - a)(\gamma + y_{i+1})}{y'_{i+}(4\gamma - a)} \\
&\quad - \frac{2\gamma + (2\gamma - a)u_i}{4\gamma - a} + \frac{(2\gamma - a)(\gamma + y_i)}{y'_{i+}(4\gamma - a)} \\
&= \frac{(2\gamma - a)(u_{i+1} - u_i)}{4\gamma - a} - \frac{(2\gamma - a)(y_{i+1} - y_i)}{y'_{i+}(4\gamma - a)} = 0.
\end{aligned}$$

In particular if $v_{i+1} \leq u_{i+1}$, then $v_i < u_{i+1}$. This implies that the set

$$\mathcal{I} := \{i \in \{0, 1, \dots, n\}; u_i \leq v_i < u_{i+1}\}$$

is not empty. In fact, if $v_n < u_n$, then $v_{n-1} < u_n$. If $v_{n-1} < u_{n-1}$, then $v_{n-2} < u_{n-1}$. If \mathcal{I} is empty, then we conclude $v_0 < u_0 = 0$ by induction, which contradicts (2.B.15). Now, let

$$k = \min\{i \in \{0, 1, 2, \dots, n\}; u_i \leq v_i < u_{i+1}\}.$$

Then $v_{i-1} \geq u_i$ for all $i \leq k$. For $z \geq y'_{i-}$, define $y_i(\cdot, z)$ by

$$y_i(u, z) = \begin{cases} y(u) & \text{if } 0 \leq u \leq u_i \\ y_i + z(u - u_i) & \text{if } u_i < u \leq v_i(z) \\ r_i(z)(1 - u)^{a/2\gamma-1} - \gamma & \text{if } v_i(z) < u \leq 1, \end{cases} \quad (2.B.16)$$

where $(v_i(z), r_i(z))$ is the solution of

$$y_i + z(v_i(z) - u_i) = r_i(z)(1 - v_i(z))^{a/2\gamma-1} - \gamma, \quad z = r_i(z) \left(1 - \frac{a}{2\gamma}\right) (1 - v_i(z))^{a/2\gamma-2}.$$

The solution uniquely exists as before and we have

$$v_i(y'_{i+}) = v_i, \quad r_i(y'_{i+}) = r_i$$

and

$$v_i(y'_{i-}) = v_{i-1}, \quad r_i(y'_{i-}) = r_{i-1}.$$

Since

$$v_i(z) = \frac{2\gamma + (2\gamma - a)u_i}{4\gamma - a} - \frac{(2\gamma - a)(\gamma + y_i)}{z(4\gamma - a)},$$

$v_i(z)$ is an increasing function. This implies that $v_i(z) \geq v_{i-1} \geq u_i$ for all $i \leq k$. Note also that $y_k(u, y'_{k+}) = y(u)$ for $u \in [0, v_k]$. Recall that a function of the form (2.B.12) for $u \geq v_k$ minimizes (2.B.11) with $u_0 = v_k$. This implies that

$$\eta_a[y_k(\cdot, y'_{k+})] < \eta_a[y].$$

Further,

$$\begin{aligned}
&\partial_z \{\eta_a[y_i(\cdot, z)]\} \\
&= \frac{(4\gamma z(1 - u_i) + (2\gamma - a)(\gamma + y_i))(2\gamma z(1 - u_i) - (2\gamma - a)(\gamma + y_i))^2}{3z^2 a^2 (4\gamma - a)} \\
&> 0
\end{aligned} \quad (2.B.17)$$

for all $z \geq y'_{i-}$ (see Appendix). It means that

$$\eta_a[y_k(\cdot, y'_{k-})] < \eta_a[y_k(\cdot, y'_{k+})] < \eta_a[y].$$

The function $y_k(\cdot, y'_{k-})$ is continuously differentiable at u_k and coincides with $y_{k-1}(\cdot, y'_{k-1+})$. Again by (2.B.17), we have

$$\eta_a[y_{k-1}(\cdot, y'_{k-1-})] < \eta_a[y_{k-1}(\cdot, y'_{k-1+})] < \eta_a[y].$$

We can repeat this argument to conclude

$$\eta_a[y_1(\cdot, y'_{1-})] < \eta_a[y].$$

Note that $y'_{1-} = a/2$ and so, $y_1(\cdot, y'_{1-})$ coincides with the limit of y_ϵ defined by (2.B.13) as $\epsilon \rightarrow 0$. It is not difficult see $\eta_a[y_1(\cdot, y'_{1-})] = \gamma^2 \eta_{\dagger}(a/\gamma)$, which contradicts how y was chosen. This completes Case 3.

Case 4). Assume $a/\gamma \geq 2$. Let $\hat{y}(u) = au/2$. Then, $\hat{y} \in \mathcal{Y}_a$ and $\eta_a[\hat{y}] = \gamma^2 \eta_{\dagger}(a/\gamma)$. Suppose there exists $y \in \mathcal{Y}_a$ such that $\eta_a[y] < \gamma^2 \eta_{\dagger}(a/\gamma)$. Since a convex function is approximated by piecewise linear convex functions arbitrarily close, we can and do assume y itself is piecewise linear without loss of generality. Let $u_0 \in (0, 1)$ be the last point where y' jumps. Denote $y_0 = y(u_0)$, $y'_- = \lim_{u \uparrow u_0} y'(u)$ and $y'_+ = \lim_{u \downarrow u_0} y'(u)$. For $z \geq y'_-$, define $y(\cdot, z)$ as

$$y(u, z) = \begin{cases} y(u) & \text{if } 0 \leq u \leq u_0 \\ y_0 + z(u - u_0) & \text{if } u_0 < u \leq 1. \end{cases}$$

Note that $y(\cdot, y'_+) = y$. As seen before,

$$\int_{u_0}^1 \left(y(u, z) + \gamma + \frac{2\gamma}{a}(u - 1)y'(u, z) \right)^2 du = \eta(u_0, y_0, z, 2\gamma/a)$$

and η is given by (2.B.8). Since $1 - \theta = 1 - 2\gamma/a \geq 0$, the first term of (2.B.8) is minimized by $z = y'_-$ on the region $z \geq y'_-$. This implies

$$\eta_a[y(\cdot, y'_-)] < \eta_a[y(\cdot, y'_+)] = \eta_a[y].$$

The function $y(\cdot, y'_-)$ is continuously differentiable at u_0 . We can repeat the same argument with y replaced by $y(\cdot, y'_-)$ to obtain a smaller value of η_a . Eventually, all discontinuity points are removed and the final product coincides with $\hat{y}(u) = au/2$. This contradicts how y was chosen. This completes Case 4.

Case 5). Assume $a/\gamma \leq -2$. The idea is the same as in the previous case. Let $\hat{y}(u) = au/2$. As before, $\hat{y} \in \mathcal{Y}_a$ and $\eta_a[\hat{y}] = \gamma^2 \eta_{\dagger}(a/\gamma)$. Suppose there exists $y \in \mathcal{Y}_a$ such that $\eta_a[y] < \gamma^2 \eta_{\dagger}(a/\gamma)$. Since a convex function is approximated by piecewise linear convex functions arbitrarily close, we can and do assume y itself is piecewise linear without loss of generality. Let $u_0 \in (0, 1)$ be the last point where y' jumps. Denote $y_0 = y(u_0)$, $y'_- = \lim_{u \uparrow u_0} y'(u)$ and $y'_+ = \lim_{u \downarrow u_0} y'(u)$. For $z \geq y'_-$, define $y(\cdot, z)$ as

$$y(u, z) = \begin{cases} y(u) & \text{if } 0 \leq u \leq u_0 \\ y_0 + z(u - u_0) & \text{if } u_0 < u \leq 1. \end{cases}$$

As seen before, $y(\cdot, y'_+) = y$ and

$$\int_{u_0}^1 \left(y(u, z) + \gamma + \frac{2\gamma}{a}(u - 1)y'(u, z) \right)^2 du = \eta(u_0, y_0, z, 2\gamma/a)$$

with η given by (2.B.8). Note that $y'_- \geq a/2$ by the convexity of y . Therefore, the same argument as in the previous case remains valid here once we prove that

$$-\frac{3}{2} \frac{1-\theta}{\theta^2-\theta+1} \frac{\gamma+w}{1-v} \leq \frac{a}{2} \quad (2.B.18)$$

for $\theta = 2\gamma/a$, $v = u_0$ and $w = y_0$. By the convexity, we have $y_0 \geq u_0 a/2$. Note also that

$$\partial_v \left\{ \frac{\gamma + va/2}{1-v} \right\} = \frac{a + 2\gamma}{2(1-v)^2} \geq 0.$$

Therefore,

$$\begin{aligned} -\frac{3}{2} \frac{1-\theta}{\theta^2-\theta+1} \frac{\gamma+y_0}{1-u_0} &\leq -\frac{3}{2} \frac{1-\theta}{\theta^2-\theta+1} \frac{\gamma+u_0 a/2}{1-u_0} \\ &\leq -\frac{3}{2} \frac{1-\theta}{\theta^2-\theta+1} \gamma. \end{aligned}$$

To show (2.B.18), it suffices then to see

$$\frac{a}{2} + \frac{3}{2} \frac{1-\theta}{\theta^2-\theta+1} \gamma = \frac{a(2-\theta)(1+\theta)}{4(\theta^2-\theta+1)} \geq 0.$$

Now, we can deduce a contradiction as in the previous case, which completes the proof. \square

Proof of Theorem 2.3.2. In order to bound the infimum of $\eta^{b,c}$, it suffices to consider c which is continuous. Then by the preceding lemmas, we conclude

$$\eta^{b,c}(s, t) \geq \lim_{x \rightarrow \gamma(s,t)} x^2 \eta_{\dagger}(a(s, t)/x)$$

for any $(b, c) \in \mathcal{S}_a$. It remains to show that this lower bound is asymptotically attained by a sequence $(b_n, c_n) \in \mathcal{S}_a$. By a localization argument, we can and do assume a , $1/a$ and γ are bounded without loss of generality. It suffices then to construct $(b_n, c_n) \in \mathcal{S}_a$ in such a way that $\eta^{b_n, c_n}(s, t)$ is bounded in (n, s, t) and

$$\eta^{b_n, c_n}(s, t) \rightarrow \lim_{x \rightarrow \gamma(s,t)} x^2 \eta_{\dagger}(a(s, t)/x)$$

for all (s, t) . We often omit the dependence in (s, t) in the following. Define $l = l(s, t)$ by (2.3.12). Note that l can be expressed as

$$l(s, t) = a(s, t) \hat{l} \left(\frac{\gamma(s, t)}{a(s, t)} \right),$$

where

$$\hat{l}(x) = \begin{cases} 0 & \text{if } x \geq 1 \\ 2x(1-x)/(4x-1) & \text{if } 1/2 > x \geq 1/2 \\ x & \text{if } 1/2 > x \geq 0 \\ 0 & \text{if } 0 > x. \end{cases}$$

Note that \hat{l} is continuous and $0 \leq \hat{l} \leq 1/2$. Let $l_n = a \hat{l}_n(\gamma/a)$, where \hat{l}_n is a sequence of C^4 functions such that

$$\max \left\{ \hat{l}(x) - \frac{1}{n}, \frac{4n}{5n-2} \hat{l}(x) \right\} \leq \hat{l}_n(x) < \min \left\{ \hat{l}(x) + \frac{1}{n}, \frac{1}{2} \right\}$$

for all x and $n \geq 3$ and that $\hat{l}_n(x) = 0$ when $x \leq -1/2$. Let $r_n = a \hat{r}_n(\gamma/a)$, where

$$\hat{r}_n(x) = \hat{l}_n(x) + \exp\{n^2\}(x - 1/2)_+^5 + |x|(1 - \exp\{-n(x + 1/2)_-\}^5).$$

Note that \hat{r}_n is a C^4 function, $\hat{r}_n(x) = \hat{l}_n(x)$ when $|x| \leq 1/2$, and that the pointwise limit of \hat{r}_n is

$$\hat{r}_\infty(x) = \begin{cases} \infty & \text{if } x > 1/2 \\ \hat{l}(x) & \text{if } -1/2 \leq x \leq 1/2 \\ |x| & \text{if } x < -1/2. \end{cases}$$

Let $\psi_n = 1 - 2/n$ and

$$\hat{\beta}_n(x) = \begin{cases} \psi_n(1 - 2\hat{l}_n(x))/(x + \psi_n\hat{l}_n(x)), & \text{if } |x| \geq 1/2, \\ (x + \frac{1}{2})\hat{\beta}_n(-\frac{1}{2}) + (x - \frac{1}{2})\hat{\beta}_n(\frac{1}{2}) & \text{if } |x| < 1/2. \end{cases}$$

Define $h_n^*(y) = h_n^*(y, s, t)$ for $y \geq 0$ by

$$h_n^*(y, s, t) = \hat{h}_n^*\left(\frac{y}{a(s, t)}, \frac{\gamma(s, t)}{a(s, t)}\right),$$

where

$$\hat{h}_n^*(z, x) = \begin{cases} -2z & \text{if } 0 \leq z < \hat{l}_n(x) \\ \hat{\beta}_n(x)(z - \hat{l}_n(x)) - 2\hat{l}_n(x) & \text{if } \hat{l}_n(x) \leq z < \hat{r}_n(x) \\ -2(z - \hat{r}_n(x)) + \hat{\beta}_n(x)(\hat{r}_n(x) - \hat{l}_n(x)) - 2\hat{l}_n(x) & \text{if } z \geq \hat{r}_n(x). \end{cases}$$

Note that \hat{h}_n^* is continuous in (z, x) and therefore, so is h_n^* in (y, s, t) . Let

$$b_n^* = \inf\{y > 0; h_n^*(y) = -1\}$$

and

$$g_n^*(x) = \frac{1}{1 + h_n^*(x)} \exp\left\{-\frac{2}{a} \int_0^x \frac{dy}{1 + h_n^*(y)}\right\},$$

$$\eta_n^* = \frac{2}{a} \int_0^{b_n^*} (x - \gamma h_n^*(x))^2 g_n^*(x) dx.$$

It is straightforward to see

$$b_n^* = a\hat{b}_n(\gamma/a) = r_n + \frac{\gamma + \psi_n r_n}{\gamma + \psi_n l_n} \left(\frac{a}{2} - l_n\right),$$

where

$$\hat{b}_n(x) = \hat{r}_n(x) + \frac{x + \psi_n \hat{r}_n(x)}{x + \psi_n \hat{l}_n(x)} \left(\frac{1}{2} - \hat{l}_n(x)\right).$$

This is a sequence of $C^{3,1}$ bounded functions, since

$$\frac{\gamma + \psi_n r_n}{\gamma + \psi_n l_n} = 1$$

if $|\gamma| \leq a/2$ and we have

$$|\gamma + l_n \psi_n| \geq \frac{a}{2} \tag{2.B.19}$$

otherwise. We always have $l_n \leq r_n < b_n^*$.

Define $\epsilon_n^*(y)$ for $y \geq 0$ by $\epsilon_n^*(y) = \hat{\epsilon}_n(y/a, \gamma/a)$, where

$$\hat{\epsilon}_n(z, x) = \begin{cases} 0 & \text{if } 0 \leq z \leq \hat{l}_n(x), \\ 1/n & \text{if } \hat{l}_n(x) < z < (\hat{r}_n(x) + \hat{b}_n(x))/2, \\ 0 & \text{if } (\hat{r}_n(x) + \hat{b}_n(x))/2 \leq z. \end{cases}$$

Take a sequence of $C^{5,4}$ functions \hat{h}_n such that

$$1 - \hat{\epsilon}_n(z, x) \leq \frac{1 + \hat{h}_n(z, x)}{1 + \hat{h}_n^*(z, x)} \leq 1 + \hat{\epsilon}_n(z, x), \quad \partial_z \hat{h}_n(z, x) \geq -2.$$

Such a sequence exists because $\hat{\beta} \geq -2$ and

$$\inf_x \left\{ \frac{\hat{r}_n(x) + \hat{b}_n(x)}{2} - \hat{l}_n(x) \right\} > 0$$

for each n . Define $h_n(y) = h_n(y, s, t)$ for $y \geq 0$ by $h_n(y) = \hat{h}_n(y/a, \gamma/a)$. Then, h_n is a sequence of $C^{1,3,1}$ functions such that

$$1 - \epsilon_n^*(y) \leq \frac{1 + h_n(y, s, t)}{1 + h_n^*(y, s, t)} \leq 1 + \epsilon_n^*(y), \quad \partial_y h_n(y, s, t) \geq -2/a$$

for all (y, s, t) . Define b_n , g_n and η_n as in same manner as b_n^* , g_n^* and η_n^* with h_n^* replaced by h_n . It is easy to see $b_n = b_n^*$. Since $h_n^*(x)$ is piecewise linear in x , it is also easy to see

$$\int_0^{b_n^*} \frac{dy}{1 + h_n^*(y)} = \infty.$$

Therefore,

$$\int_0^{b_n} \frac{dy}{1 + h_n(y)} = \infty$$

and so,

$$\int_0^{b_n} g_n(x) dx = \int_0^\infty e^{-2t/a} dt = \frac{a}{2}.$$

This means $(b_n, c_n) \in \mathcal{S}_a$, where c_n is defined by

$$c_n(x) = \frac{h_n'(x) + 2/a}{2(1 + h_n(x))}.$$

Now, we show that (b_n, c_n) is an asymptotically optimal sequence. When $|\gamma| \leq a/2$, we have $r_n = l_n$, $b_n = a/2$ and so, $h_n(x) = h_n^*(x) = -2x/a$. By a straightforward computation, we have

$$\eta_n = \eta_n^* = \frac{1}{3} \left(\gamma + \frac{a}{2} \right)^2 = \gamma^2 \eta_n^*(a/\gamma).$$

We therefore assume $|\gamma| > a/2$ in the sequel. Define

$$\epsilon_n(y) = \begin{cases} 0 & \text{if } y \leq l_n \\ 1/n & \text{if } y > l_n. \end{cases}$$

Note that $\epsilon_n(y) \geq \epsilon_n^*(y)$ and so,

$$\eta_n \leq \frac{2}{a} \int_0^{b_n^*} (x - \gamma h_n(x))^2 \frac{1 + \epsilon_n(x)}{1 + h_n^*(x)} \exp \left\{ -\frac{2}{a} \int_0^x \frac{1 - \epsilon_n(y)}{1 + h_n^*(y)} dy \right\} dx.$$

Since the square root of the right hand side is less than or equal

$$\left\{ \frac{2}{a} \int_0^{b_n^*} (x - \gamma h_n^*(x))^2 \frac{1 + \epsilon_n(x)}{1 + h_n^*(x)} \exp \left\{ -\frac{2}{a} \int_0^x \frac{1 - \epsilon_n(y)}{1 + h_n^*(y)} dy \right\} dx \right\}^{1/2} \\ + \frac{\gamma}{n} \left\{ \frac{2}{a} \int_0^{b_n^*} (1 + h_n^*(x))^2 \frac{1 + \epsilon_n(x)}{1 + h_n^*(x)} \exp \left\{ -\frac{2}{a} \int_0^x \frac{1 - \epsilon_n(y)}{1 + h_n^*(y)} dy \right\} dx \right\}^{1/2},$$

it suffices to show that

$$\frac{2}{a} \int_0^{b_n^*} (x - \gamma h_n^*(x))^2 \frac{1 + \epsilon_n(x)}{1 + h_n^*(x)} \exp \left\{ -\frac{2}{a} \int_0^x \frac{1 - \epsilon_n(y)}{1 + h_n^*(y)} dy \right\} dx \quad (2.B.20)$$

is bounded in (n, s, t) and $\rightarrow \gamma^2 \eta_{\dagger}(a/\gamma)$ as $n \rightarrow \infty$ for all (s, t) , and that

$$\frac{1}{n} \frac{2}{a} \int_0^{b_n^*} (1 + h_n^*(x))(1 + \epsilon_n(x)) \exp \left\{ -\frac{2}{a} \int_0^x \frac{1 - \epsilon_n(y)}{1 + h_n^*(y)} dy \right\} dx \quad (2.B.21)$$

is bounded in (n, s, t) . Let

$$t_n(x) = \int_0^x \frac{1 - \epsilon_n(y)}{1 + h_n^*(y)} dy.$$

By a straightforward computation,

$$t_n(x) = \begin{cases} -\frac{a}{2} \log(1 - 2x/a) & \text{if } x < l_n, \\ -\frac{a}{2} \log(1 - 2l_n/a) + \frac{1}{\beta_n} \left(1 - \frac{1}{n}\right) \log \left(\frac{1 - 2l_n/a + \beta_n(x - l_n)}{1 - 2l_n/a} \right) & \text{if } l_n \leq x \leq r_n, \end{cases}$$

where $\beta_n = \hat{\beta}_n(\gamma/a)/a$. Note that $l_n < a/2$ by definition. Put

$$u_n = 1 - \exp\{-2t_n(r_n)/a\}$$

and consider changing variable as

$$-\frac{a}{2} \log(1 - u) = t_n(x),$$

that is,

$$x = \begin{cases} au/2 & \text{if } u < 2l_n/a, \\ \left(l_n + \frac{\gamma}{\psi_n}\right) \left(\frac{1-u}{1-2l_n/a}\right)^{\delta_n} - \frac{\gamma}{\psi_n} & \text{if } 2l_n/a \leq u < u_n, \end{cases}$$

where

$$\delta_n = -\frac{a}{2} \frac{n}{n-1} \beta_n = -\frac{1}{2} \frac{n-2}{n-1} \frac{a-2l_n}{\gamma + l_n \psi_n}.$$

For $u \in (u_n, 1]$, $x = x(u)$ has the form of linear extrapolation because $t'_n(x)$ is continuous at $x = r_n$ and the slope of $h_n^*(x)$ is $-2/a$ for $x > r_n$. Then, we have that (2.B.20) is equal to

$$\int_0^{2l_n/a} \left(x(u) + \gamma + \frac{2\gamma}{a}(u-1)x'(u) \right)^2 du \\ + \frac{n+1}{n-1} \int_{2l_n/a}^1 \left(x(u) + \gamma + \frac{2\gamma}{a} \left(1 - \frac{1}{n}\right) (u-1)x'(u) \right)^2 du.$$

The first term is equal to

$$\left(\gamma + \frac{a}{2} \right)^2 \frac{1}{3} \left(\frac{2l_n}{a} \right)^3,$$

which is bounded and converges to

$$\left(\gamma + \frac{a}{2}\right)^2 \frac{1}{3} \left(\frac{2l}{a}\right)^3. \quad (2.B.22)$$

The second terms is dominated by $(n+1)/(n-1)$ times

$$\int_{2l_n/a}^1 \left(\gamma \left(1 - \frac{1}{\psi_n}\right) + \left(l_n + \frac{\gamma}{\psi_n}\right) (1 - \gamma\beta_n) \left(\frac{1-u}{1-2l_n/a}\right)^{\delta_n} \right)^2 du \quad (2.B.23)$$

$$+ \eta(u_n, w_n, z_n, 2\gamma(1-1/n)/a),$$

where η is defined by (2.B.8) and

$$w_n = \left(l_n + \frac{\gamma}{\psi_n}\right) \left(\frac{1-u_n}{1-2l_n/a}\right)^{\delta_n} - \frac{\gamma}{\psi_n},$$

$$z_n = -\frac{a\delta_n}{a-2l_n} \left(l_n + \frac{\gamma}{\psi_n}\right) \left(\frac{1-u_n}{1-2l_n/a}\right)^{\delta_n-1}.$$

Looking at the explicit expression of $\eta(u_n, w_n, z_n, 2\gamma(1-1/n)a)$, nontrivial estimates are necessary only for

$$(1-u_n)^{1+2\delta_n}$$

and

$$(a-2l_n)^{-2\delta_n}.$$

Since $\min_{x>0} x \log x = 1/e$, the latter is dominated by

$$\exp \left\{ \frac{2}{e} \frac{n-2}{n-1} \frac{1}{\gamma + l_n \psi_n} \right\}$$

and so, bounded thanks to (2.B.19). By definition of l_n and ψ_n , we have that

$$l_n \geq 0, \quad a - 2l_n > 0, \quad \left(2 + \frac{\psi_n}{2}\right) l_n \geq 2l$$

and that $l_n = 0$ when $\gamma < -a/2$. This implies that

$$1 + 2\delta_n = 1 - \frac{n-2}{n-1} \frac{a-2l_n}{\gamma + l_n \psi_n} > \begin{cases} 1 & \text{if } \gamma < -a/2, \\ \frac{\gamma-a+2l}{\gamma+a/2} \geq 0 & \text{if } \gamma > a/2. \end{cases}$$

It follows then that $\eta(u_n, w_n, z_n, 2\gamma(1-1/n)/a)$ is bounded in (n, s, t) . Moreover, since $u_n \rightarrow 1$, it converges to 0 for all (s, t) and this is easy to see if $\gamma \neq a$. In case $\gamma = a$, note that there exist constants C, C' such that

$$1 - u_n \leq C(C' + \beta_n r_n)^{-2(1-1/n)/(a\beta_n)}$$

and $\beta_n \rightarrow 1/a$. Since

$$1 + 2\delta_n > 1 - \frac{n-2}{n-1}$$

and $1/r_n = O(e^{-n^2})$, we have

$$(1-u_n)^{1+2\delta_n} \leq C(C' + \beta_n r_n)^{-2/(na\beta_n)} \rightarrow 0.$$

Now, let us see the first term of (2.B.23) is bounded in (n, s, t) and converges to

$$\frac{l^2(a+2\gamma)^2(\gamma+l)(a-2l)}{a^3(\gamma-a+3l)}. \quad (2.B.24)$$

By a straightforward computation,

$$\begin{aligned} & \left(l_n + \frac{\gamma}{\psi_n} \right)^2 (1 - \gamma\beta_n)^2 \int_{2l_n/a}^1 \left(\frac{1-u}{1-2l_n/a} \right)^{2\delta_n} du \\ &= \frac{(l_n\psi_n(a+2\gamma) + a\gamma(1-\psi_n))^2 a - 2l_n}{a^3\psi_n^2} \frac{1}{1+2\delta_n}. \end{aligned} \quad (2.B.25)$$

When $\gamma < -a/2$, we have $1 + 2\delta_n \geq 1$ as already seen and so, (2.B.25) is bounded. When $\gamma > a/2$, note that

$$\frac{l_n}{1+2\delta_n} \leq \frac{(\gamma + l_n\psi_n)l_n}{\gamma - a + l_n(2 + \psi_n)} \leq \frac{(\gamma + l_n\psi_n)l_n}{\gamma - a + 2l + l_n\psi_n/2} \leq 2 \frac{\gamma + l_n\psi_n}{\psi_n}$$

since $\gamma - a + 2l \geq 0$. Further, if $a/\gamma \leq 1 + 1/(2n)$, then

$$1 + 2\delta_n \geq 1 - \frac{n-2}{n-1} \frac{a}{\gamma} \geq 1 - \frac{n-2}{n-1} \left(1 + \frac{1}{2n} \right) = \frac{n+2}{2n(n-1)}$$

and if $a/\gamma > 1 + 1/(2n)$, then

$$1 + 2\delta_n > \frac{\gamma - a + 2l}{\gamma + a/2} = \frac{a(a-\gamma)}{(4\gamma-a)(\gamma+a/2)} > \frac{1}{2n+1} \frac{a^2}{(4\gamma-a)(\gamma+a/2)}$$

These estimates imply that $l_n/(1+2\delta_n)$ and $(1-\psi_n)/(1+2\delta_n)$ are bounded in (n, t, s) . It follows then that (2.B.25) is bounded in (n, t, s) . It is easy to see that (2.B.25) converges to (2.B.24) for all (s, t) . By a straightforward computation, we see that the sum of (2.B.22) and (2.B.24) coincides with $\gamma^2 \eta_{\dagger}(a/\gamma)$.

It only remains to show that (2.B.21) is bounded. By the same change of variable, we have that (2.B.21) is equal to

$$\frac{1}{n} \int_0^{2l_n/a} \left(\frac{2}{a}(u-1)x'(u) \right)^2 du + \frac{n+1}{n(n-1)} \int_{2l_n/a}^1 \left(\frac{2}{a} \left(1 - \frac{1}{n} \right) (u-1)x'(u) \right)^2 du.$$

The first term is less than $1/n$. It is then sufficient to estimate

$$\begin{aligned} & \frac{1}{n} \int_{2l_n/a}^1 (1-u)^2 |x'(u)|^2 du \\ & \leq \frac{\delta_n^2}{n} \left(l_n + \frac{\gamma}{\psi_n} \right)^2 \int_{2l_n/a}^1 \left(\frac{1-u}{1-2l_n/a} \right)^{2\delta_n} du + \frac{z_n^2}{3n} (1-u_n)^3. \end{aligned}$$

The right hand terms are shown to be bounded by using the above estimates to deal with (2.B.23). This completes the proof. \square

Part II

**ASYMPTOTIC OPTIMAL
TRACKING**

Chapter 3

Asymptotic lower bounds

Contents

3.1	Introduction	95
3.2	Tracking with combined regular and impulse control	100
3.2.1	Asymptotic framework	101
3.2.2	Lower bound	104
3.3	Extensions of Theorem 3.2.1 to other types of control	105
3.3.1	Combined regular and singular control	105
3.3.2	Impulse control	106
3.3.3	Singular control	107
3.3.4	Regular control	108
3.4	Interpretation of lower bounds and examples	108
3.4.1	Martingale problem associated to controlled Brownian motion	109
3.4.2	Time-average control of Brownian motion	110
3.4.3	Explicit examples in dimension one	113
3.5	Proof of Theorem 3.2.1	119
3.5.1	Reduction to local time-average control problem	119
3.5.2	Proof of Theorem 3.2.1	121
3.5.3	Proof of Lemma 3.5.3	122
3.6	Proof of Theorem 3.3.1	124
3.7	Proof of Propositions 3.4.1-3.4.5	124
3.7.1	Verification theorem in \mathbb{R}^d	124
3.7.2	Verification of Proposition 3.4.4	126
Appendix 3.A	Kummer confluent hypergeometric function ${}_1F_1$	128
Appendix 3.B	Proof of Lemma 3.7.2	129
Appendix 3.C	Tightness function	133
Appendix 3.D	Convergence in probability, stable convergence	134

3.1 Introduction

We consider the problem of tracking a target whose dynamics (X_t°) is modeled by a continuous Itô semi-martingale defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ with values in \mathbb{R}^d such that

$$dX_t^\circ = b_t dt + \sqrt{a_t} dW_t, \quad X_0^\circ = 0.$$

Here, (W_t) is a d -dimensional Brownian motion and (b_t) , (a_t) are predictable processes with values in \mathbb{R}^d and \mathcal{S}_+^d the set of $d \times d$ symmetric positive definite matrices respectively. An agent

observes X_t° and adjusts her position in order to follow X_t° . However, she has to pay certain intervention costs for position adjustments. The objective is to stay close to the target X_t° while minimizing the tracking efforts. This problem arises naturally in various situations such as discretization of option hedging strategies [Fuk11a, RT14, GL14a], management of an index fund [PS04, Kor99], control of exchange rate [MØ97, CZ00], portfolio selection under transaction costs [KMK15, ST13, PST15, AMKS15], trading under market impact [BSV15, MMKS14, LMKW14] or illiquidity costs [NW11, RS10].

More precisely, let (Y_t^ψ) be the position of the agent determined by the control ψ , with $Y_t^\psi \in \mathbb{R}^d$ and let (X_t) be the deviation of the agent from the target (X_t°) , so that

$$X_t = -X_t^\circ + Y_t^\psi. \quad (3.1.1)$$

Let $H_0(X)$ be the penalty functional for the deviation from the target and $H(\psi)$ the cost incurred by the control ψ up to a finite horizon T . Then the problem of tracking can be formulated as

$$\inf_{\psi \in \mathcal{A}} J(\psi), \quad J(\psi) = H_0(X) + H(\psi), \quad (3.1.2)$$

where \mathcal{A} is the set of admissible strategies. As is usually done in the literature (see for example [Kar83, MRT92]), we consider a penalty $H_0(X)$ for the deviation from the target of additive form

$$H_0(X) = \int_0^T r_t D(X_t) dt,$$

where (r_t) is a random weight process and $D(x)$ a deterministic function. For example, we can take $D(x) = \langle x, \Sigma^D x \rangle$ where Σ^D is positive definite and $\langle \cdot, \cdot \rangle$ is the inner product in \mathbb{R}^d . On the other hand, depending on the nature of the costs, the agent can either control her speed at all times or jump towards the target instantaneously. The control ψ and the cost functional $H(\psi)$ belong to one of the following classes:

1. *Impulse Control.* There is a fixed cost component for each action, so the agent has to intervene in a discrete way. The class \mathcal{A} of admissible controls contains all sequences $\{(\tau_j, \xi_j), j \in \mathbb{N}^*\}$ where $\{\tau_j, j \in \mathbb{N}^*\}$ is a strictly increasing sequence of stopping times representing the jump times and satisfying $\lim_{j \rightarrow \infty} \tau_j = +\infty$, and for each $j \in \mathbb{N}^*$, $\xi_j \in \mathbb{R}^d$ is a \mathcal{F}_{τ_j} -measurable random vector representing the size of j -th jump. The position of the agent is given by

$$Y_t = \sum_{0 < \tau_j \leq t} \xi_j$$

and the cumulated cost is then given by

$$H(\psi) = \sum_{0 < \tau_j \leq T} k_{\tau_j} F(\xi_j),$$

where (k_t) is a random weight process and $F(\xi) > 0$ is the cost of a jump with size $\xi \neq 0$. If we take $k_t = 1$ and $F(\xi) = \sum_{i=1}^d \mathbb{1}_{\{\xi^i \neq 0\}}$ where ξ^i is the i -th component of ξ , then $H(\psi)$ represents the total number of actions on each component over the time interval $[0, T]$, see [Fuk11a, GL14a]. If $F(\xi) = \sum_{i=1}^d \mathbb{1}_{\{\xi^i \neq 0\}} + \sum_{i=1}^d P_i |\xi^i|$ where $P_i \geq 0$, we say that the cost has a fixed component and a proportional component.

2. *Singular Control.* If the cost is proportional to the size of the jump, then infinitesimal displacement is also allowed and it is natural to model (Y_t) by a process with bounded variation. In this case, the class \mathcal{A} of admissible controls contains all couples (γ, φ) where φ is a progressively measurable increasing process with $\varphi_{0-} = 0$, which represents the cumulated amount of intervention and γ is a progressively measurable process with $\gamma_t \in \Delta := \{n \in \mathbb{R}^d \mid \sum_{i=1}^d |n^i| = 1\}$ for all $t \geq 0$, which represents the distribution of

the control effort in each direction. In other words, $\varphi_t = \sum_{i=1}^d \|Y^i\|_t$ where $\|\cdot\|$ denotes the absolute variation of a process, and γ_t^i is the Radon-Nikodym derivative of Y_t^i with respect to φ_t . The position of the agent is given by

$$Y_t = \int_0^t \gamma_s d\varphi_s,$$

and the corresponding cost is usually given as (see for example [KMK15, ST13])

$$H(\psi) = \int_0^T h_t P(\gamma_t) d\varphi_t,$$

where (h_t) is a random weight process and we take (for example) $P(\gamma) = \langle P, |\gamma| \rangle$ with $P \in \mathbb{R}_+^d$ and $|\gamma| = (|\gamma^1|, \dots, |\gamma^d|)^T$. The vector $P = (P_1, \dots, P_d)^T$ represents the coefficients of proportional costs in each direction.

3. *Regular Control.* Most often, the process (Y_t) is required to be absolutely continuous with respect to time, see for example [RS10, MMKS14] among many others. In this case, the class \mathcal{A} of admissible controls contains all progressively measurable integrable processes u with values in \mathbb{R}^d , representing the speed of the agent, the position of the agent is given by

$$Y_t = \int_0^t u_s ds,$$

and the cost functional is

$$H(\psi) = \int_0^T l_t Q(u_t) dt,$$

where (l_t) is a random weight process and, for example, $Q(u) = \langle u, \Sigma^Q u \rangle$ with Σ^Q a positive definite matrix. Comparing to the case of singular control where the control variables are (γ_t) and (φ_t) , here we optimize over (u_t) .

4. *Combined control.* It is possible that several types of control are available to the agent. In that case, $\psi = (\psi_1, \dots, \psi_n)$ where for each i , ψ_i belongs to one of the classes introduced before. For example, in the case of combined regular and impulse control (see [MØ97]), the position of the agent is given by

$$Y_t = \sum_{0 < \tau_j \leq t} \xi_j + \int_0^t u_s ds,$$

while the cost functional is given by

$$H(\psi) = \sum_{0 < \tau_j \leq T} k_{\tau_j} F(\xi_j) + \int_0^T l_t Q(u_t) dt.$$

Similarly, one can consider other combinations of controls.

The problem (3.1.1)-(3.1.2) rarely admits an explicit solution. In this chapter, we propose an asymptotic framework where the tracking costs are small and derive a lower bound for (3.1.1)-(3.1.2) under this setting. More precisely we introduce a parameter ε tending to zero and consider a family of cost functionals $H^\varepsilon(\psi)$. For example, we can have $H^\varepsilon(\psi) = \varepsilon^{\beta_\psi} H(\psi)$ for some constant β_ψ , but different components of the cost functional may also scale with ε at different rates. We define the control problem

$$X_t^\varepsilon = -X_t^\circ + Y_t^{\psi^\varepsilon}, \tag{3.1.1-ε}$$

and objective function

$$\inf_{\psi^\varepsilon \in \mathcal{A}} J^\varepsilon(\psi^\varepsilon), \quad J^\varepsilon(\psi^\varepsilon) = H_0(X^\varepsilon) + H^\varepsilon(\psi^\varepsilon). \quad (3.1.2-\varepsilon)$$

Moreover, we assume that the functions D, Q, F, P possess a homogeneity property.

The main result of this chapter is a precise asymptotic relation between J^ε and the time-average control problem of Brownian motion with constant parameters, in a variety of settings. Let us give a flavor of the main result in the case of combined regular and impulse control (note that situations involving singular control are considered in Section 3.3). In this case, the dynamics of the controlled Brownian motion is given by

$$X_s = \sqrt{a}W_s + \int_0^s u_\nu d\nu + \sum_{0 < \tau_j \leq s} \xi_j, \quad (3.1.1\text{-local})$$

and the time-average control problem can be formulated as

$$\tilde{I}(a, r, l, k) = \inf_{(\tau_j, \xi_j, u)} \overline{\lim}_{S \rightarrow \infty} \frac{1}{S} \mathbb{E} \left[\int_0^S (rD(X_s) + lQ(u_s)) ds + k \sum_{0 < \tau_j \leq S} F(\xi_j) \right]. \quad (3.1.2\text{-local})$$

At the level of generality that we are interested in, we need to consider a relaxed formulation of the above control problem, as a linear programming problem on the space of measures. Following [KS01], we introduce the occupation measures

$$\begin{aligned} \mu_t(H_1) &= \frac{1}{t} \mathbb{E} \int_0^t \mathbb{1}_{H_1}(X_s, u_s) ds, \quad H_1 \in \mathcal{B}(\mathbb{R}^d \times \mathbb{R}^d), \\ \rho_t(H_2) &= \frac{1}{t} \mathbb{E} \sum_{0 < \tau_j \leq t} \mathbb{1}_{H_2}(X_{s-}, \xi_j), \quad H_2 \in \mathcal{B}(\mathbb{R}^d \times \mathbb{R}^d). \end{aligned}$$

If the process X and the controls are stationary, these measures do not depend on time and therefore

$$\begin{aligned} \overline{\lim}_{S \rightarrow \infty} \frac{1}{S} \mathbb{E} \left[\int_0^S (rD(X_s) + lQ(u_s)) ds + k \sum_{0 < \tau_j \leq S} F(\xi_j) \right] \\ = \int_{\mathbb{R}^d \times \mathbb{R}^d} (rD(x) + lQ(u)) \mu(dx \times du) + \int_{\mathbb{R}^d \times \mathbb{R}^d} kF(\xi) \rho(dx \times d\xi). \end{aligned}$$

On the other hand, by Itô's formula, for any $f \in C_0^2$,

$$f(X_t) = f(X_0) + \sqrt{a} \int_0^t f'(X_s) dW_s + \int_0^t Af(X_s, u_s) ds + \sum_{0 < \tau_j \leq t} Bf(X_{\tau_j-}, \xi_j), \quad (3.1.3)$$

where

$$\begin{aligned} Af(x, u) &= \frac{1}{2} \sum_{i,j} a_{ij} \partial_{ij}^2 f(x) + \langle u, \nabla f(x) \rangle, \\ Bf(x, \xi) &= f(x + \xi) - f(x). \end{aligned}$$

Taking expectation in (3.1.3) and assuming once again the stationarity of controls, we see that under adequate integrability conditions the measures μ and ρ satisfy the constraint

$$\int_{\mathbb{R}^d \times \mathbb{R}^d} Af(x, u) \mu(dx \times du) + \int_{\mathbb{R}^d \times \mathbb{R}^d} Bf(x, \xi) \rho(dx \times d\xi) = 0. \quad (3.1.4)$$

Therefore, the time-average control problem of Brownian motion (3.1.1-local)-(3.1.2-local) is closely related to the problem of computing

$$I(a, r, l, k) = \inf_{\mu, \rho} \int_{\mathbb{R}^d \times \mathbb{R}^d} (rD(x) + lQ(u))\mu(dx \times du) + \int_{\mathbb{R}^d \times \mathbb{R}^d} kF(\xi)\rho(dx \times d\xi), \quad (3.1.5)$$

where μ is a probability measure and ρ is a finite positive measure satisfying the constraint (3.1.4). In Section 3.4 we shall see that this characterization is essentially equivalent to (3.1.1-local)-(3.1.2-local) if we formulate the optimal control problem for the Brownian motion as a controlled martingale problem. In the considered case of combined regular and impulse control, our main result is the following.

Main result, combined regular and impulse control. *There exists $\beta^* > 0$ explicitly determined by the cost structure H_0 and $(H^\varepsilon)_{\varepsilon > 0}$ such that for all $\delta > 0$ and any sequence of admissible strategies $\{\psi^\varepsilon \in \mathcal{A}, \varepsilon > 0\}$, we have*

$$\lim_{\varepsilon \rightarrow 0^+} \mathbb{P} \left[\frac{1}{\varepsilon^{\beta^*}} J^\varepsilon(\psi^\varepsilon) \geq \int_0^T I_t dt - \delta \right] = 1, \quad (3.1.6)$$

where $I_t = I(a_t, r_t, k_t, l_t)$ is the optimal cost of the linear programming formulation (3.1.5) of the time-average control problem of Brownian motion (3.1.1-local)-(3.1.2-local) with parameters frozen at time t .

The original problem (3.1.1)-(3.1.2) is therefore simplified in the sense that the local problem (3.1.1-local)-(3.1.2-local) is easier to analyze since the dynamics of the target is reduced to that of a Brownian motion and the cost parameters become constant. In many practically important cases (see Examples 3.4.3-3.4.7), we are able to solve explicitly (3.1.1-local)-(3.1.2-local) and show that the two formulations of the time-average control problem are equivalent and therefore $I = \tilde{I}$. Moreover, in Chapter 4, we show that for the examples where (3.1.1-local)-(3.1.2-local) admits an explicit solution, the lower bound is tight (see Remark 3.3.1).

Our result enables us to revisit the asymptotic lower bounds for the discretization of hedging strategies in [Fuk11a, GL14a]. In these papers, the lower bounds are deduced by using subtle inequalities. In Chapter 4 we show that these bounds can be interpreted in a simple manner through the time-average control problem of Brownian motion.

The local control problem (3.1.1-local)-(3.1.2-local) also arises in the study of utility maximization under transaction costs, see [ST13, PST15, AMKS15, MMKS14]. This is not surprising since at first order, these problems and the tracking problem are essentially the same, see Section 4.4.3. In the above references, the authors derive the PDE associated to the first order correction of the value function, which turns out to be the HJB equation associated to the time-average control of Brownian motion. Inspired by [KS99] and [KM93], our approach, based on weak convergence of empirical occupation measures, is very different from the PDE-based method and enables us to treat more general situations. Contrary to [KS99], where the lower bound holds under expectation, we obtain pathwise lower bounds. Compared to [KM93], we are able to treat impulse control and general dynamics for the target.

The paper is organized as follows. In Section 3.2, we introduce our asymptotic framework and establish heuristically the lower bound for the case of combined regular and impulse control. Various extensions are then discussed in Section 3.3. In Section 3.4, we provide an accurate definition for the time-average control of Brownian motion using a relaxed martingale formulation and collect a comprehensive list of explicit solutions in dimension one. The proofs are given in Sections 3.5, 3.6 and 3.7.

Notation. For a complete, separable, metric space S , we define $C(S)$ the set of continuous functions on S , $C_b(S)$ the set of bounded, continuous functions on S , $\mathcal{M}(S)$ the set of finite nonnegative Borel measures on S and $\mathcal{P}(S)$ the set of probability measures on S . The sets $\mathcal{M}(S)$ and $\mathcal{P}(S)$ are equipped with the topology of weak convergence. We define $\mathcal{L}(S)$ the set of nonnegative Borel measures Γ on $S \times [0, \infty)$ such that $\Gamma(S \times [0, t]) < \infty$. Denote Γ_t the restriction of Γ to $S \times [0, t]$. We use \mathbb{R}_x^d , \mathbb{R}_u^d and \mathbb{R}_ξ^d to indicate the state space corresponding to the variables x , u and ξ . Finally, $C_0^2(\mathbb{R}^d)$ denotes the set of twice differentiable real functions on \mathbb{R}^d with compact support, equipped with the norm

$$\|f\|_{C_0^2} = \|f\|_\infty + \sum_{i=1}^d \|\partial_i f\|_\infty + \sum_{i,j=1}^d \|\partial_{ij}^2 f\|_\infty.$$

3.2 Tracking with combined regular and impulse control

Instead of giving directly a general result, which would lead to a cumbersome presentation, we focus in this section on the tracking problem with combined regular and impulse control. This allows us to illustrate our key ideas. Other situations, such as singular control, are discussed in Section 3.3.

In the case of combined regular and impulse control, a tracking strategy (u, τ, ξ) is given by a progressively measurable process $u = (u_t)_{t \geq 0}$ with values in \mathbb{R}^d and $(\tau, \xi) = \{(\tau_j, \xi_j), j \in \mathbb{N}^*\}$, with (τ_j) an increasing sequence of stopping times and (ξ_j) a sequence of \mathcal{F}_{τ_j} -measurable random variables with values in \mathbb{R}^d . The process (u_t) represents the speed of the agent. The stopping time τ_j represents the timing of j -th jump towards the target and ξ_j the size of the jump. The tracking error obtained by following the strategy (u, τ, ξ) is given by

$$X_t = -X_t^\circ + \int_0^t u_s ds + \sum_{0 < \tau_j \leq t} \xi_j.$$

At any time the agent is paying a cost for maintaining the speed u_t and each jump ξ_j incurs a positive cost. We are interested in the following type of cost functional

$$J(u, \tau, \xi) = \int_0^T (r_t D(X_t) + l_t^\circ Q(u_t)) dt + \sum_{j: 0 < \tau_j \leq T} (k_{\tau_j}^\circ F(\xi_j) + h_{\tau_j}^\circ P(\xi_j)),$$

where $T \in \mathbb{R}^+$ and (r_t) , (l_t°) , (k_t°) and (h_t°) are random weight processes. The cost functions D , Q , F , P are deterministic functions which satisfy the following homogeneity property

$$D(\varepsilon x) = \varepsilon^{\zeta_D} D(x), \quad Q(\varepsilon u) = \varepsilon^{\zeta_Q} Q(u), \quad F(\varepsilon \xi) = \varepsilon^{\zeta_F} F(\xi), \quad P(\varepsilon \xi) = \varepsilon^{\zeta_P} P(\xi), \quad (3.2.1)$$

for any $\varepsilon > 0$ and

$$\zeta_D > 0, \quad \zeta_Q > 1, \quad \zeta_F = 0, \quad \zeta_P = 1.$$

Note that here we slightly extend the setting of the previous section by introducing two functions F and P which typically represent the fixed and the proportional costs respectively.

In this chapter, we essentially have in mind the case where

$$D(x) = \langle x, \Sigma^D x \rangle, \quad Q(u) = \langle u, \Sigma^Q u \rangle, \quad F(\xi) = \sum_{i=1}^d F_i \mathbb{1}_{\{\xi^i \neq 0\}}, \quad P(\xi) = \sum_{i=1}^d P_i |\xi^i|,$$

with $F_i, P_i \in \mathbb{R}_+$ such that

$$\min_i F_i > 0, \quad (3.2.2)$$

and $\Sigma^D, \Sigma^Q \in \mathcal{S}_+^d$. Note that in this situation, we have $\zeta_D = \zeta_Q = 2$.

3.2.1 Asymptotic framework

We now explain our asymptotic setting where the costs are small and provide a heuristic proof of our main result. We assume that there exist $\varepsilon > 0$ and $\beta_Q, \beta_F, \beta_P > 0$ such that

$$l_t^\circ = \varepsilon^{\beta_Q} l_t, \quad k_t^\circ = \varepsilon^{\beta_F} k_t, \quad h_t^\circ = \varepsilon^{\beta_P} h_t. \quad (3.2.3)$$

Then the asymptotic framework consists in considering the sequence of optimization problems indexed by $\varepsilon \rightarrow 0$

$$\inf_{(u^\varepsilon, \tau^\varepsilon, \xi^\varepsilon) \in \mathcal{A}} J^\varepsilon(u^\varepsilon, \tau^\varepsilon, \xi^\varepsilon),$$

with

$$J^\varepsilon(u^\varepsilon, \tau^\varepsilon, \xi^\varepsilon) = \int_0^T (r_t D(X_t^\varepsilon) + \varepsilon^{\beta_Q} l_t Q(u_t^\varepsilon)) dt + \sum_{j:0 < \tau_j^\varepsilon \leq T} (\varepsilon^{\beta_F} k_{\tau_j^\varepsilon} F(\xi_j^\varepsilon) + \varepsilon^{\beta_P} h_{\tau_j^\varepsilon} P(\xi_j^\varepsilon)),$$

and

$$X_t^\varepsilon = -X_t^\circ + \int_0^t u_s^\varepsilon ds + \sum_{j:0 < \tau_j^\varepsilon \leq t} \xi_j^\varepsilon.$$

The key observation is that under such setting, the tracking problem can be decomposed into a sequence of local problems. More precisely, let $\{t_k^\varepsilon = k\delta^\varepsilon, k = 0, 1, \dots, K^\varepsilon\}$ be a partition of the interval $[0, T]$ with $\delta^\varepsilon \rightarrow 0$ as $\varepsilon \rightarrow 0$. Then we can write

$$\begin{aligned} J^\varepsilon(u^\varepsilon, \tau^\varepsilon, \xi^\varepsilon) &= \sum_{k=0}^{K^\varepsilon-1} \left(\int_{t_k^\varepsilon}^{t_k^\varepsilon + \delta^\varepsilon} (r_t D(X_t^\varepsilon) + \varepsilon^{\beta_Q} l_t Q(u_t^\varepsilon)) dt \right. \\ &\quad \left. + \sum_{j:t_k^\varepsilon < \tau_j^\varepsilon \leq t_k^\varepsilon + \delta^\varepsilon} (\varepsilon^{\beta_F} k_{\tau_j^\varepsilon} F(\xi_j^\varepsilon) + \varepsilon^{\beta_P} h_{\tau_j^\varepsilon} P(\xi_j^\varepsilon)) \right) \\ &= \sum_{k=0}^{K^\varepsilon-1} \frac{1}{\delta^\varepsilon} \left(\int_{t_k^\varepsilon}^{t_k^\varepsilon + \delta^\varepsilon} (r_t D(X_t^\varepsilon) + \varepsilon^{\beta_Q} l_t Q(u_t^\varepsilon)) dt \right. \\ &\quad \left. + \sum_{j:t_k^\varepsilon < \tau_j^\varepsilon \leq t_k^\varepsilon + \delta^\varepsilon} (\varepsilon^{\beta_F} k_{\tau_j^\varepsilon} F(\xi_j^\varepsilon) + \varepsilon^{\beta_P} h_{\tau_j^\varepsilon} P(\xi_j^\varepsilon)) \right) (t_{k+1}^\varepsilon - t_k^\varepsilon) \\ &= \sum_{k=0}^{K^\varepsilon-1} j_{t_k^\varepsilon}^\varepsilon (t_{k+1}^\varepsilon - t_k^\varepsilon), \end{aligned}$$

with

$$j_{t_k^\varepsilon}^\varepsilon = \frac{1}{\delta^\varepsilon} \left(\int_{t_k^\varepsilon}^{t_k^\varepsilon + \delta^\varepsilon} (r_t D(X_t^\varepsilon) + \varepsilon^{\beta_Q} l_t Q(u_t^\varepsilon)) dt + \sum_{j:t_k^\varepsilon < \tau_j^\varepsilon \leq t_k^\varepsilon + \delta^\varepsilon} (\varepsilon^{\beta_F} k_{\tau_j^\varepsilon} F(\xi_j^\varepsilon) + \varepsilon^{\beta_P} h_{\tau_j^\varepsilon} P(\xi_j^\varepsilon)) \right).$$

As ε tends to zero, we approximately have

$$J^\varepsilon(u^\varepsilon, \tau^\varepsilon, \xi^\varepsilon) \simeq \int_0^T j_t^\varepsilon dt.$$

We are hence led to study j_t^ε as $\varepsilon \rightarrow 0$, which is closely related to the time-average control problem of Brownian motion. To see this, consider the following rescaling of X^ε over the horizon $(t, t + \delta^\varepsilon]$:

$$\tilde{X}_s^{\varepsilon, t} = \frac{1}{\varepsilon^\beta} X_{t+\varepsilon^\alpha \beta s}^\varepsilon, \quad s \in (0, T^\varepsilon],$$

where $T^\varepsilon = \varepsilon^{-\alpha\beta} \delta^\varepsilon$, $\alpha = 2$ and $\beta > 0$ is to be determined (here $\alpha = 2$ is related to the scaling property of Brownian motion). We use the superscript t to indicate that the scaled

systems correspond to the horizon $(t, t + \delta^\varepsilon]$. Then the dynamics of $\tilde{X}^{\varepsilon, t}$ is given by, see [RY99, Proposition V.1.5],

$$\tilde{X}_s^{\varepsilon, t} = \tilde{X}_0^{\varepsilon, t} + \int_0^s \tilde{b}_\nu^{\varepsilon, t} d\nu + \int_0^s \sqrt{\tilde{a}_\nu^{\varepsilon, t}} d\tilde{W}_\nu^{\varepsilon, t} + \int_0^s \tilde{u}_\nu^{\varepsilon, t} d\nu + \sum_{0 < \tilde{\tau}_j^{\varepsilon, t} \leq s} \tilde{\xi}_j^\varepsilon,$$

with

$$\tilde{b}_s^{\varepsilon, t} = -\varepsilon^{(\alpha-1)\beta} b_{t+\varepsilon\alpha\beta s}, \quad \tilde{a}_s^{\varepsilon, t} = a_{t+\varepsilon\alpha\beta s}, \quad \tilde{W}_s^{\varepsilon, t} = -\frac{1}{\varepsilon\beta} (W_{t+\varepsilon\alpha\beta s} - W_t),$$

and

$$\tilde{u}_s^{\varepsilon, t} = \varepsilon^{(\alpha-1)\beta} u_{t+\varepsilon\alpha\beta s}^\varepsilon, \quad \tilde{\xi}_j^\varepsilon = \frac{1}{\varepsilon\beta} \xi_j^\varepsilon, \quad \tilde{\tau}_j^{\varepsilon, t} = \frac{1}{\varepsilon\alpha\beta} (\tau_j^\varepsilon - t) \vee 0.$$

Note that $(\tilde{W}_s^{\varepsilon, t})$ a Brownian motion with respect to $\tilde{\mathcal{F}}_s^{\varepsilon, t} = \mathcal{F}_{t+\varepsilon\alpha\beta s}$. Abusing notation slightly, we write

$$d\tilde{X}_s^{\varepsilon, t} = \tilde{b}_s^{\varepsilon, t} ds + \sqrt{\tilde{a}_s^{\varepsilon, t}} d\tilde{W}_s^{\varepsilon, t} + \tilde{u}_s^{\varepsilon, t} ds + d\left(\sum_{0 < \tilde{\tau}_j^{\varepsilon, t} \leq s} \tilde{\xi}_j^\varepsilon\right), \quad s \in (0, T^\varepsilon]. \quad (3.2.4)$$

Using the homogeneity properties (3.2.1) of the cost functions, we obtain

$$\begin{aligned} j_t^\varepsilon &= \frac{1}{T^\varepsilon} \left(\int_0^{T^\varepsilon} (\varepsilon^{\beta\zeta_D} r_{t+\varepsilon\alpha\beta s} D(\tilde{X}_s^{\varepsilon, t}) + \varepsilon^{\beta_Q - (\alpha-1)\zeta_Q\beta} l_{t+\varepsilon\alpha\beta s} Q(\tilde{u}_s^{\varepsilon, t})) ds \right. \\ &\quad \left. + \sum_{0 < \tilde{\tau}_j^{\varepsilon, t} \leq T^\varepsilon} (\varepsilon^{\beta_F - (\alpha - \zeta_F)\beta} k_{t+\varepsilon\alpha\beta\tilde{\tau}_j^{\varepsilon, t}} F(\tilde{\xi}_j^\varepsilon) + \varepsilon^{\beta_P - (\alpha - \zeta_P)\beta} h_{t+\varepsilon\alpha\beta\tilde{\tau}_j^{\varepsilon, t}} P(\tilde{\xi}_j^\varepsilon)) \right) \\ &\simeq \frac{1}{T^\varepsilon} \left(\int_0^{T^\varepsilon} (\varepsilon^{\beta\zeta_D} r_t D(\tilde{X}_s^{\varepsilon, t}) + \varepsilon^{\beta_Q - (\alpha-1)\zeta_Q\beta} l_t Q(\tilde{u}_s^{\varepsilon, t})) ds \right. \\ &\quad \left. + \sum_{0 < \tilde{\tau}_j^{\varepsilon, t} \leq T^\varepsilon} (\varepsilon^{\beta_F - (\alpha - \zeta_F)\beta} k_t F(\tilde{\xi}_j^\varepsilon) + \varepsilon^{\beta_P - (\alpha - \zeta_P)\beta} h_t P(\tilde{\xi}_j^\varepsilon)) \right). \end{aligned}$$

The second approximation can be justified by the continuity of the cost coefficients r_t , l_t , k_t and h_t .

Now, if there exists $\beta > 0$ such that

$$\beta\zeta_D = \beta_Q - (\alpha - 1)\zeta_Q\beta = \beta_F - (\alpha - \zeta_F)\beta = \beta_P - (\alpha - \zeta_P)\beta,$$

that is,

$$\beta = \frac{\beta_F}{\zeta_D + \alpha - \zeta_F} = \frac{\beta_P}{\zeta_D + \alpha - \zeta_P} = \frac{\beta_Q}{\zeta_D + (\alpha - 1)\zeta_Q}, \quad (3.2.5)$$

where $\alpha = 2$, then we have

$$j_t^\varepsilon \simeq \varepsilon^{\beta\zeta_D} I_t^\varepsilon,$$

with

$$I_t^\varepsilon = \frac{1}{T^\varepsilon} \left(\int_0^{T^\varepsilon} (r_t D(\tilde{X}_s^{\varepsilon, t}) + l_t Q(\tilde{u}_s^{\varepsilon, t})) ds + \sum_{0 < \tilde{\tau}_j^{\varepsilon, t} \leq T^\varepsilon} (k_t F(\tilde{\xi}_j^\varepsilon) + h_t P(\tilde{\xi}_j^\varepsilon)) \right). \quad (3.2.6)$$

By suitably choosing δ^ε , we have

$$\delta^\varepsilon \rightarrow 0, \quad T^\varepsilon \rightarrow \infty.$$

It follows that $\tilde{b}_s^{\varepsilon, t} \simeq 0$ and $\tilde{a}_s^{\varepsilon, t} \simeq a_t$ for $s \in (0, T^\varepsilon]$. Therefore, the dynamics of (3.2.4) is approximately a controlled Brownian motion with diffusion matrix a_t . We deduce that

$$I_t^\varepsilon \gtrsim I(a_t, r_t, l_t, k_t, h_t), \quad (3.2.7)$$

where the term in the right-hand side is defined by

$$I(a, r, l, k, h) = \inf_{(u, \tau, \xi)} \limsup_{S \rightarrow \infty} \frac{1}{S} \left[\int_0^S (rD(X_s) + lQ(u_s)) ds + \sum_{0 \leq \tau_j \leq S} (kF(\xi_j) + hP(\xi_j)) \right], \quad (3.2.8)$$

with

$$X_s = \sqrt{a}W_s + \int_0^s u_r dr + \sum_{0 \leq \tau_j \leq s} \xi_j. \quad (3.2.9)$$

Therefore, we obtain that as $\varepsilon \rightarrow 0$:

$$J^\varepsilon \simeq \int_0^T j_t^\varepsilon dt \simeq \varepsilon^{\beta\zeta_D} \int_0^T I_t^\varepsilon dt \gtrsim \varepsilon^{\beta\zeta_D} \int_0^T I_t dt.$$

Then we may expect that (3.2.8) is equal to the following expected cost criterion

$$I(a, r, l, k, h) = \inf_{(u, \tau, \xi)} \limsup_{S \rightarrow \infty} \frac{1}{S} \mathbb{E} \left[\int_0^S (rD(X_s) + lQ(u_s)) ds + \sum_{0 \leq \tau_j \leq S} (kF(\xi_j) + hP(\xi_j)) \right], \quad (3.2.10)$$

see for example [BG88, JZ06c, JZ06a]. Therefore, we will use the latter version to characterize the lower bound since it is easier to manipulate.

Remark 3.2.1. *The approach of weak convergence is classical for proving inequalities similar to (3.2.7), in particular in the study of heavy traffic networks (see [Kus14, section 9] for an overview). The usual weak convergence theorems enable one to show that the perturbed system converges in the Skorohod topology to the controlled Brownian motion as ε tends to zero. However, since the time horizon tends to infinity, this does not immediately imply the convergence of time-average cost functionals like I_t^ε .*

In [KM93], the authors consider pathwise average cost problems for controlled queues in the heavy traffic limit, where the control term is absolutely continuous. They use the empirical “functional occupation measure” on the canonical path space and characterize the limit as a controlled Brownian motion. The same method has also been used in [BGL11] in the study of single class queueing networks.

However, this approach cannot be applied directly to singular/impulse controls for which the tightness of the occupation measures is difficult to establish. In fact, the usual Skorohod topology is not suitable for the impulse control term

$$\tilde{Y}_t^\varepsilon := \sum_{0 < \tilde{\tau}_j^\varepsilon \leq t} \tilde{\xi}_j^\varepsilon.$$

Indeed, in the case of singular/impulse control, the component $\{\tilde{Y}^\varepsilon\}$ is generally not tight under the Skorohod topology. For example (see [Kus01, p.72]), consider the family (Y_t^ε) where the function Y_t^ε equals zero for $t < 1$ and jumps upward by an amount $\sqrt{\varepsilon}$ at times $1 + i\varepsilon$, $i = 0, 1, \dots, \varepsilon^{-1/2}$ until it reaches the value unity. The natural limit of Y^ε is of course $\mathbb{1}_{\{t \geq 1\}}$ but this sequence is not tight in the Skorohod topology. The nature of this convergence is discussed in [Kur91] and a corresponding topology is provided in [Jak97].

This difficulty could be avoided by introducing a random time change after which the (suitably interpolated) control term becomes uniformly Lipschitz and hence converges under the Skorohod topology. This technique is used in [BG06, BG12, Kus01] to study the convergence of controlled queues with discounted costs. This seems to be a possible alternative way to extend the approach

of [KM93] to singular/impulse controls. Nevertheless, the analysis would probably be quite involved.

Instead of proving the tightness of control terms $(\tilde{Y}_t^\varepsilon)$ in weaker topologies, we shall use an alternative characterization of the time-average control problem of Brownian motion. In [KS99], the authors characterize the time-average control of a Jackson network in the heavy traffic limit as the solution of a linear program. The use of occupation measure on the state space instead of the path space turns out to be sufficient to describe the limiting stochastic control problem. However, the optimization criterion is not pathwise.

In this chapter, we use a combination of the techniques in [KS99] and [KM93] to obtain pathwise lower bounds.

3.2.2 Lower bound

In order to properly state our result for the case of combined regular and impulse control, we first introduce the solution $I = I(a, r, l, k, h)$ of the following linear programming problem:

$$I(a, r, l, k, h) = \inf_{(\mu, \rho)} \int_{\mathbb{R}_x^d \times \mathbb{R}_u^d} (rD(x) + lQ(u))\mu(dx \times du) + \int_{\mathbb{R}_x^d \times \mathbb{R}_\xi^d \setminus \{0_\xi\}} (kF(\xi) + hP(\xi))\rho(dx \times d\xi), \quad (3.2.11)$$

with $(\mu, \rho) \in \mathcal{P}(\mathbb{R}_x^d \times \mathbb{R}_u^d) \times \mathcal{M}(\mathbb{R}_x^d \times \mathbb{R}_\xi^d \setminus \{0_\xi\})$ satisfying the following constraint

$$\int_{\mathbb{R}_x^d \times \mathbb{R}_u^d} A^a f(x, u)\mu(dx \times du) + \int_{\mathbb{R}_x^d \times \mathbb{R}_\xi^d \setminus \{0_\xi\}} Bf(x, \xi)\rho(dx \times d\xi) = 0, \quad \forall f \in C_0^2(\mathbb{R}_x^d), \quad (3.2.12)$$

where

$$A^a f(x, u) = \frac{1}{2} \sum_{i,j} a_{ij} \partial_{ij}^2 f(x) + \langle u, \nabla f(x) \rangle, \quad Bf(x, \xi) = f(x + \xi) - f(x).$$

We will see in Section 3.4.2 that it is essentially an equivalent characterization of the time-average control problem (3.2.9)-(3.2.10). In Example 3.4.6, we consider a particular case for which I and the optimal solution μ^*, ρ^* can be explicitly determined. From now on, we make the following assumptions.

Assumption 3.2.1 (Regularity of linear programming). *The function $I = I(a, r, l, k, h)$ defined by (3.2.11)-(3.2.12) is measurable.*

Assumption 3.2.2 (Model). *The predictable processes (a_t) and (b_t) are continuous and (a_t) is positive definite on $[0, T]$.*

Assumption 3.2.3 (Optimization criterion). *The parameters of the cost functional (r_t) , (l_t) , (k_t) and (h_t) are continuous and positive on $[0, T]$.*

Assumption 3.2.4 (Asymptotic framework). *The cost functionals satisfy the homogeneity property (3.2.1) and the relation (3.2.5) holds for some $\beta > 0$.*

Remark 3.2.2. *Let us comment briefly the above assumptions. Assumption 3.2.1 is necessary to avoid pathological cases. In most examples, the function I is continuous (see Examples 3.4.3-3.4.7). Assumptions 3.2.2-3.2.3 impose minimal regularity on the dynamics of X° and cost parameters. Assumption 3.2.4 ensures that all the costs have similar order of magnitude.*

Second, we introduce the following notion.

Definition 3.2.1. Let $\{Z, (Z^\varepsilon)_\varepsilon, \varepsilon > 0\}$ be random variables on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We say that Z^ε is asymptotically bounded from below by Z in probability if

$$\forall \delta > 0, \quad \lim_{\varepsilon \rightarrow 0} \mathbb{P}[Z^\varepsilon > Z - \delta] = 1.$$

We write $\liminf_{\varepsilon \rightarrow 0} Z^\varepsilon \geq_p Z$.

We now give the version of our main result for the case of combined regular and impulse control.

Theorem 3.2.1 (Asymptotic lower bound for combined regular and impulse control). *Under Assumptions 3.2.1, 3.2.2, 3.2.3 and 3.2.4, we have*

$$\liminf_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon^{\beta\zeta_D}} J^\varepsilon(u^\varepsilon, \tau^\varepsilon, \xi^\varepsilon) \geq_p \int_0^T I(a_t, r_t, l_t, k_t, h_t) dt, \quad (3.2.13)$$

for any sequence of admissible tracking strategies $\{(u^\varepsilon, \tau^\varepsilon, \xi^\varepsilon) \in \mathcal{A}, \varepsilon > 0\}$.

Thus, in Theorem 3.2.1 we have expressed the lower bound for the tracking problem in terms of the integral of the solution I of a linear program, which will be interpreted as time-average control of Brownian motion in Section 3.4. For any subsequence ε' , we can always pick a further subsequence ε'' such that

$$\liminf_{\varepsilon'' \rightarrow 0} \frac{1}{(\varepsilon'')^{\beta\zeta_D}} J^{\varepsilon''}(u^{\varepsilon''}, \tau^{\varepsilon''}, \xi^{\varepsilon''}) \geq \int_0^T I(a_t, r_t, l_t, k_t, h_t) dt - \delta,$$

almost surely. Therefore, by Fatou's lemma, the following corollary holds.

Corollary 3.2.1. *We have*

$$\liminf_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon^{\zeta_D \beta}} \mathbb{E}[J^\varepsilon(u^\varepsilon, \tau^\varepsilon, \xi^\varepsilon)] \geq \mathbb{E}\left[\int_0^T I(a_t, r_t, l_t, k_t, h_t) dt\right].$$

3.3 Extensions of Theorem 3.2.1 to other types of control

In this section, we consider the case of combined regular and singular control and those with only one type of control. In particular, we will see that in the presence of singular control, the operator B is different. Formally we could give similar results for the combination of all three controls or even in the presence of several controls of the same type with different cost functions and scaling properties. To avoid cumbersome notation, we restrict ourselves to the cases meaningful in practice, which are illustrated by explicit examples in Section 3.4.

3.3.1 Combined regular and singular control

When the fixed cost component is absent, that is $F = 0$, impulse control and singular control can be merged. In that case, the natural way to formulate the tracking problem is to consider a strategy $(u^\varepsilon, \gamma^\varepsilon, \varphi^\varepsilon)$ with u^ε a progressively measurable process as before, $\gamma_t^\varepsilon \in \Delta$ and (φ_t^ε) a possibly discontinuous non-decreasing process such that

$$X_t^\varepsilon = -X_t^o + \int_0^t u_s^\varepsilon ds + \int_0^t \gamma_s^\varepsilon d\varphi_s^\varepsilon,$$

and

$$J^\varepsilon(u^\varepsilon, \gamma^\varepsilon, \varphi^\varepsilon) = \int_0^T (r_t D(X_t^\varepsilon) + \varepsilon^{\beta_Q} l_t Q(u_t^\varepsilon)) dt + \int_0^T \varepsilon^{\beta_P} h_t P(\gamma_t^\varepsilon) d\varphi_t^\varepsilon.$$

To avoid degeneracy, we assume that for any $\gamma \in \Delta$,

$$P(\gamma) > 0. \quad (3.3.1)$$

Using similar heuristic arguments as those in the previous section, we are led to consider the time-average control of Brownian motion with combined regular and singular control

$$I(a, r, l, h) = \inf_{(u, \gamma, \varphi)} \limsup_{S \rightarrow \infty} \frac{1}{S} \mathbb{E} \left[\int_0^S (rD(X_s) + lQ(u_s)) ds + \int_0^S h_s P(\gamma_s) d\varphi_s \right], \quad (3.3.2)$$

where

$$X_s = \sqrt{a}W_s + \int_0^s u_r dr + \int_0^s \gamma_r d\varphi_r. \quad (3.3.3)$$

The corresponding linear programming problem is given by

$$I(a, r, l, h) = \inf_{(\mu, \rho)} \int_{\mathbb{R}_x^d \times \mathbb{R}_u^d} (rD(x) + lQ(u)) \mu(dx \times du) + \int_{\mathbb{R}_x^d \times \Delta \times \mathbb{R}_\delta^+} hP(\gamma) \rho(dx \times d\gamma \times d\delta), \quad (3.3.4)$$

with $(\mu, \rho) \in \mathcal{P}(\mathbb{R}_x^d \times \mathbb{R}_u^d) \times \mathcal{M}(\mathbb{R}_x^d \times \Delta \times \mathbb{R}_\delta^+)$ satisfying the following constraint

$$\int_{\mathbb{R}_x^d \times \mathbb{R}_u^d} A^a f(x, u) \mu(dx \times du) + \int_{\mathbb{R}_x^d \times \Delta \times \mathbb{R}_\delta^+} Bf(x, \gamma, \delta) \rho(dx \times d\gamma \times d\delta) = 0, \quad \forall f \in C_0^2(\mathbb{R}_x^d), \quad (3.3.5)$$

where

$$A^a f(x, u) = \frac{1}{2} \sum_{ij} a_{ij} \partial_{ij}^2 f(x) + \langle u, \nabla f(x) \rangle, \quad Bf(x, \gamma, \delta) = \begin{cases} \langle \gamma, \nabla f(x) \rangle, & \delta = 0, \\ \delta^{-1} (f(x + \delta\gamma) - f(x)), & \delta > 0. \end{cases}$$

We have the following theorem.

Theorem 3.3.1 (Asymptotic lower bound for combined regular and singular control). *Assume that $I(a, r, l, h)$ is measurable, that the parameters (r_t) , (l_t) and (h_t) are continuous and positive, that Assumption 3.2.2 holds true and that Assumption 3.2.4 is satisfied for some $\beta > 0$. Then,*

$$\liminf_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon^{\beta} \zeta_D} J^\varepsilon(u^\varepsilon, \gamma^\varepsilon, \varphi^\varepsilon) \geq_p \int_0^T I(a_t, r_t, l_t, h_t) dt, \quad (3.3.6)$$

for any sequence of admissible tracking strategies $\{(u^\varepsilon, \gamma^\varepsilon, \varphi^\varepsilon) \in \mathcal{A}, \varepsilon > 0\}$.

Adapting the proofs of Theorem 3.2.1 and Theorem 3.3.1 in an obvious way, we easily obtain the following bounds when only one control is present.

3.3.2 Impulse control

Consider

$$\inf_{(\tau^\varepsilon, \xi^\varepsilon) \in \mathcal{A}} J^\varepsilon(u^\varepsilon, \tau^\varepsilon, \xi^\varepsilon),$$

with

$$J^\varepsilon(u^\varepsilon, \tau^\varepsilon, \xi^\varepsilon) = \int_0^T r_t D(X_t^\varepsilon) dt + \sum_{0 < \tau_j^\varepsilon \leq T} (\varepsilon^{\beta_F} k_{\tau_j^\varepsilon} F(\xi_j^\varepsilon) + \varepsilon^{\beta_P} h_{\tau_j^\varepsilon} P(\xi_j^\varepsilon)),$$

and

$$X_t^\varepsilon = -X_t^o + \sum_{0 < \tau_j^\varepsilon \leq t} \xi_j^\varepsilon.$$

We have the following theorem.

Theorem 3.3.2 (Asymptotic lower bound for impulse control). *Let $I = I(a, r, k, h)$ be given by*

$$I(a, r, k, h) = \inf_{(\mu, \rho)} \int_{\mathbb{R}_x^d} rD(x) \mu(dx) + \int_{\mathbb{R}_x^d \times \mathbb{R}_\xi^d \setminus \{0_\xi\}} (kF(\xi) + hP(\xi)) \rho(dx \times d\xi),$$

with $(\mu, \rho) \in \mathcal{P}(\mathbb{R}_x^d) \times \mathcal{M}(\mathbb{R}_x^d \times \mathbb{R}_\xi^d \setminus \{0_\xi\})$ satisfying the following constraint

$$\int_{\mathbb{R}_x^d} A^a f(x) \mu(dx) + \int_{\mathbb{R}_x^d \times \mathbb{R}_\xi^d \setminus \{0_\xi\}} Bf(x, \xi) \rho(dx \times d\xi) = 0, \quad \forall f \in C_0^2(\mathbb{R}_x^d),$$

where

$$A^a f(x) = \frac{1}{2} \sum_{i,j} a_{ij} \partial_{ij}^2 f(x), \quad Bf(x, \xi) = f(x + \xi) - f(x).$$

Assume that $I(a, r, k, h)$ is measurable, that the parameters (r_t) , (k_t) and (h_t) are continuous and positive, that Assumption 3.2.2 holds true, and that Assumption 3.2.4 is satisfied for some $\beta > 0$. Then,

$$\liminf_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon^{\beta \zeta_D}} J^\varepsilon(\tau^\varepsilon, \xi^\varepsilon) \geq_p \int_0^T I(a_t, r_t, k_t, h_t) dt.$$

See Example 3.4.5 for a closed form solution of I .

3.3.3 Singular control

Consider

$$\inf_{(\gamma^\varepsilon, \varphi^\varepsilon) \in \mathcal{A}} J^\varepsilon(\gamma^\varepsilon, \varphi^\varepsilon),$$

with

$$J^\varepsilon(\gamma^\varepsilon, \varphi^\varepsilon) = \int_0^T r_t D(X_t^\varepsilon) dt + \int_0^T \varepsilon^{\beta P} h_t P(\gamma_t^\varepsilon) d\varphi_t^\varepsilon.$$

and

$$X_t^\varepsilon = -X_t^o + \int_0^t \gamma_s^\varepsilon d\varphi_s^\varepsilon.$$

We have the following theorem.

Theorem 3.3.3 (Asymptotic lower bound for singular control). *Let $I = I(a, r, h)$ be given by*

$$I(a, r, h) = \inf_{(\mu, \rho)} \int_{\mathbb{R}_x^d} r D(x) \mu(dx) + \int_{\mathbb{R}_x^d \times \Delta \times \mathbb{R}_\delta^+} h P(\gamma) \rho(dx \times d\gamma \times d\delta),$$

with $(\mu, \rho) \in \mathcal{P}(\mathbb{R}_x^d) \times \mathcal{M}(\mathbb{R}_x^d \times \Delta \times \mathbb{R}_\delta^+)$ satisfying the following constraint

$$\int_{\mathbb{R}_x^d} A^a f(x) \mu(dx) + \int_{\mathbb{R}_x^d \times \Delta \times \mathbb{R}_\delta^+} Bf(x, \gamma, \delta) \rho(dx \times d\gamma \times d\delta) = 0, \quad \forall f \in C_0^2(\mathbb{R}_x^d),$$

where

$$A^a f(x) = \frac{1}{2} \sum_{i,j} a_{ij} \partial_{ij}^2 f(x), \quad Bf(x, \gamma, \delta) = \begin{cases} \langle \gamma, \nabla f(x) \rangle, & \delta = 0, \\ \delta^{-1} (f(x + \delta\gamma) - f(x)), & \delta > 0. \end{cases}$$

Assume that $I(a, r, h)$ is measurable, that the parameters (r_t) and (h_t) are continuous and positive, that Assumption 3.2.2 holds true and that Assumption 3.2.4 is satisfied for some $\beta > 0$. Then,

$$\liminf_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon^{\beta \zeta_D}} J^\varepsilon(\gamma^\varepsilon, \varphi^\varepsilon) \geq_p \int_0^T I(a_t, r_t, h_t) dt.$$

See Example 3.4.4 for a closed form solution of I .

3.3.4 Regular control

Consider

$$\inf_{u^\varepsilon \in \mathcal{A}} J^\varepsilon(u^\varepsilon),$$

with

$$J^\varepsilon(u^\varepsilon) = \int_0^T (r_t D(X_t^\varepsilon) + \varepsilon^{\beta_Q} l_t Q(u_t^\varepsilon)) dt,$$

and

$$X_t^\varepsilon = -X_t^\circ + \int_0^t u_s^\varepsilon ds.$$

We have the following theorem.

Theorem 3.3.4 (Asymptotic lower bound for regular control). *Let $I = I(a, r, l)$ be given by*

$$I(a, r, l) = \inf_{\mu} \int_{\mathbb{R}_x^d \times \mathbb{R}_u^d} (rD(x) + lQ(u)) \mu(dx, du),$$

with $\mu \in \mathcal{P}(\mathbb{R}_x^d \times \mathbb{R}_u^d)$ satisfying the following constraint

$$\int_{\mathbb{R}_x^d \times \mathbb{R}_u^d} A^a f(x, u) \mu(dx, du) = 0, \quad \forall f \in C_0^2(\mathbb{R}_x^d),$$

where

$$A^a f(x, u) = \frac{1}{2} \sum_{i,j} a_{ij} \partial_{ij}^2 f(x) + \langle u, \nabla f(x) \rangle.$$

Assume that $I(a, r, l)$ is measurable, that the parameters (r_t) and (l_t) are continuous and positive on $[0, T]$, that Assumption 3.2.2 holds true and that Assumption 3.2.4 is satisfied for some $\beta > 0$. Then,

$$\liminf_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon^{\beta_{\zeta_D}}} J^\varepsilon(u^\varepsilon) \geq_p \int_0^T I(a_t, r_t, l_t) dt.$$

See Example 3.4.3 for a closed form solution of I .

Remark 3.3.1 (Upper bound). *It is natural to wonder whether the lower bounds in our theorems are tight and if it is the case, what are the strategies that attain them. In next chapter, we show that for the examples provided in Section 3.4, there are closed form strategies attaining asymptotically the lower bounds. For instance, in the case of combined regular and impulse control, it means that there exist $(u^{\varepsilon,*}, \tau^{\varepsilon,*}, \xi^{\varepsilon,*}) \in \mathcal{A}$ such that*

$$\lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon^{\beta_{\zeta_D}}} J^\varepsilon(u^{\varepsilon,*}, \tau^{\varepsilon,*}, \xi^{\varepsilon,*}) \rightarrow_p \int_0^T I(a_t, r_t, l_t, k_t, h_t) dt.$$

These optimal strategies are essentially time-varying versions of the optimal strategies for the time-average control of Brownian motion.

3.4 Interpretation of lower bounds and examples

Our goal in this section is to provide a probabilistic interpretation of the lower bounds in Theorems 3.2.1, 3.3.1, 3.3.2, 3.3.3 and 3.3.4, which are expressed in terms of linear programming. In particular, we want to connect them with the time-average control problem of Brownian motion. To our knowledge, there is no general result available for the equivalence between time-average control problem and linear programming. Partial results exist in [BG88, KS98, KS99, KS15, HSZ14] but do not cover all the cases we need. Here we provide a brief self-contained study enabling us to also treat the cases of singular/impulse controls and their combinations with

regular control. We first introduce controlled martingale problems and show that they can be seen as a relaxed version of the controlled Brownian motion (3.1.1-local). Then we formulate the time-average control problem in this martingale framework. We finally show that this problem has an equivalent description in terms of infinite dimensional linear program. While essential ingredients and arguments for obtaining these results are borrowed from [KS98] and [KS01], we provide sharp conditions which guarantee the equivalence of these two formulations.

3.4.1 Martingale problem associated to controlled Brownian motion

In [ST13, PST15, MMKS14, AMKS15], the authors obtain a HJB equation in the first order expansion for the value function of the utility maximization problem under transaction costs, which essentially provides a lower bound for their control problems. They mention a connection between the HJB equation and the time-average control problem of Brownian motion, see also [Hyn12]. Here we wish to rigorously establish an equivalence between the linear programs in our lower bounds and the time-average control of Brownian motion. This leads us to introduce a relaxed version for the controlled Brownian motion. We shall see that the optimal costs for all these formulations coincide in the examples provided in the next section.

We place ourselves in the setting of [KS01], from which we borrow and rephrase several elements, and assume that the state space E and control spaces U and V are complete, separable, metric spaces. Consider an operator $A : \mathcal{D} \subset C_b(E) \rightarrow C(E \times U)$ and an operator $B : \mathcal{D} \subset C_b(E) \rightarrow C(E \times V)$.

Definition 3.4.1 (Controlled martingale problem). *A triplet (X, Λ, Γ) with (X, Λ) an $E \times \mathcal{P}(U)$ -valued process and Γ an $\mathcal{L}(E \times V)$ -valued random variable is a solution of the controlled martingale problem for (A, B) with initial distribution $\nu_0 \in \mathcal{P}(E)$ if there exists a filtration (\mathcal{F}_t) such that the process (X, Λ, Γ_t) is \mathcal{F}_t -progressive, X_0 has distribution ν_0 and for every $f \in \mathcal{D}$,*

$$f(X_t) - \int_0^t \int_U Af(X_s, u) \Lambda_s(du) ds - \int_{E \times V \times [0, t]} Bf(x, v) \Gamma(dx \times dv \times ds) \quad (3.4.1)$$

is an \mathcal{F}_t -martingale.

We now consider two specific cases for the operators A and B , which will be relevant in order to express our lower bounds. Furthermore, we explain why these specific choices of A and B are connected to combined regular and singular/impulse control of Brownian motion.

Example 3.4.1 (Combined regular and impulse control of Brownian motion). *Let $\mathcal{D} = C_0^2(\mathbb{R}^d) \oplus \mathbb{R}$ and define $A : \mathcal{D} \rightarrow C(E \times U)$ and $B : \mathcal{D} \rightarrow C(E \times V)$ by*

$$Af(x, u) = \frac{1}{2} \sum_{i,j} a_{ij} \partial_{ij}^2 f(x) + \langle u, \nabla f(x) \rangle, \quad (3.4.2)$$

$$Bf(x, \xi) = f(x + \xi) - f(x). \quad (3.4.3)$$

Here $E = \mathbb{R}_x^d$, $U = \mathbb{R}_u^d$ and $V = \mathbb{R}_\xi^d \setminus \{0_\xi\}$. We call any solution of this martingale problem the combined regular and impulse control of Brownian motion.

Indeed, consider the following process

$$X_t = X_0 + \sqrt{a} W_t + \int_0^t u_s ds + \sum_{0 < \tau_j \leq t} \xi_j,$$

with (u_t) a progressively measurable process, (τ_j) a sequence of stopping times and (ξ_j) a sequence of \mathcal{F}_{τ_j} -measurable random variables. Define

$$N_t = \sum_j \mathbb{1}_{\{\tau_j \leq t\}}, \quad \xi_t = \xi_j, \quad t \in (\tau_{j-1}, \tau_j].$$

Then for any $f \in \mathcal{D}$, by Itô's formula,

$$\begin{aligned} f(X_t) - \int_0^t Af(X_s, u_s)ds - \int_0^t Bf(X_{s-}, \xi_{s-})dN_s \\ = f(X_0) - \int_0^t \nabla f(X_s)^T \sqrt{a}dW_s, \end{aligned}$$

which is a martingale. Let

$$\Lambda_t = \delta_{u_t}(du), \quad \Gamma(H \times [0, t]) = \int_0^t \mathbb{1}_H(X_{s-}, \xi_{s-})dN_s, \quad H \in \mathcal{B}(\mathbb{R}_x^d \times \mathbb{R}_\xi^d \setminus \{0_\xi\}).$$

Then (X, Λ, Γ) solves the martingale problem (A, B) with initial distribution $\mathcal{L}(X_0)$.

Example 3.4.2 (Combined regular and singular control of Brownian motion). Take $\mathcal{D} = C_0^2(\mathbb{R}^d) \oplus \mathbb{R}$ and define

$$Af(x, u) = \frac{1}{2} \sum_{i,j} a_{ij} \partial_{ij}^2 f(x) + \langle u, \nabla f(x) \rangle, \quad (3.4.4)$$

$$Bf(x, \gamma, \delta) = \begin{cases} \langle \gamma, \nabla f(x) \rangle, & \delta = 0 \\ \delta^{-1}(f(x + \delta\gamma) - f(x)), & \delta > 0. \end{cases} \quad (3.4.5)$$

Here $E = \mathbb{R}_x^d$, $U = \mathbb{R}_u^d$ and $V = \Delta \times \mathbb{R}_\delta^+$. Any solution of this martingale problem is called combined regular and singular control of Brownian motion.

Indeed, let X be given by

$$X_t = X_0 + \sqrt{a}W_t + \int_0^t u_s ds + \int_0^t \gamma_s d\varphi_s,$$

with u a progressively measurable process, $\gamma_s \in \Delta$ and φ_s non-decreasing. By Itô's formula, we have

$$\begin{aligned} f(X_t) - \int_0^t Af(X_s, u_s)ds - \int_0^t Bf(X_{s-}, \gamma_s, \delta\varphi_s)d\varphi_s \\ = f(X_0) - \int_0^t \nabla f(X_s)^T \sqrt{a}dW_s, \end{aligned}$$

which is a martingale for any $f \in \mathcal{D}$. Let

$$\Lambda_t = \delta_{u_t}(du), \quad \Gamma(H \times [0, t]) = \int_0^t \mathbb{1}_H(X_{s-}, \gamma_s, \delta\varphi_s)d\varphi_s, \quad H \in \mathcal{B}(\mathbb{R}_x^d \times \Delta \times \mathbb{R}_\delta^+).$$

Then (X, Λ, Γ) solves the martingale problem (A, B) with initial distribution $\mathcal{L}(X_0)$.

3.4.2 Time-average control of Brownian motion

Now we formulate a relaxed version of the time-average control problem of Brownian motion in terms of a controlled martingale problem. This generalizes [KS98, HSZ14] to combined regular and singular/impulse control of martingale problems, see also [KS15]. Recall that A and B are two operators where $A : \mathcal{D} \subset C_b(E) \rightarrow C(E \times U)$ and $B : \mathcal{D} \subset C_b(E) \rightarrow C(E \times V)$. Consider two cost functionals $C_A : E \times U \rightarrow \mathbb{R}_+$ and $C_B : E \times V \rightarrow \mathbb{R}_+$.

Definition 3.4.2 (Martingale formulation of time-average control problem). *The time-average control problem under the martingale formulation is given by*

$$I^M = \inf_{(X, \Lambda, \Gamma)} \limsup_{t \rightarrow \infty} \frac{1}{t} \mathbb{E} \left[\int_0^t \int_U C_A(X_s, u) \Lambda_s(du) ds + \int_{E \times V \times [0, t]} C_B(x, v) \Gamma(dx \times dv \times ds) \right], \quad (3.4.6)$$

where the inf is taken over all solutions of the martingale problem (A, B) with any initial distribution $\nu_0 \in \mathcal{P}(E)$.

Now, let (X, Λ, Γ) be any solution of the martingale problem with operators A and B . Define $(\mu_t, \rho_t) \in \mathcal{P}(E \times U) \times \mathcal{M}(E \times V)$ as

$$\mu_t(H_1) = \frac{1}{t} \mathbb{E} \left[\int_0^t \int_U \mathbb{1}_{H_1}(X_s, u) \Lambda_s(du) ds \right], \quad (3.4.7)$$

$$\rho_t(H_2) = \frac{1}{t} \mathbb{E} [\Gamma(H_2 \times [0, t])], \quad (3.4.8)$$

for $H_1 \in \mathcal{B}(E \times U)$ and $H_2 \in \mathcal{B}(E \times V)$. Then the average cost up to time t in (3.4.6) can be expressed as

$$\int_{E \times U} C_A(x, u) \mu_t(dx \times du) + \int_{E \times V} C_B(x, v) \rho_t(dx \times dv).$$

On the other hand, for $f \in \mathcal{D}$, (3.4.1) defines a martingale. Taking the expectation, we obtain

$$\frac{1}{t} \mathbb{E} \left[\int_0^t \int_U Af(X_s, u) \Lambda_s(du) ds + \int_{E \times V \times [0, t]} Bf(x, v) \Gamma(dx \times dv \times ds) \right] = \frac{1}{t} \mathbb{E} [f(X_t) - f(X_0)].$$

If X_t is stationary, we have

$$\int_{E \times U} Af(x, u) \mu_t(dx \times du) + \int_{E \times V} Bf(x, v) \rho_t(dx \times dv) = 0.$$

Letting t tend to infinity, this leads us to introduce the following linear programming problem.

Definition 3.4.3 (Linear programming (LP) formulation of time-average control). *The time-average control problem under the LP formulation is given by*

$$I^P = \inf_{(\mu, \rho)} c(\mu, \rho), \quad (3.4.9)$$

with

$$c : \mathcal{P}(E \times U) \times \mathcal{M}(E \times V) \rightarrow \mathbb{R}_+ \\ (\mu, \rho) \mapsto \int_{E \times U} C_A(x, u) \mu(dx \times du) + \int_{E \times V} C_B(x, v) \rho(dx \times dv), \quad (3.4.10)$$

where the inf is computed over all $\mu \in \mathcal{P}(E \times U)$ and $\rho \in \mathcal{M}(E \times V)$ satisfying the constraint

$$\int_{E \times U} Af(x, u) \mu(dx \times du) + \int_{E \times V} Bf(x, v) \rho(dx \times dv) = 0, \quad \forall f \in \mathcal{D}. \quad (3.4.11)$$

We now present the theorem which connects linear programming and time-average control of Brownian motion.

Theorem 3.4.1 (Equivalence between I^M and I^P). *Assume that*

1. (Condition on the operators A and B) *The operators A and B satisfy Condition 1.2 in [KS01]. In particular, there exist $\psi_A \in C(E \times U)$, $\psi_B \in C(E \times V)$, and constants a_f, b_f depending on $f \in \mathcal{D}$ such that*

$$|Af(x, u)| \leq a_f \psi_A(x, u), \quad |Bf(x, v)| \leq b_f \psi_B(x, v), \quad \forall x \in E, u \in U, v \in V. \quad (3.4.12)$$

2. (Condition on the cost function C_A) *The cost function C_A is non-negative and inf-compact, that is $\{(x, u) \in E \times U \mid C_A(x, u) \leq c\}$ is a compact set for each $c \in \mathbb{R}_+$. In particular, C_A is lower semi-continuous.*
3. (Condition on cost function C_B) *The cost function C_B is non-negative and lower semi-continuous. Moreover, C_B satisfies*

$$\inf_{(x, v) \in E \times V} C_B(x, v) > 0. \quad (3.4.13)$$

4. (Relation between operators and cost functions) There exist constants θ and $0 < \beta < 1$ such that

$$\psi_A(x, u)^{1/\beta} \leq \theta(1 + C_A(x, u)), \quad \psi_B(x, v)^{1/\beta} \leq \theta C_B(x, v), \quad (3.4.14)$$

for ψ_A and ψ_B given by (3.4.12).

Then the two formulations above of the time-average control problem are equivalent in the sense that

$$I^M = I^P.$$

We therefore write I for both I^M and I^P .

Proof. The idea of the proof is the same as in [KS98], with a key ingredient provided by [KS01, Theorem 1.7].

We first show that $I^P \leq I^M$. Given any solution (X, Λ, Γ) of the martingale problem, we define the occupation measures as (3.4.7) and (3.4.8). Without loss of generality, we can assume that $\limsup_{t \rightarrow \infty} c(\mu_t, \rho_t) < \infty$ where c is defined in (3.4.10) (otherwise we would have $I^M = \infty \geq I^P$). We will show that $I^P \leq \limsup_{t \rightarrow \infty} c(\mu_t, \rho_t)$.

We consider the one-point compactification $\overline{E \times V} = E \times V \cup \{\infty = (x_\infty, v_\infty)\}$ and extend C_B to $\overline{E \times V}$ by

$$C_B(x_\infty, v_\infty) = \liminf_{(x, v) \rightarrow (x_\infty, v_\infty)} C_B(x, v) > 0,$$

where the last inequality is guaranteed by (3.4.13). Since C_B is lower semi-continuous, the level sets $\{(x, v) \in \overline{E \times V} \mid C_B(x, v) \leq c\}$ are compact. By Lemma 3.C.1, we deduce that c is a tightness function on $\mathcal{P}(E \times U) \times \mathcal{M}(\overline{E \times V})$. So the family of occupation measures $(\mu_t, \rho_t)_{t \geq 0}$ is tight if ρ_t is viewed as a measure on $\overline{E \times V}$. It follows that the family of occupation measures indexed by t is relatively compact.

Let $(\mu, \bar{\rho}) \in \mathcal{P}(E \times U) \times \mathcal{M}(\overline{E \times V})$ be any limit point of (μ_t, ρ_t) with canonical decomposition $\bar{\rho} = \rho + \theta_{\bar{\rho}} \delta_\infty$. We claim that (μ, ρ) satisfies the linear constraint (3.4.11). Indeed, by (3.4.12) and (3.4.14), we have

$$|Af|^{1/\beta} \leq \theta_f(1 + C_A), \quad |Bf|^{1/\beta} \leq \theta_f C_B,$$

where θ_f is a non-negative real number depending on f . Then $\sup_t c(\mu_t, \rho_t) < \infty$ implies that Af and Bf are uniformly integrable with respect to μ and $\bar{\rho}$. We therefore have

$$\int_{E \times U} Af d\mu + \int_{\overline{E \times V}} \mathbb{1}_{E \times V} Bf d\bar{\rho} = \lim_{t \rightarrow \infty} \int_{E \times U} Af d\mu_t + \int_{\overline{E \times V}} \mathbb{1}_{E \times V} Bf d\rho_t.$$

The right hand side being equal to zero, we obtain

$$\int_{E \times U} Af d\mu + \int_{E \times V} Bf d\rho = 0.$$

Since C_A and C_B are lower semi-continuous, it follows that (see [DE11, Theorem A.3.12])

$$I^P \leq c(\mu, \rho) \leq c(\mu, \bar{\rho}) \leq \liminf_{t \rightarrow \infty} c(\mu_t, \rho_t) \leq \limsup_{t \rightarrow \infty} c(\mu_t, \rho_t).$$

As the choice of the solution of the controlled martingale problem (X, Λ, Γ) is arbitrary, we conclude that $I^P \leq I^M$.

We now show that $I^M \leq I^P$. Given any (μ, ρ) satisfying the linear constraint (3.4.11) such that $c(\mu, \rho) < \infty$, Theorem 1.7 in [KS01] together with the condition on the operators A and B provides the existence of a stationary solution (X, Λ, Γ) for the martingale problem (A, B) with marginals (μ, ρ) , hence $I^M \leq I^P$. \square

Remark 3.4.1. *Compared with the results in [KS98, BG88], no near-monotone condition is necessary for the singular component. This is due to the fact that in the linear constraint (3.4.11), ρ belongs to $\mathcal{M}(E \times V)$ instead of $\mathcal{P}(E \times V)$.*

We now give natural examples for which $I^P = I^M$.

Corollary 3.4.1 (Time-average control of Brownian motion with quadratic costs). *Assume that C_A is given by*

$$C_A(x, u) = r\langle x, \Sigma^D x \rangle + l\langle u, \Sigma^Q u \rangle, \quad \forall (x, u) \in \mathbb{R}_x^d \times \mathbb{R}_u^d, \quad (3.4.15)$$

where $\Sigma^D, \Sigma^Q \in \mathcal{S}_d^+$. Then for both following cases, we have $I^M = I^P$:

1. The operators A and B are given by (3.4.2)-(3.4.3) and the cost function C_B is given by

$$C_B(x, \xi) = k \sum_{i=1}^d F_i \mathbb{1}_{\{\xi^i \neq 0\}} + h\langle P, |\xi| \rangle, \quad (x, \xi) \in \mathbb{R}_x^d \times \mathbb{R}_\xi^d \setminus \{0_\xi\}, \quad (3.4.16)$$

with $\min_i F_i > 0$.

2. The operators A and B are given by (3.4.4)-(3.4.5) and the cost function C_B is given by

$$C_A(x, \gamma, \delta) = h\langle P, |\gamma| \rangle, \quad (x, \gamma, \delta) \in \mathbb{R}_x^d \times \Delta \times \mathbb{R}_\delta^+, \quad (3.4.17)$$

with $\min_i P_i > 0$.

Proof. We consider the impulse case. First, we show that A and B satisfy [KS01, Condition 1.2]. (i) It is clear that $\mathbb{1}_{\mathbb{R}^d} \in \mathcal{D}$, and $A\mathbb{1}_{\mathbb{R}^d} = 0, B\mathbb{1}_{\mathbb{R}^d} = 0$. (ii) Define

$$\psi_A(x, u) = 1 \vee \sum_{i=1}^d |u_i|, \quad \psi_B(x, \xi) = 1,$$

then (3.4.12) is satisfied. (iii) Since $C_0^2(\mathbb{R}_x^d)$ equipped with $\|\cdot\|_{C_0^2}$ is separable, the third condition is satisfied. (iv) $A_u f(x) = Af(x, u)$ and $B_\xi f(x) = Bf(x, \xi)$ satisfy the positive maximum principle, so they are dissipative. It is obvious that they verify [KS01, (1.10)]. Hence they are pre-generators. (v) Obvious. Second, since C_A and C_B are l.s.c. and $\min_i F_i > 0$, the conditions on C_A and C_B are verified. Third, (3.4.14) holds with $\beta = 1/2$.

The proof for the singular case is similar. Note that $\min_i P_i > 0$ is necessary for (3.4.13) and the second bound in (3.4.14) to hold. \square

3.4.3 Explicit examples in dimension one

We collect here a comprehensive list of examples in dimension one for which explicit solutions are available. Most of these results exist already in the literature under the classical SDE formulation (see for example [DY13, JZ06c, JZ06a, JZ06b, ST13, AMKS15, MMKS14, GL14a, HSZ14]), but Examples 3.4.6 and 3.4.7 are presented here for the first time. The basic idea is to solve explicitly the HJB equation corresponding to the time-average control problem and apply a verification theorem. Similar methods apply under the linear programming framework. However, for completeness we provide in Section 3.7 detailed proofs tailored to the formulation in terms of linear programming. In fact, we prove only the case of combined regular and impulse control, that is Example 3.4.6. The proofs for the other examples are similar and hence omitted.

Example 3.4.3 (Regular control of Brownian motion). *Let $r, l > 0$ and consider the following linear programming problem*

$$I(a, r, l) = \inf_{\mu} \int_{\mathbb{R}_x \times \mathbb{R}_u} (rx^2 + lu^2) \mu(dx, du), \quad (3.4.18)$$

where $\mu \in \mathcal{P}(\mathbb{R}_x \times \mathbb{R}_u)$ satisfies

$$\int_{\mathbb{R}_x \times \mathbb{R}_u} \left(\frac{1}{2} a f''(x) + u f'(x) \right) \mu(dx, du) = 0, \quad \forall f \in C_0^2(\mathbb{R}_x). \quad (3.4.19)$$

By Corollary 3.4.1, this is equivalent to the time-average control of Brownian motion with quadratic costs in the sense of Definition 3.4.2.

Let us explain heuristically how to obtain the optimal solution (a rigorous verification argument for the linear programming formulation is provided in Section 3.7). Roughly speaking, Definition 3.4.2 describes the following dynamics

$$dX_t^u = \sqrt{a} dW_t + u_t dt.$$

The optimization objective is

$$\inf_{(u_t) \in \mathcal{A}} \limsup_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left[\int_0^T (r(X_t^u)^2 + l u_t^2) dt \right],$$

where the set \mathcal{A} of admissible controls contains all progressively measurable processes (u_t) such that

$$\limsup_{T \rightarrow \infty} \mathbb{E}[(X_T^u)^2] < \infty$$

Consider the associated HJB equation

$$\inf_{u \in \mathbb{R}_u} \frac{1}{2} a w''(x) + u w'(x) + l u^2 + r x^2 - I^V = 0,$$

where the constant I^V must be found as part of the solution. It is easy to find the explicit solution (see also [MMKS14, Equation (3.18)]):

$$w(x) = \sqrt{r} l x^2, \quad I^V = \sqrt{a^2 r l}.$$

Now, let (u_t) be an admissible control and apply Itô's formula to $w(X_T)$:

$$w(X_T) = w(X_0) + \int_0^T \left(\frac{1}{2} a w''(X_t) + u_t w'(X_t) \right) dt + \int_0^T w'(X_t) dW_t.$$

It follows that

$$\begin{aligned} \int_0^T (r X_t^2 + l u_t^2) dt &\geq \int_0^T \left(I^V - \frac{1}{2} a w''(X_t) - u_t w'(X_t) \right) dt \\ &= T I^V + w(X_0) - w(X_T) + \int_0^T w'(X_t) dW_t. \end{aligned} \quad (3.4.20)$$

Taking expectation, dividing by T on both sides, and using the admissibility conditions, we obtain

$$\limsup_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left[\int_0^T (r X_t^2 + l u_t^2) dt \right] \geq I^V.$$

To show that I^V is indeed the optimal cost, it is enough to show that equality holds in (3.4.20) for the optimal feedback control given by

$$u^*(x) = -\frac{w'(x)}{2l} = -\theta x, \quad \theta = \sqrt{\frac{r}{l}}.$$

Therefore, the optimally controlled process is an Ornstein-Uhlenbeck process

$$dX^* = \sqrt{a} dW_t - \theta X_t^* dt.$$

Naturally, the stationary distribution of $(X^*, u^*(X_t^*))$ is the solution μ^* of the linear programming problem. We have the following result.

Proposition 3.4.1. *The solution of (3.4.18)-(3.4.19) is given by*

$$I(a, r, l) = \sqrt{a^2 r l},$$

and the optimum is attained by

$$\mu^*(dx, du) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{x^2}{2\sigma^2}} dx \otimes \delta_{\{-\theta x\}}(du),$$

where

$$\sigma^2 = \frac{a}{2\theta}, \quad \theta = \sqrt{\frac{r}{l}}.$$

Example 3.4.4 (Singular control of Brownian motion). *For any parameters $r, h > 0$, consider the following linear programming problem*

$$I(a, r, h) = \inf_{(\mu, \rho)} \int_{\mathbb{R}_x} r x^2 \mu(dx) + \int_{\mathbb{R}_x \times \{\pm 1\} \times \mathbb{R}_\delta^+} h |\gamma| \rho(dx, d\gamma, d\delta), \quad (3.4.21)$$

where $\mu \in \mathcal{P}(\mathbb{R}_x)$ and $\rho \in \mathcal{M}(\mathbb{R}_x \times \{\pm 1\} \times \mathbb{R}_\delta^+)$ satisfy

$$\int_{\mathbb{R}_x} \frac{1}{2} a f''(x) \mu(dx) + \int_{\mathbb{R}_x \times \{\pm 1\} \times \mathbb{R}_\delta^+} \gamma f'(x) \rho(dx, d\gamma, d\delta) = 0, \quad \forall f \in C_0^2(\mathbb{R}_x). \quad (3.4.22)$$

By Corollary 3.4.1, this is equivalent to the time-average control of Brownian motion with quadratic deviation penalty and proportional costs in the sense of Definition 3.4.2.

The dynamics of the solution of the controlled martingale problem is heuristically

$$dX_t = \sqrt{a} dW_t + \gamma_t d\varphi_t,$$

where $\gamma_t \in \{\pm 1\}$ and φ_t is a non-decreasing process. The optimization objective is

$$\inf_{(\gamma_t, \varphi_t) \in \mathcal{A}} \limsup_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left[\int_0^T r X_t^2 dt + h \varphi_T \right].$$

The associated HJB equation is

$$\left(\frac{1}{2} a w''(x) + r x^2 - I^V \right) \wedge \left(\inf_{\gamma \in \{\pm 1\}} \gamma w'(x) + h \right) = 0.$$

An explicit solution for w is provided in [ST13] (see also [JZ06a, Kar83, DY13]):

$$w(x) = \begin{cases} Ax^4 + Bx^2, & -U \leq x \leq U, \\ w(-U) + h(-U - x), & x < -U, \\ w(U) + h(x - U), & x > U, \end{cases} \quad (3.4.23)$$

with

$$A = -\frac{1}{6} \frac{r}{a}, \quad B = \frac{I}{a},$$

and

$$I = \left(\frac{3}{4} a r^{1/2} h \right)^{2/3}, \quad U = \left(\frac{3}{4} a r^{-1} h \right)^{1/3}.$$

The optimally controlled process is

$$dX_t^* = \sqrt{a} dW_t + d\varphi_t^- - d\varphi_t^+,$$

where φ^\pm are the local times keeping $X_t^* \in [-U, U]$ such that

$$\int_0^t \mathbb{1}_{\{X_s^* \neq -U\}} d\varphi_s^- + \int_0^t \mathbb{1}_{\{X_s^* \neq U\}} d\varphi_s^+ = 0.$$

In other words, this is a Brownian motion with reflection on the interval $[-U, U]$. The optimal solution (μ^*, ρ^*) is the stationary distribution of X_t^* and the limit of boundary measures

$$\frac{1}{t} \int_0^t (\delta_{(U, -1, 0)} d\varphi_s^+ + \delta_{(-U, 1, 0)} d\varphi_s^-),$$

as $t \rightarrow \infty$. We have the following result.

Proposition 3.4.2. *The solution of (3.4.21)-(3.4.22) is given by*

$$I(a, r, h) = \left(\frac{3}{4}ar^{1/2}h\right)^{2/3}.$$

and the optimum is attained by

$$\mu^*(dx) = \frac{1}{2U} \mathbb{1}_{[-U, U]}(x) dx, \quad (3.4.24)$$

$$\rho^*(dx, d\gamma, d\delta) = \frac{a}{2U} \left(\frac{1}{2} \delta_{(-U, 1, 0)} + \frac{1}{2} \delta_{(U, -1, 0)} \right), \quad (3.4.25)$$

where

$$U = \left(\frac{3}{4}ar^{-1}h\right)^{1/3}.$$

Example 3.4.5 (Impulse control of Brownian motion). *For any parameters $r, k > 0$ and $h \geq 0$, consider the following linear programming problem*

$$I(a, r, k, h) = \inf_{(\mu, \rho)} \int_{\mathbb{R}_x} rx^2 \mu(dx) + \int_{\mathbb{R}_x \times \mathbb{R}_\xi \setminus \{0_\xi\}} (k + h|\xi|) \rho(dx, d\xi), \quad (3.4.26)$$

where $\mu \in \mathcal{P}(\mathbb{R}_x)$ and $\rho \in \mathcal{M}(\mathbb{R}_x \times \mathbb{R}_\xi \setminus \{0_\xi\})$ satisfy

$$\int_{\mathbb{R}_x} \frac{1}{2} a f''(x) \mu(dx) + \int_{\mathbb{R}_x \times \mathbb{R}_\xi \setminus \{0_\xi\}} (f(x + \xi) - f(x)) \rho(dx, d\xi) = 0, \quad \forall f \in C_0^2(\mathbb{R}). \quad (3.4.27)$$

By Corollary 3.4.1, this is equivalent to the time-average control problem of Brownian motion in the sense of Definition 3.4.2.

The associated HJB equation is (see also [DY13, JZ06c, AMKS15, GL14a])

$$\left(\frac{1}{2}aw''(x) + rx^2 - I^V\right) \wedge \left(\inf_{\xi \in \mathbb{R}_\xi \setminus \{0_\xi\}} (w(x + \xi) + k + h|\xi|) - w(x)\right) = 0.$$

Let θ_1, θ_2 and $0 < \tilde{x}^* < x^*$ be solutions of the following polynomial system

$$\begin{cases} 6a\theta_1 + r = 0, \\ P(x^*) - P(\tilde{x}^*) = k + h(x^* - \tilde{x}^*), \\ P'(x^*) = h, \quad P'(\tilde{x}^*) = h, \end{cases}$$

where $P(x) = \theta_1 x^4 + \theta_2 x^2$. Let $U = x^*$. We can show that the solution of the HJB equation is given by

$$w(x) = \begin{cases} P(x), & |x| \leq U, \\ w(U) + h(|x| - U), & |x| > U, \end{cases}$$

and

$$I^V = a\theta_2.$$

Let $\xi^* = x^* - \tilde{x}^*$. The optimally controlled process is a Brownian motion on the interval $[-U, U]$, which jumps to $\pm U \mp \xi^*$ when reaching the boundary point $\pm U$. Such processes have been studied in [BAP09, GK02]. We have the following result.

Proposition 3.4.3. *The solution of (3.4.26)-(3.4.27) is given by*

$$I(a, r, k, h) = a\theta_2,$$

and the optimum is attained by

$$\mu^* = p(x)dx, \quad p(x) = \begin{cases} \frac{1}{(x^*)^2 - (\tilde{x}^*)^2}(x + x^*), & -x^* \leq x < -\tilde{x}^*, \\ \frac{1}{x^* + \tilde{x}^*}, & -\tilde{x}^* \leq x \leq \tilde{x}^*, \\ \frac{1}{(x^*)^2 - (\tilde{x}^*)^2}(x^* - x), & \tilde{x}^* < x \leq x^*, \end{cases} \quad (3.4.28)$$

$$\rho^* = \frac{a}{(x^*)^2 - (\tilde{x}^*)^2} \left(\frac{1}{2} \delta_{(x^*, \tilde{x}^* - x^*)} + \frac{1}{2} \delta_{(-x^*, -\tilde{x}^* + x^*)} \right). \quad (3.4.29)$$

which correspond to the stationary distribution and boundary measure of Brownian motion with jumps from the boundary on the interval $[-U, U]$.

Example 3.4.6 (Combined regular and impulse control of Brownian motion). *For any parameters $r, l, k > 0$ and $h \geq 0$, consider the following linear programming problem*

$$I(a, r, l, k, h) = \inf_{(\mu, \rho)} \int_{\mathbb{R}_x \times \mathbb{R}_u} (rx^2 + lu^2) \mu(dx, du) + \int_{\mathbb{R}_x \times \mathbb{R}_\xi \setminus \{0_\xi\}} (k + h|\xi|) \rho(dx, d\xi), \quad (3.4.30)$$

where $\mu \in \mathcal{P}(\mathbb{R}_x \times \mathbb{R}_u)$ and $\rho \in \mathcal{M}(\mathbb{R}_x \times \mathbb{R}_\xi \setminus \{0_\xi\})$ satisfy

$$\int_{\mathbb{R}_x \times \mathbb{R}_u} \left(\frac{1}{2} af''(x) + uf'(x) \right) \mu(dx, du) + \int_{\mathbb{R}_x \times \mathbb{R}_\xi \setminus \{0_\xi\}} (f(x + \xi) - f(x)) \rho(dx, d\xi) = 0, \quad \forall f \in C_0^2(\mathbb{R}). \quad (3.4.31)$$

By Corollary 3.4.1, this is equivalent to the time-average control problem of Brownian motion in the sense of Definition 3.4.2.

The corresponding HJB equation is

$$\left(\frac{1}{2} aw''(x) + \inf_u (uw'(x) + lu^2) + rx^2 - I^V \right) \wedge \left(\inf_\xi (w(x + \xi) + k + h|\xi|) - w(x) \right) = 0.$$

In Section 3.7, we show that this equation admits a classical solution

$$w(x) = \begin{cases} (rl)^{1/2} x^2 - 2al \ln {}_1F_1\left(\frac{1-\iota}{4}; \frac{1}{2}; \left(\frac{r}{a^2 l}\right)^{1/2} x^2\right), & |x| \leq U, \\ w(U) + h(|x| - U), & |x| > U, \end{cases} \quad (3.4.32)$$

where ${}_1F_1$ is the Kummer confluent hypergeometric function (see Section 3.A) and ξ^* and U are such that $0 < \xi^* < U$ and

$$w(\pm U \mp \xi^*) + k + h|\xi^*| - w(\pm U) = 0.$$

Moreover,

$$I^V = \iota a \sqrt{rl},$$

for some $\iota \in (0, 1)$.

Let u^* be defined as

$$u^*(x) = \underset{u}{\operatorname{Argmin}} Aw(x, u) + C_A(x, u) = -\frac{w'(x)}{2l}.$$

The optimally controlled process is given by

$$dX_t^* = \sqrt{a} dW_t + u^*(X_t^*) dt + d\left(\sum_{\tau_j \leq t} (\mathbb{1}_{\{X_{\tau_j^-}^* = -U\}} \xi^* - \mathbb{1}_{\{X_{\tau_j^-}^* = U\}} \xi^*) \right).$$

We have the following result.

Proposition 3.4.4. *The solution of (3.4.30)-(3.4.31) is given by*

$$I(a, r, l, k, h) = \iota a \sqrt{rl},$$

and the optimum is attained by (μ^*, ρ^*) where

$$\begin{aligned} \mu^*(dx, du) &= p^*(x)dx \otimes \delta_{u^*(x)}(du), \\ \rho^*(dx, d\xi) &= (\rho_-^* \delta_{(-U, \xi^*)} + \rho_+^* \delta_{(U, -\xi^*)})(dx, d\xi), \end{aligned}$$

with $p^*(x) \in C^0([L, U]) \cap C^2([L, U] \setminus \{L + \xi^*(L), U + \xi^*(U)\})$ solution of

$$\begin{cases} \frac{1}{2}ap''(x) - (u^*(x)p(x))' = 0, & x \in (-U, U) \setminus \{-U + \xi^*, U - \xi^*\}, \\ p(-U) = p(U) = 0, \\ \frac{1}{2}ap'((-U)+) = p'((-U + \xi^*)-) - \frac{1}{2}a(p'((-U + \xi^*)+)), \\ \frac{1}{2}ap'(U-) = p'((U - \xi^*)+) - \frac{1}{2}a(p'((U - \xi^*)-)), \\ \int_{-U}^U p(x) = 1, \end{cases}$$

and $\rho_-^*, \rho_+^* \in \mathbb{R}^+$ given by

$$\rho_-^* = \frac{1}{2}ap'((-U)+), \quad \rho_+^* = -\frac{1}{2}ap'(U-).$$

Note that μ^* and ρ^* are the stationary distribution and boundary measure of the optimally controlled process X_t^* .

Example 3.4.7 (Combined regular and singular control of Brownian motion). *For any parameters $r, l, h > 0$, consider the following linear programming problem*

$$I = \inf_{(\mu, \rho)} \int_{\mathbb{R}_x \times \mathbb{R}_u} (rx^2 + lu^2)\mu(dx, du) + \int_{\mathbb{R}_x \times \{\pm 1\} \times \mathbb{R}_\delta^+} h|\gamma|\rho(dx, d\gamma, d\delta), \quad (3.4.33)$$

where $\mu \in \mathcal{P}(\mathbb{R}_x \times \mathbb{R}_u)$ and $\rho \in \mathcal{M}(\mathbb{R}_x \times \{\pm 1\} \times \mathbb{R}_\delta^+)$ satisfy

$$\int_{\mathbb{R}_x \times \mathbb{R}_u} \left(\frac{1}{2}af''(x) + uf'(x) \right) \mu(dx, du) + \int_{\mathbb{R}_x \times \{\pm 1\} \times \mathbb{R}_\delta^+} \gamma f'(x)(dx, d\gamma, d\delta) = 0, \quad \forall f \in C_0^2(\mathbb{R}_x). \quad (3.4.34)$$

Again by Corollary 3.4.1, this is equivalent to the time-average control problem of Brownian motion in the sense of Definition 3.4.2.

The constant I^V and $w : \mathbb{R} \rightarrow \mathbb{R}$ exist with the same expressions as in Example 3.4.6. Similarly, we have the following result.

Proposition 3.4.5. *The solution of (3.4.33)-(3.4.7) is given by*

$$I = \iota a \sqrt{rl},$$

and the optimum is attained by (μ^*, ρ^*) where

$$\begin{aligned} \mu^*(dx, du) &= p^*(x)dx \otimes \delta_{u^*(x)}(du), \\ \rho^*(dx, d\gamma, d\delta) &= (\rho_-^* \delta_{(-U, 1, 0)} + \rho_+^* \delta_{(U, -1, 0)})(dx, d\gamma, d\delta), \end{aligned}$$

with $p^*(x) \in C^2([-U, U])$ solution of

$$\begin{cases} \frac{1}{2}ap''(x) - (u^*(x)p(x))' = 0, & x \in (-U, U), \\ \frac{1}{2}ap'(x) + u^*(x)p(x) = 0, & x \in \{-U, U\}, \\ \int_{-U}^U p(x) = 1, \end{cases} \quad (3.4.35)$$

and $\rho_-^*, \rho_+^* \in \mathbb{R}^+$ given by

$$\rho_-^* = \frac{1}{2}ap(-U), \quad \rho_+^* = \frac{1}{2}ap(U). \quad (3.4.36)$$

Remark 3.4.2 (Higher dimension examples). *To our knowledge, examples with closed-form solutions for time-average control of Brownian motion in higher dimension are not available except [GL14a, MMKS14, AMKS15, JS10, GW15c].*

3.5 Proof of Theorem 3.2.1

This section is devoted to the proof of Theorem 3.2.1. In Section 3.5.1, we first rigorously establish the arguments outlined in Section 3.2.1, showing that it is enough to consider a small horizon $(t, t + \delta^\varepsilon]$. Then we prove Theorem 3.2.1 in Section 3.5.2. Our proof is inspired by the approaches in [KS99] and [KM93]. An essential ingredient is Lemma 3.5.3, whose proof is given in Section 3.5.3.

3.5.1 Reduction to local time-average control problem

We first show that, to obtain (3.2.13), it is enough to study the local time-average control problem (note that the parameters r_t, l_t, k_t, h_t are frozen at time t)

$$I_t^\varepsilon = \frac{1}{T^\varepsilon} \left(\int_0^{T^\varepsilon} (r_t D(\tilde{X}_s^{\varepsilon,t}) + l_t Q(\tilde{u}_s^{\varepsilon,t})) ds + \sum_{0 < \tilde{\tau}_j^{\varepsilon,t} \leq T^\varepsilon} (k_t F(\tilde{\xi}_j^\varepsilon) + h_t P(\tilde{\xi}_j^\varepsilon)) \right), \quad (3.5.1)$$

where

$$d\tilde{X}_s^{\varepsilon,t} = \tilde{b}_s^{\varepsilon,t} ds + \sqrt{\tilde{a}_s^{\varepsilon,t}} d\tilde{W}_s^{\varepsilon,t} + \tilde{u}_s^{\varepsilon,t} ds + d\left(\sum_{0 < \tilde{\tau}_j^{\varepsilon,t} \leq s} \tilde{\xi}_j^\varepsilon \right), \quad \tilde{X}_0^{\varepsilon,t} = 0 \quad (3.5.2)$$

with $T^\varepsilon = \varepsilon^{-\alpha\beta} \delta^\varepsilon$ and $\delta^\varepsilon \in \mathbb{R}_+$ depending on ε in such a way that

$$\delta^\varepsilon \rightarrow 0, \quad T^\varepsilon \rightarrow \infty, \quad (3.5.3)$$

as $\varepsilon \rightarrow 0$. Recall that $\alpha = 2$, which is due to the scaling property of Brownian motion. We can simply put $\delta^\varepsilon = \varepsilon^\beta$.

Localization Since we are interested in convergence results in probability, under Assumptions 3.2.2 and 3.2.3, it is enough to consider the situation where the following assumption holds.

Assumption 3.5.1. *There exists a positive constant $M \in \mathbb{R}_+^*$ such that*

$$\sup_{(t,\omega) \in [0,T] \times \Omega} |a_t(\omega)| \vee r_t(\omega)^{\pm 1} \vee l_t(\omega)^{\pm 1} \vee h_t(\omega)^{\pm 1} \vee k_t(\omega)^{\pm 1} < M < \infty.$$

Furthermore, X° is a martingale ($b_t \equiv 0$).

Indeed, set $T_m = \inf\{t > 0, \sup_{s \in [0,t]} |b_s| \vee |a_s| \vee r_s^{\pm 1} \vee l_s(\omega)^{\pm 1} \vee h_s^{\pm 1} \vee k_s^{\pm 1} \leq m\}$. Then we have $\lim_{m \rightarrow \infty} \mathbb{P}[T_m = T] = 1$. By standard localization procedure, we can assume that all the parameters are bounded as in Assumption 3.5.1. Let

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \exp \left\{ - \int_0^T a_t^{-1} b_t dW_t - \frac{1}{2} \int_0^T b_t^T a_t^{-2} b_t dt \right\},$$

then by Girsanov theorem, X° is a martingale under \mathbb{Q} . Since \mathbb{Q} is equivalent to \mathbb{P} , we only need to prove (3.2.13) under \mathbb{Q} . Consequently, we can assume that X° is a martingale without loss of generality.

From now on, we will suppose that Assumption 3.5.1 holds.

Locally averaged cost

Lemma 3.5.1. *Under Assumption 3.5.1, we have almost surely*

$$\liminf_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon \zeta_D \beta} J^\varepsilon \geq \liminf_{\varepsilon \rightarrow 0} \int_0^T I_t^\varepsilon dt.$$

Proof. We introduce an auxiliary cost functional:

$$\begin{aligned} \bar{J}^\varepsilon &= \int_0^T \left(\frac{1}{\delta^\varepsilon} \int_t^{(t+\delta^\varepsilon) \wedge T} (r_s D(X_s^\varepsilon) + l_s Q(u_s^\varepsilon)) ds \right. \\ &\quad \left. + \frac{1}{\delta^\varepsilon} \sum_{t < \tau_j^\varepsilon \leq (t+\delta^\varepsilon) \wedge T} (\varepsilon^{\beta_F} k_{\tau_j^\varepsilon} F(\xi_j^\varepsilon) + \varepsilon^{\beta_P} h_{\tau_j^\varepsilon} P(\xi_j^\varepsilon)) \right) dt. \end{aligned}$$

Note that the parameters r, l, k, h inside the integral are not frozen at t . Using Fubini theorem, we have

$$\begin{aligned} \bar{J}^\varepsilon &= \int_0^T \frac{s}{\delta^\varepsilon} \mathbb{1}_{\{0 < s < \delta^\varepsilon\}} (r_s D(X_s^\varepsilon) + l_s Q(u_s^\varepsilon)) ds \\ &\quad + \sum_{0 < \tau_j^\varepsilon \leq T} \frac{\tau_j^\varepsilon}{\delta^\varepsilon} \mathbb{1}_{\{0 < \tau_j^\varepsilon < \delta^\varepsilon\}} (\varepsilon^{\beta_F} k_{\tau_j^\varepsilon} F(\xi_j^\varepsilon) + \varepsilon^{\beta_P} h_{\tau_j^\varepsilon} P(\xi_j^\varepsilon)). \end{aligned}$$

Hence

$$0 \leq \frac{1}{\varepsilon \zeta_D \beta} (J^\varepsilon - \bar{J}^\varepsilon) \leq \delta^\varepsilon I_0^\varepsilon + \mathbb{1}_{\{\tau_0^\varepsilon = 0\}} (\varepsilon^{\beta_F} k_{\tau_0^\varepsilon} F(\xi_1^\varepsilon) + \varepsilon^{\beta_P} h_{\tau_0^\varepsilon} P(\xi_1^\varepsilon)), \quad (3.5.4)$$

where I_0^ε is given by (3.5.1) with $t = 0$.

On the other hand, we have

$$\begin{aligned} \int_0^T I_t^\varepsilon dt &= \frac{1}{\varepsilon \zeta_D \beta} \int_0^T \left(\frac{1}{\delta^\varepsilon} \int_t^{(t+\delta^\varepsilon) \wedge T} (r_t D(X_s^\varepsilon) + l_t Q(u_s^\varepsilon)) ds \right. \\ &\quad \left. + \frac{1}{\delta^\varepsilon} \sum_{t < \tau_j^\varepsilon \leq (t+\delta^\varepsilon) \wedge T} (\varepsilon^{\beta_F} k_t F(\xi_j^\varepsilon) + \varepsilon^{\beta_P} h_t P(\xi_j^\varepsilon)) \right) dt, \end{aligned}$$

where the parameters are frozen at time t . It follows that

$$\left| \frac{1}{\varepsilon \zeta_D \beta} \bar{J}^\varepsilon - \int_0^T I_t^\varepsilon dt \right| \leq M \cdot w(r, l, k, h; \delta^\varepsilon) \cdot \left(\frac{1}{\varepsilon \zeta_D \beta} \bar{J}^\varepsilon \wedge \tilde{J}^\varepsilon \right), \quad (3.5.5)$$

with w the modulus of continuity and M the constant in Assumption 3.5.1. Note that we have $w(r, l, k, h; \delta) \rightarrow 0+$ as $\delta \rightarrow 0+$ by the continuity of r, l, k, h . Combining (3.5.4) and (3.5.5), the inequality follows. \square

Reduction to local problems Using previous lemma, we can reduce the problem to the study of local problems as stated below.

Lemma 3.5.2 (Reduction). *For the proof of Theorem 3.2.1, it is enough to show that*

$$\liminf_{\varepsilon \rightarrow 0} \mathbb{E} [I_t^\varepsilon] \geq \mathbb{E} [I_t]. \quad (3.5.6)$$

Proof. In view of Lemma 3.5.1, to obtain Theorem 3.2.1, we need to prove that

$$\liminf_{\varepsilon \rightarrow 0} \int_0^T I_t^\varepsilon dt \geq_p \int_0^T I_t dt.$$

By Assumption 3.2.1 and Lemma 3.D.1, it is enough to show that

$$\liminf_{\varepsilon \rightarrow 0} \mathbb{E} [Y I_t^\varepsilon] \geq \mathbb{E} [Y I_t],$$

for any bounded random variable Y . Up to a change of notation $r_t \rightarrow Y r_t$, $l_t \rightarrow Y l_t$, $h_t \rightarrow Y h_t$ and $k_t \rightarrow Y k_t$ (note that this is allowed since we do not require r_t, l_t, h_t and k_t to be adapted), it suffices to show (3.5.6). \square

3.5.2 Proof of Theorem 3.2.1

After Section 3.5.1, it suffices to prove (3.5.6) where I_t^ε is given by (3.5.1)-(3.5.2). In particular, we can assume that

$$\sup_{\varepsilon>0} \mathbb{E} [I_t^\varepsilon] < \infty. \quad (3.5.7)$$

Combining ideas from [KM93, KS99], we first consider the empirical occupation measures of $(\tilde{X}_s^{\varepsilon,t})$. Define the following random occupation measures with natural inclusion

$$\begin{aligned} \mu_t^\varepsilon &= \frac{1}{T^\varepsilon} \int_0^{T^\varepsilon} \delta_{\{(\tilde{X}_s^{\varepsilon,t}, \tilde{u}_s^{\varepsilon,t})\}} ds \in \mathcal{P}(\mathbb{R}_x^d \times \mathbb{R}_u^d), \\ \rho_t^\varepsilon &= \frac{1}{T^\varepsilon} \sum_{0 < \tilde{\tau}_j^{\varepsilon,t} \leq T^\varepsilon} \delta_{\{(\tilde{X}_{\tilde{\tau}_j^{\varepsilon,t}}^{\varepsilon,t}, \tilde{\xi}_j^{\varepsilon,t})\}} \in \mathcal{M}(\mathbb{R}_x^d \times \mathbb{R}_\xi^d \setminus \{0_\xi\}) \hookrightarrow \mathcal{M}(\overline{\mathbb{R}_x^d \times \mathbb{R}_\xi^d \setminus \{0_\xi\}}), \end{aligned}$$

where $\overline{E} = E \cup \{\infty\}$ is the one-point compactification of E . Such compactification of state space appears in [BG88]. See also the proof of Corollary 3.4.1 where the compactification of state space is used. Note that for $\bar{m} \in \mathcal{M}(\overline{E})$ we have the canonical decomposition

$$\bar{m}(de) = m(de) + \theta \delta_\infty,$$

with $m \in \mathcal{M}(E)$ and $\theta \in \mathbb{R}^+$. Second, we define $c_t : \Omega \times \mathcal{P}(\mathbb{R}_x^d \times \mathbb{R}_u^d) \times \mathcal{M}(\overline{\mathbb{R}_x^d \times \mathbb{R}_\xi^d \setminus \{0_\xi\}}) \rightarrow \mathbb{R}$,

$$\begin{aligned} (\omega, \mu, \bar{\rho}) &\mapsto \int_{\mathbb{R}_x^d \times \mathbb{R}_u^d} (r_t(\omega)D(x) + l_t(\omega)Q(u))\mu(dx \times du) \\ &\quad + \int_{\overline{\mathbb{R}_x^d \times \mathbb{R}_\xi^d \setminus \{0_\xi\}}} (k_t(\omega)F(\xi) + h_t(\omega)P(\xi))\bar{\rho}(dx \times d\xi), \end{aligned}$$

where the cost functions F and P are extended to $\overline{\mathbb{R}_x^d \times \mathbb{R}_\xi^d \setminus \{0_\xi\}}$ by setting

$$F(\infty_{(x,\xi)}) = \inf_{\xi \in \mathbb{R}_\xi^d \setminus \{0_\xi\}} F(\xi) > 0, \quad P(\infty_{(x,\xi)}) = 0, \quad (3.5.8)$$

with $\infty_{(x,\xi)}$ the point of compactification for $\mathbb{R}_x^d \times \mathbb{R}_\xi^d \setminus \{0_\xi\}$. Note that the functions F and P remain l.s.c. on the compactified space, which is an important property we will need in the following. Moreover, for any $\bar{\rho} = \rho + \theta_{\bar{\rho}} \delta_{\infty_{(x,\xi)}}$, we have

$$c_t(\omega, \mu, \bar{\rho}) \geq c_t(\omega, \mu, \rho). \quad (3.5.9)$$

Now we have

$$I_t^\varepsilon = c_t(\mu_t^\varepsilon, \rho_t^\varepsilon), \quad (3.5.10)$$

and we can write (3.5.7) as

$$\sup_{\varepsilon>0} \mathbb{E} [c_t(\mu_t^\varepsilon, \rho_t^\varepsilon)] < \infty. \quad (3.5.11)$$

The following lemma, proved in Section 3.5.3, is the key of the demonstration for Theorem 3.2.1.

Lemma 3.5.3 (Characterization of limits). *Assume that (3.5.11) holds, then*

1. *The sequence $\{(\mu_t^\varepsilon, \rho_t^\varepsilon)\}$ is tight as a sequence of random variables with values in $\mathcal{P}(\mathbb{R}_x^d \times \mathbb{R}_u^d) \times \mathcal{M}(\mathbb{R}_x^d \times \mathbb{R}_\xi^d \setminus \{0_\xi\})$, equipped with the topology of weak convergence. In particular, $\{(\mu_t^\varepsilon, \rho_t^\varepsilon)\}$ is relatively compact in $\mathcal{P}^\mathbb{P}(\Omega \times \mathcal{P}(\mathbb{R}_x^d \times \mathbb{R}_u^d) \times \mathcal{M}(\overline{\mathbb{R}_x^d \times \mathbb{R}_\xi^d \setminus \{0_\xi\}}))$, equipped with the topology of stable convergence (see Appendix 3.D).*

2. Let $\mathbb{Q}_t \in \mathcal{P}^{\mathbb{P}}(\Omega \times \mathcal{P}(\mathbb{R}_x^d \times \mathbb{R}_u^d) \times \mathcal{M}(\overline{\mathbb{R}_x^d \times \mathbb{R}_\xi^d \setminus \{0_\xi\}}))$ be any \mathcal{F} -stable limit of $\{(\mu_t^\varepsilon, \rho_t^\varepsilon)\}$ with disintegration form

$$\mathbb{Q}_t(d\omega, d\mu, d\bar{\rho}) = \mathbb{P}(d\omega)\mathbb{Q}_t^\omega(d\mu, d\bar{\rho}).$$

and

$$\begin{aligned} S(a) = & \left\{ (\mu, \bar{\rho}) \in \mathcal{P}(\mathbb{R}_x^d \times \mathbb{R}_u^d) \times \mathcal{M}(\overline{\mathbb{R}_x^d \times \mathbb{R}_\xi^d \setminus \{0_\xi\}}), \right. \\ & \bar{\rho} = \rho + \theta_{\bar{\rho}}\delta_\infty \text{ with } \rho \in \mathcal{M}(\mathbb{R}_x^d \times \mathbb{R}_\xi^d \setminus \{0_\xi\}), \\ & \left. \int_{\mathbb{R}_x^d \times \mathbb{R}_u^d} A^a f(x, u)\mu(dx, du) + \int_{\mathbb{R}_x^d \times \mathbb{R}_\xi^d \setminus \{0_\xi\}} Bf(x, \xi)\rho(dx, d\xi) = 0, \forall f \in C_0^2(\mathbb{R}_x^d) \right\}, \end{aligned}$$

where A^a and B are given by (3.4.2) and (3.4.3). Then we have \mathbb{P} -almost surely,

$$(\mu, \bar{\rho}) \in S(a_t(\omega)), \quad \mathbb{Q}_t^\omega\text{-almost surely.}$$

By Lemma 3.5.3, we have, up to a subsequence,

$$(\mu_t^\varepsilon, \rho_t^\varepsilon) \rightarrow_{\mathcal{F}} \mathbb{Q}_t \in \mathcal{P}^{\mathbb{P}}(\mathcal{P}(\mathbb{R}_x^d \times \mathbb{R}_u^d) \times \mathcal{M}(\overline{\mathbb{R}_x^d \times \mathbb{R}_\xi^d \setminus \{0_\xi\}})).$$

Write \mathbb{Q}_t in disintegration form, we have

$$\mathbb{Q}_t(d\omega, d\mu, d\bar{\rho}) = \mathbb{P}(d\omega)\mathbb{Q}_t^\omega(d\mu, d\bar{\rho}). \quad (3.5.12)$$

and \mathbb{Q}_t^ω -almost surely,

$$(\mu, \rho) \in S(a_t(\omega)). \quad (3.5.13)$$

Since the cost functional c_t is lower semi-continuous, we have

$$\begin{aligned} \liminf_{\varepsilon \rightarrow 0} \mathbb{E}[I_t^\varepsilon] & \stackrel{(3.5.10)}{=} \liminf_{\varepsilon \rightarrow 0} \mathbb{E}[c_t(\mu_t^\varepsilon, \rho_t^\varepsilon)] \\ & \stackrel{(3.D.1)}{\geq} \mathbb{E}^{\mathbb{Q}_t} [c_t(\omega, \mu, \bar{\rho})] \\ & \stackrel{(3.5.9)}{\geq} \mathbb{E}^{\mathbb{Q}_t} [c_t(\omega, \mu, \rho)] \\ & \stackrel{(3.5.12)}{=} \int_{\Omega} \mathbb{P}(d\omega) \int_{\mathcal{P}(\mathbb{R}_x \times \mathbb{R}_u) \times \mathcal{M}(\overline{\mathbb{R}_x \times \mathbb{R}_\xi \setminus \{0_\xi\}})} c_t(\omega, \mu, \rho)\mathbb{Q}_t^\omega(d\mu, d\bar{\rho}) \\ & \stackrel{(3.5.13)}{=} \int_{\Omega} \mathbb{P}(d\omega) \int_{S(a_t(\omega))} c_t(\omega, \mu, \rho)\mathbb{Q}_t^\omega(d\mu, d\bar{\rho}) \\ & \geq \int_{\Omega} \mathbb{P}(d\omega) \inf_{(\mu, \rho) \in S(a_t(\omega))} c_t(\omega, \mu, \rho). \end{aligned}$$

Finally, by definition of I , we have

$$\liminf_{\varepsilon \rightarrow 0} \mathbb{E}[I_t^\varepsilon] \geq \mathbb{E}[I(a_t, r_t, l_t, k_t, h_t)].$$

3.5.3 Proof of Lemma 3.5.3

First we show the tightness of $\{(\mu^\varepsilon, \rho^\varepsilon)\}$ in $\mathcal{P}(\mathbb{R}_x^d \times \mathbb{R}_u^d) \times \mathcal{M}(\overline{\mathbb{R}_x^d \times \mathbb{R}_\xi^d \setminus \{0_\xi\}})$. A common method is to use tightness functions (see Section 3.C).

Recall that the cost functions F and P are extended to $\overline{\mathbb{R}_x^d \times \mathbb{R}_\xi^d \setminus \{0_\xi\}}$ by (3.5.8) such that c is lower semi-continuous. Moreover, c is a tightness function under Assumption 3.5.1, see Section 3.C or [DE11, pp. 309].

Consequently, if (3.5.11) holds, the family of random measures $\{(\mu^\varepsilon, \rho^\varepsilon)\}$ is tight. Furthermore, by Proposition 3.D.1, we have

$$(\mu_t^\varepsilon, \rho_t^\varepsilon) \rightarrow_{\mathcal{F}} \mathbb{Q}_t \in \mathcal{P}(\mathcal{P}(\mathbb{R}_x^d \times \mathbb{R}_u^d) \times \overline{\mathcal{M}(\mathbb{R}_x^d \times \mathbb{R}_\xi^d \setminus \{0_\xi\})}),$$

up to a subsequence, with

$$\mathbb{Q}_t(d\omega, d\mu, d\bar{\rho}) = \mathbb{P}(d\omega)\mathbb{Q}_t^\omega(d\mu, d\bar{\rho}).$$

For the rest of the lemma, we use a combination of the arguments in [KS99] and [KM93]. Recall that

$$A^a f(x, u) = \frac{1}{2} \sum_{i,j} a_{ij} \partial_{ij}^2 f(x) + u^T \nabla f(x), \quad Bf(x, \xi) = f(x + \xi) - f(x).$$

For $f \in C_0^2(\mathbb{R}_x^d)$, define

$$\Psi_t^f(\omega, \mu, \bar{\rho}) := \int_{\mathbb{R}_x^d \times \mathbb{R}_u^d} A^{a_t(\omega)} f(x, u) \mu(dx \times du) + \int_{\mathbb{R}_x^d \times \mathbb{R}_\xi^d \setminus \{0_\xi\}} Bf(x, \xi) \rho(dx, d\xi), \quad \bar{\rho} = \rho + \delta_{\bar{\rho}} \delta_\infty.$$

Note that Ψ_t^f is well-defined since $\rho \in \mathcal{M}(\mathbb{R}_x^d \times \mathbb{R}_\xi^d \setminus \{0_\xi\})$. Then we claim that

$$\mathbb{E}^{\mathbb{Q}}[|\Psi_t^f(\omega, \mu, \bar{\rho})|] = \lim_{\varepsilon \rightarrow 0} \mathbb{E}[|\Psi_t^f(\omega, \mu_t^\varepsilon(\omega), \rho_t^\varepsilon(\omega))|] = 0. \quad (3.5.14)$$

Although $\Psi^f(\omega, \cdot, \cdot) \in C(\mathcal{P}(\mathbb{R}_x^d \times \mathbb{R}_u^d) \times \overline{\mathcal{M}(\mathbb{R}_x^d \times \mathbb{R}_\xi^d \setminus \{0_\xi\})})$, it is not bounded. The first equality in (3.5.14) does not follow directly from the definition of stable convergence. However, by Corollary 3.4.1, Condition (3.4.14) holds, that is there exists $\beta \in (0, 1)$ and θ_f non-negative real number depending on f such that

$$(Af)^{1/\beta} \leq \theta_f(1 + C_A), \quad (Bf)^{1/\beta} \leq \theta_f C_B.$$

By (3.5.11), we deduce that $\{\Psi_t^f(\omega, \mu_t^\varepsilon, \rho_t^\varepsilon)\}$ is uniformly integrable and obtain the first equality in (3.5.14) by [JM81, Theorem 2.16].

For the second equality in (3.5.14), we apply Itô's formula to $f(\tilde{X}_{T^\varepsilon}^{\varepsilon,t})$ (recall that the dynamics of $\tilde{X}^{\varepsilon,t}$ is given by (3.2.4)) and obtain that

$$\begin{aligned} f(\tilde{X}_{T^\varepsilon}^{\varepsilon,t}) &= f(\tilde{X}_{0+}^{\varepsilon,t}) + \int_0^{T^\varepsilon} f'(\tilde{X}_s^{\varepsilon,t}) \sqrt{\tilde{a}_s^{\varepsilon,t}} d\tilde{W}_s^{\varepsilon,t} \\ &\quad + \int_0^{T^\varepsilon} \frac{1}{2} \sum_{ij} \tilde{a}_{ij,s}^{\varepsilon,t} \partial_{ij}^2 f(\tilde{X}_s^{\varepsilon,t}) ds + \int_0^{T^\varepsilon} \sum_i \tilde{u}_{i,s}^{\varepsilon,t} \partial_i f(\tilde{X}_s^{\varepsilon,t}) ds \\ &\quad + \sum_{0 < \tilde{\tau}_j^{\varepsilon,t} \leq T^\varepsilon} (f(\tilde{X}_{\tilde{\tau}_j^{\varepsilon,t}-}^{\varepsilon,t} + \tilde{\xi}_j^\varepsilon) - f(\tilde{X}_{\tilde{\tau}_j^{\varepsilon,t}-}^{\varepsilon,t})). \end{aligned}$$

Combining the definitions of μ^ε , ρ^ε and Ψ_t^f , we have

$$\begin{aligned} \mathbb{E}[|\Psi_t^f(\omega, \mu_t^\varepsilon(\omega), \rho_t^\varepsilon(\omega))|] &\leq \frac{1}{T^\varepsilon} \mathbb{E}[|f(\tilde{X}_{T^\varepsilon}^{\varepsilon,t}) - f(\tilde{X}_{0+}^{\varepsilon,t})|] + \frac{1}{T^\varepsilon} \mathbb{E}\left[\left|\int_0^{T^\varepsilon} f'(\tilde{X}_s^{\varepsilon,t}) \sqrt{\tilde{a}_s^{\varepsilon,t}} d\tilde{W}_s^{\varepsilon,t}\right|\right] \\ &\quad + \frac{1}{T^\varepsilon} \mathbb{E}\left[\int_0^{T^\varepsilon} \frac{1}{2} \sum_{ij} |\tilde{a}_{ij,s}^{\varepsilon,t} - \tilde{a}_{ij,0}^{\varepsilon,t}| \partial_{ij}^2 f(\tilde{X}_s^{\varepsilon,t}) ds\right]. \end{aligned}$$

By Assumptions 3.2.2 and 3.5.1 and dominated convergence, the term on the right hand side converges to zero. Therefore (3.5.14) holds.

By definition of \mathbb{Q}^ω and Fubini theorem, we have

$$\mathbb{E}^{\mathbb{Q}}[|\Psi_t^f(\omega, \mu, \bar{\rho})|] = \mathbb{E}^{\mathbb{P}}[\mathbb{E}^{\mathbb{Q}^\omega}[|\Psi_t^f(\omega, \mu, \bar{\rho})|]] = 0.$$

Hence we have \mathbb{P} -a.e.- ω , $\mathbb{E}^{\mathbb{Q}^\omega}[|\Psi_t^f(\omega, \mu, \bar{\rho})|] = 0$. Let D be a countable dense subset of C_0^2 . Since D is countable we have \mathbb{P} -a.e.- ω , $\mathbb{E}^{\mathbb{Q}^\omega}[|\Psi_t^f(\omega, \mu, \bar{\rho})|] = 0$ for all $f \in D$. Fix $\omega \in \Omega \setminus \mathcal{N}$ for which the property holds. Again by the same argument, we have \mathbb{Q}^ω -a.e.- $(\mu, \bar{\rho})$, $\Psi_t^f(\omega, \mu, \bar{\rho}) = 0$ for all $f \in D$. Since D is dense in C_0^2 , $\Psi_t^f(\omega, \mu, \bar{\rho}) = 0$ holds for $f \in C_0^2$.

3.6 Proof of Theorem 3.3.1

The rescaled process $(\tilde{X}_s^{\varepsilon,t})$ is given by

$$d\tilde{X}_s^{\varepsilon,t} = \tilde{b}_s^{\varepsilon,t} ds + \sqrt{\tilde{a}_s^{\varepsilon,t}} d\tilde{W}_s^{\varepsilon,t} + \tilde{u}_s^{\varepsilon,t} ds + \tilde{\gamma}_s^{\varepsilon,t} d\tilde{\varphi}_s^{\varepsilon,t},$$

with

$$\tilde{\gamma}_s^{\varepsilon,t} = \gamma_{t+\varepsilon^{2\beta}s}^\varepsilon, \quad \tilde{\varphi}_s^{\varepsilon,t} = \frac{1}{\varepsilon^\beta} (\varphi_{t+\varepsilon^{2\beta}s}^\varepsilon - \varphi_t^\varepsilon).$$

The empirical occupation measure of singular control ν_t^ε is defined by

$$\nu_t^\varepsilon = \frac{1}{T\varepsilon} \int_0^{T\varepsilon} \delta_{\{\tilde{X}_s^{\varepsilon,t}, \tilde{\gamma}_s^{\varepsilon,t}, \Delta\tilde{\varphi}_s^{\varepsilon,t}\}} d\tilde{\varphi}_s^{\varepsilon,t}.$$

while μ_t^ε is defined in the same way as previously.

Define the cost functional $c_t : \Omega \times \mathcal{P}(\mathbb{R}_x^d \times \mathbb{R}_u^d) \times \mathcal{M}(\overline{\mathbb{R}_x^d \times \Delta \times \mathbb{R}_\delta^+}) \rightarrow \mathbb{R}$:

$$\begin{aligned} (\omega, \mu, \bar{\nu}) \mapsto & \int_{\mathbb{R}_x^d \times \mathbb{R}_u^d} (r_t(\omega)D(x) + l_t(\omega)Q(u))\mu(dx \times du) \\ & + \int_{\mathbb{R}_x^d \times \Delta \times \mathbb{R}_\delta^+} h_t(\omega)P(\gamma)\bar{\nu}(dx \times d\gamma \times d\delta). \end{aligned}$$

Then we can show a similar version of Lemma 3.5.3 and prove Theorem 3.3.1 with the operator B replaced by (3.4.5). The key ingredients are

- The functional c is a tightness function.
- The condition (3.4.14) holds for A and B given by (3.4.4)-(3.4.5).

To satisfy these two properties is the main reason why we need to use a different operator B .

3.7 Proof of Propositions 3.4.1-3.4.5

In this section, we prove Propositions 3.4.1-3.4.5. First, we provide a verification argument tailored to the linear programming formulation in \mathbb{R}^d . Second, we give full details for the proof of Proposition 3.4.4. The proofs in the remaining cases are exactly the same hence omitted.

3.7.1 Verification theorem in \mathbb{R}^d

Consider $A : \mathcal{D} \rightarrow C(\mathbb{R}_x^d \times \mathbb{R}_u^d)$ and $B : \mathcal{D} \rightarrow C(\mathbb{R}_x^d \times V)$ with $\mathcal{D} = C_0^2(\mathbb{R}_x^d)$. The operator A is given by

$$Af(x, u) = \frac{1}{2} \sum_{ij} a_{ij} \partial_{ij}^2 f(x) + \sum_i u_i \partial_i f(x), \quad f \in C_0^2(\mathbb{R}_x^d),$$

The operator B is given by

$$Bf(x, \xi) = f(x + \xi) - f(x),$$

if $V = \mathbb{R}_\xi^d \setminus \{0_\xi\}$, and by

$$Bf(x, \gamma, \delta) = \begin{cases} \langle \gamma, \nabla f(x) \rangle, & \delta = 0, \\ \delta^{-1}(f(x + \delta\gamma) - f(x)), & \delta > 0, \end{cases}$$

if $V = \Delta \times \mathbb{R}_\delta^+$.

Let $C_A : \mathbb{R}_x^d \times \mathbb{R}_u^d \rightarrow \mathbb{R}^+$ and $C_B : \mathbb{R}_x^d \times V \rightarrow \mathbb{R}^+$ be two cost functions. We consider the following optimization problem:

$$I = \inf_{(\mu, \rho)} c(\mu, \rho) := \int_{\mathbb{R}_x^d \times \mathbb{R}_u^d} C_A(x, u) \mu(dx, du) + \int_{\mathbb{R}_x^d \times V} C_B(x, v) \rho(dx, dv), \quad (3.7.1)$$

where $(\mu, \rho) \in \mathcal{P}(\mathbb{R}_x^d \times \mathbb{R}_u^d) \times \mathcal{M}(\mathbb{R}_x^d \times V)$ satisfies

$$\int_{\mathbb{R}_x^d \times \mathbb{R}_u^d} Af(x, u) \mu(dx, du) + \int_{\mathbb{R}_x^d \times V} Bf(x, v) \rho(dx, dv) = 0, \quad \forall f \in C_0^2(\mathbb{R}_x^d). \quad (3.7.2)$$

Lemma 3.7.1 (Verification). *Let $w \in C^1(\mathbb{R}_x^d) \cap C^2(\mathbb{R}_x^d \setminus N)$ so that Aw is well-defined pointwisely for $x \notin N$ and Bw is well-defined for $x \in \mathbb{R}_x^d$. Assume that*

1. *For each $(\mu, \rho) \in \mathcal{P}(\mathbb{R}_x^d \times \mathbb{R}_u^d) \times \mathcal{M}(\mathbb{R}_x^d \times V)$ satisfying (3.7.2) and $c(\mu, \rho) < \infty$, we have $\mu(N \times \mathbb{R}_u^d) = 0$.*
2. *There exists $w_n \in C_0^2(\mathbb{R}_x^d)$ such that*

$$\begin{aligned} Aw_n(x, u) &\rightarrow Aw(x, u), \quad \forall (x, u) \in \mathbb{R}_x^d \setminus N \times \mathbb{R}_u^d, \\ Bw_n(x, v) &\rightarrow Bw(x, v), \quad \forall (x, v) \in \mathbb{R}_x^d \times V, \end{aligned}$$

and there exist $\theta \in \mathbb{R}^+$ such that

$$\begin{aligned} |Aw_n(x, u)| &\leq \theta(1 + C_A(x, u)), \quad \forall (x, u) \in (\mathbb{R}_x^d \setminus N) \times \mathbb{R}_u^d \\ |Bw_n(x, v)| &\leq \theta C_B(x, v), \quad \forall (x, v) \in \mathbb{R}_x^d \times V. \end{aligned}$$

3. *There exists a constant $I^V \in \mathbb{R}$ such that*

$$\inf_{u \in \mathbb{R}_u^d} Aw(x, u) + C_A(x, u) \geq I^V, \quad x \in \mathbb{R}_x^d \setminus N, \quad (3.7.3)$$

$$\inf_{v \in V} Bw(x, v) + C_B(x, v) \geq 0, \quad x \in \mathbb{R}_x^d. \quad (3.7.4)$$

Then we have $I \geq I^V$.

If there exists (μ^*, ρ^*) satisfying the LP constraint and

$$Aw(x, u) + C_A(x, u) = I^V, \quad \mu^* - a.e. \quad (3.7.5)$$

$$Bw(x, v) + C_B(x, v) = 0, \quad \rho^* - a.e. \quad (3.7.6)$$

then we have $I = I^V$. Moreover, the optimum is attained by (μ^*, ρ^*) and we call (w, I^V) the value function of the linear programming problem.

Proof of Lemma 3.7.1. Let (μ, ρ) be any pair satisfying (3.7.2) and $c(\mu, \rho) < \infty$. We have

$$\begin{aligned} &\int_{\mathbb{R}_x^d \times \mathbb{R}_u^d} Aw(x, u) \mu(dx, du) + \int_{\mathbb{R}_x^d \times V} Bw(x, v) \rho(dx, dv) \\ &= \int_{\mathbb{R}_x^d \times \mathbb{R}_u^d} Aw_n(x, u) \mu(dx, du) + \int_{\mathbb{R}_x^d \times V} Bw_n(x, v) \rho(dx, dv) \\ &= 0. \end{aligned}$$

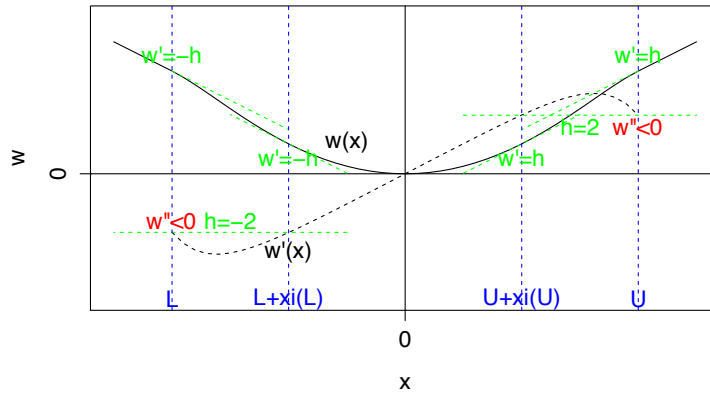


Figure 3.1 – Value function for combined regular and impulse control

The first term is well-defined since Aw is defined μ -everywhere. The second equality follows from the second condition and dominated convergence theorem. Hence

$$c(\mu, \rho) = \int_{\mathbb{R}_x^d \times \mathbb{R}_u^d} (C_A(x, u) + Aw(x, u))\mu(dx, du) + \int_{\mathbb{R}_x^d \times V} (C_B(x, v) + Bw(x, v))\rho(dx, dv) \geq I^V,$$

where the last inequality is due to (3.7.3)-(3.7.4), and the equality holds if and only if (3.7.5)-(3.7.6) are satisfied. \square

Therefore, finding an explicit solution of a linear programming is possible if we can determine a suitable value function from (3.7.5)-(3.7.6).

3.7.2 Verification of Proposition 3.4.4

In this section, we provide an explicit solution of the following linear programming problem:

$$I(a, r, l, k, h) = \inf_{(\mu, \rho)} \int_{\mathbb{R}_x \times \mathbb{R}_u} (rx^2 + lu^2)\mu(dx, du) + \int_{\mathbb{R}_x \times \mathbb{R}_\xi \setminus \{0_\xi\}} (k + h|\xi|)\rho(dx, d\xi), \quad (3.7.7)$$

where $\mu \in \mathcal{P}(\mathbb{R}_x \times \mathbb{R}_u)$ and $\rho \in \mathcal{M}(\mathbb{R}_x \times \mathbb{R}_\xi \setminus \{0_\xi\})$ satisfy

$$\int_{\mathbb{R}_x \times \mathbb{R}_u} \left(\frac{1}{2}af''(x) + uf'(x)\right)\mu(dx, du) + \int_{\mathbb{R}_x \times \mathbb{R}_\xi \setminus \{0_\xi\}} (f(x + \xi) - f(x))\rho(dx, d\xi) = 0, \quad (3.7.8)$$

for any $f \in C_0^2(\mathbb{R})$. The following lemma, whose proof is given in the Appendix, and Theorem 3.7.1 establish the existence of the value function (w, I^V) for (3.7.7)-(3.7.8).

Lemma 3.7.2 (Value function for combined regular and impulse control). *There exist $U > \xi^* > 0$, $I^V > 0$, and $w \in C^1(\mathbb{R}) \cap C^2(\mathbb{R} \setminus \{U, U\})$ such that*

$$Aw(x, u^*(x)) + C_A(x, u^*(x)) = I^V, \quad x \in (-U, U), \quad (3.7.9)$$

$$Bw(x, -\text{sgn}(x)\xi^*) + C_B(x, -\text{sgn}(x)\xi^*) = 0, \quad x \in \{-U, U\}, \quad (3.7.10)$$

where $u^*(x)$ is defined by

$$u^*(x) := \underset{u \in \mathbb{R}_u}{\text{Argmin}} Aw(x, u) + C_A(x, u) = -\frac{w'(x)}{2l}. \quad (3.7.11)$$

More precisely, we have (c.f. Figure 3.1)

$$w(x) = \begin{cases} (rl)^{1/2}x^2 - 2al \ln {}_1F_1\left(\frac{1-\iota}{4}; \frac{1}{2}; \left(\frac{r}{a^2l}\right)^{1/2}x^2\right), & |x| \leq U, \\ w(U) + h(|x| - U), & |x| > U, \end{cases} \quad (3.7.12)$$

where ${}_1F_1$ is the Kummer confluent hypergeometric function (see Section 3.A) and

$$I^V = \iota\sqrt{a^2rl},$$

for some $\iota \in (0, 1)$. Moreover, w satisfies the following conditions

$$w'(x) = w'(x - \text{sgn}(x)\xi^*) = \text{sgn}(x)h, \quad x \in \{-U, U\}, \quad (3.7.13)$$

$$w''(x) < 0, \quad x \in \{-U, U\}, \quad (3.7.14)$$

$$w'(x) \in \begin{cases} (-\infty, -h), & -U < x < -U + \xi^*, \\ (-h, h), & -U + \xi^* < x < U - \xi^*, \\ (h, \infty), & U - \xi^* < x < U, \end{cases} \quad (3.7.15)$$

and w , ξ^* , U and I^V depend continuously on the parameters (a, r, k, h) .

Remark 3.7.1. Equations (3.7.9) and (3.7.10) correspond essentially to (3.7.5) and (3.7.6). The interval $(-U, U)$ is called continuation region. Equation (3.7.13) is the so called “smooth-fit” condition and guarantees that w is a C^1 function. Equations (3.7.14) and (3.7.15) characterize the growth of the derivatives of w and will be useful in the proof of Theorem 3.7.1.

Proposition 3.4.4 is a direct consequence of the following theorem.

Theorem 3.7.1 (Combined regular and impulse control.). *For any parameters $a, r, l, k > 0$ and $h \geq 0$, we have*

1. The pair (w, I^V) in Lemma 3.7.2 is the value function of (3.7.7)-(3.7.8) in the sense of Lemma 3.7.1. In particular, the optimal cost of (3.7.7)-(3.7.8) is given by $I = I^V$.
2. Let $p^*(x) \in C^0([-U, U]) \cap C^2((-U, U) \setminus \{-U + \xi^*, U - \xi^*\})$ be a solution of

$$\begin{cases} \frac{1}{2}ap''(x) - (u^*(x)p(x))' = 0, & x \in (-U, U) \setminus \{-U + \xi^*, U - \xi^*\}, \\ p(-U) = p(U) = 0, \\ \frac{1}{2}ap'((-U)+) = p'((-U + \xi^*)-) - \frac{1}{2}a(p'((-U + \xi^*)+)), \\ \frac{1}{2}ap'(U-) = p'((U - \xi^*)+) - \frac{1}{2}a(p'((U - \xi^*)-)), \\ \int_{-U}^U p(x) = 1, \end{cases} \quad (3.7.16)$$

write $\rho_-^*, \rho_+^* \in \mathbb{R}^+$ for

$$\rho_-^* = \frac{1}{2}ap'((-U)+), \quad \rho_+^* = -\frac{1}{2}ap'(U-), \quad (3.7.17)$$

and recall that u^* is given by (3.7.11). Then the optimum of (3.7.7)-(3.7.8) is attained by

$$\mu^*(dx, du) = p^*(x)dx \otimes \delta_{u^*(x)}(du), \quad \rho^*(dx, d\xi) = \rho_-^*\delta_{(-U, \xi^*)} + \rho_+^*\delta_{(U, -\xi^*)}. \quad (3.7.18)$$

Proof.

1. Consider the function w defined in Lemma 3.7.2. First we show that the three conditions in Lemma 3.7.1 are satisfied by w .

i) Note that $N = \{-U, U\}$. For any (μ, ρ) satisfying the LP constraint, we show that $\mu(\{x\} \times \mathbb{R}_u) = 0, \forall x \in \mathbb{R}_x$. In particular, $\mu(N \times \mathbb{R}_u) = 0$. Indeed, let $f_n \in C_0^2(\mathbb{R}_x)$ be a sequence of test functions such that $f_n'' \rightarrow \mathbb{1}_{\{x\}}, \|f_n\|_\infty \vee \|f_n'\|_\infty \rightarrow 0$ and there exists $\theta \in \mathbb{R}^+$ such that

$$|Af_n(x, u)| \leq \theta(1 + C_A(x, u)), \quad |Bf_n(x, \xi)| \leq \theta C_B(x, \xi) \quad \forall x \in \mathbb{R}_x, u \in \mathbb{R}_u, \xi \in \mathbb{R}_\xi \setminus \{0_\xi\}.$$

For example, let $\varphi \in C_0^2$ with φ'' being a piece-wise linear function such that $\varphi''(\pm\infty) = \varphi''(-1) = \varphi''(1) = \varphi''(3) = \varphi''(5) = 0, \varphi''(0) = \varphi''(4) = 1$ and $\varphi''(2) = -2$ and take $f_n(z) = \frac{1}{n^2}\varphi(n(z-x))$. Since $c(\mu, \rho) < \infty$, we have by dominated convergence theorem

$$\mu(\{x\} \times \mathbb{R}_u) = \lim_n \int Af_n(z, u)\mu(dz \times du) + \int Bf_n(z, \xi)\rho(dz \times d\xi) = 0.$$

ii) Let $\varphi_n \in C_0^2$ be a sequence of indicator functions such that $\varphi_n(x) = 1$ for $|x| \leq n$ and

$$\sup_n \|\varphi_n\|_{C_0^2} := \|\varphi_n\|_\infty \vee \|\varphi_n'\|_\infty \vee \|\varphi_n''\|_\infty < \infty.$$

Let $w_n = w\varphi_n$. Then w_n is C^2 except at $\{-U, U\}$ and is of compact support. For each n , w_n satisfies also the LP constraint

$$\int Aw_n d\mu + \int Bw_n d\rho = 0.$$

Indeed, let φ_δ be any convolution kernel and $w_{n,\delta} := w_n * \varphi_\delta$. So $w_{n,\delta}$ satisfies the LP constraint. Moreover, $Aw_{n,\delta} \rightarrow Aw_n$ for $x \notin \{-U, U\}, Bw_{n,\delta} \rightarrow Bw_n$ for any (x, u) and $\sup_\delta \|w_{n,\delta}\|_{C_0^2} < \|w_n\|_{C_0^2} < \infty$. By dominated convergence, w_n satisfies the LP constraint. Finally, a direct computation shows that for some constants θ and θ' ,

$$|Aw_n| \leq \theta' \|\varphi_n\|_{C_0^2} (|w| + |w'| + |w''|) \leq \theta(1 + C_A), \quad |Bw_n| \leq 2\|\varphi_n\|_\infty |w_n| \leq \theta C_B.$$

So the second condition is satisfied.

iii) By (3.7.9) and (3.7.14), we have $Aw + C_A \geq 0$ for $x \notin \{-U, U\}$. By (3.7.10)-(3.7.15) and definition of w outside $[U, U]$, we have $Bw + C_B \geq 0$.

By Lemma 3.7.1, we then conclude that $I = I^V$.

2. We need to show that μ^* and ρ^* satisfy the LP constraint. Assume that μ^* and ρ^* are given by (3.7.18), then by integration by parts, the LP constraint holds if $p^*(x)$ is solution of (3.7.16). It is easy to see that the latter admits a unique solution. □

Appendix 3.A Kummer confluent hypergeometric function ${}_1F_1$

We collect here some properties of the Kummer confluent hypergeometric function ${}_1F_1$ which are useful to establish the existence of value functions of combined control problems in dimension one. Recall that ${}_1F_1$ is defined as

$${}_1F_1(a, b, z) = \sum_{k=0}^{\infty} \frac{(a)_k}{(b)_k} \frac{z^k}{k!},$$

with $(a)_k$ the Pochhammer symbol.

Lemma 3.A.1. *We have the following properties.*

1. *The function ${}_1F_1$ admits the following integral representation*

$${}_1F_1(a, b, z) = \frac{\Gamma(b)}{\Gamma(b-a)\Gamma(a)} \int_0^1 e^{zt} t^{a-1} (1-t)^{b-a-1} dt.$$

It is an entire function of a and z and a meromorphic function of b .

2. *We have*

$$\begin{aligned} \frac{\partial}{\partial z} {}_1F_1(a, b, z) &= \frac{a}{b} {}_1F_1(a+1, b+1, z), \\ \frac{\partial}{\partial a} {}_1F_1(a, b, z) &= \sum_{k=0}^{\infty} \frac{(a)_k}{(b)_k} \frac{z^k}{k!} \sum_{p=0}^{k-1} \frac{1}{p+a}. \end{aligned}$$

3. *We have*

$$(a+1)z {}_1F_1(a+2, b+2, z) + (b+1)(b-z) {}_1F_1(a+1, b+1, z) - b(b+1) {}_1F_1(a, b, z) = 0.$$

4. *We have*

$${}_1F_1(a, b, z) = \frac{\Gamma(a)}{\Gamma(b-a)} e^{i\pi a} z^{-a} (1 + O(\frac{1}{|z|})) + \frac{\Gamma(b)}{\Gamma(a)} e^z z^{a-b} (1 + O(\frac{1}{|z|})),$$

as $z \rightarrow \infty$.

5. *Consider the Weber differential equation*

$$w''(x) - \left(\frac{1}{4}x^2 + \theta\right)w(x) = 0. \quad (3.A.1)$$

The even and odd solutions of this equation are given, respectively, by

$$\tilde{w}(x; \theta) = e^{-\frac{1}{4}x^2} {}_1F_1\left(\frac{1}{2}\theta + \frac{1}{4}, \frac{1}{2}, \frac{1}{2}x^2\right), \quad (3.A.2)$$

$$\bar{w}(x; \theta) = x e^{-\frac{1}{4}x^2} {}_1F_1\left(\frac{1}{2}\theta + \frac{3}{4}, \frac{3}{2}, \frac{1}{2}x^2\right). \quad (3.A.3)$$

Proof. See [AS72, AG08]. □

Appendix 3.B Proof of Lemma 3.7.2

We first look for w in the continuation region $(-U, U)$. Define (the change of variable comes from [OS05, pp.260])

$$w(x) := -2al \ln \tilde{w}\left(\frac{x}{\alpha}; -\frac{\iota}{2}\right), \quad \alpha^2 = \frac{1}{2}a(r^{-1}l)^{1/2}, \quad \iota = \frac{I^V}{(a^2rl)^{1/2}},$$

where \tilde{w} is the odd solution (3.A.2) of the Weber differential equation (3.A.1). Then w satisfies the following ODE

$$\frac{1}{2}aw''(x) - \frac{1}{4l}(w'(x))^2 + rx^2 = I^V,$$

which is exactly (3.7.9). Hence we conjecture that the solution in the continuation region $(-U, U)$ is given by

$$\begin{aligned} w(x) &= -2al \ln \left(e^{-\frac{1}{4}x^2/\alpha^2} {}_1F_1\left(\frac{1-\iota}{4}, \frac{1}{2}, \frac{1}{2\alpha^2}x^2\right) \right) \\ &= (rl)^{1/2}x^2 - 2al \ln \left({}_1F_1\left(\frac{1-\iota}{4}, \frac{1}{2}, \frac{1}{2\alpha^2}x^2\right) \right). \end{aligned}$$

Now we show that there exist suitable values U , $\xi(U)$ and ι such that $0 \leq U + \xi(U) \leq U$, $\iota \in (0, 1)$ and Conditions (3.7.10)-(3.7.15) are satisfied. Let

$$h(x; \iota) := \frac{\partial w}{\partial x} = 2(rl)^{1/2} \left(1 - (1 - \iota)g\left(\frac{1}{2\alpha^2}x^2; \iota\right)\right)x,$$

with

$$g(z; \iota) = \frac{{}_1F_1\left(\frac{1-\iota}{4} + 1, \frac{1}{2} + 1, z\right)}{{}_1F_1\left(\frac{1-\iota}{4}, \frac{1}{2}, z\right)}.$$

We have the following lemma.

Lemma 3.B.1. *The function $g(z; \iota)$ satisfies*

$$g(z; \iota) \rightarrow \begin{cases} 1, & z \rightarrow 0^+, \\ \frac{2}{1-\iota}, & z \rightarrow +\infty, \end{cases}$$

and

$$g'(z; \iota) > 0, \quad \forall z \in [0, +\infty).$$

Proof. The limits of $g(z; \iota)$ follow from the asymptotic behaviour of ${}_1F_1$. To show that g is increasing, we use Properties 2 and 3 in Lemma 3.A.1 and obtain

$$g'(z) = g(z) \left(1 - \frac{1-\iota}{2}g(z)\right) + \frac{1}{2z}(1 - g(z)).$$

Note that $g'(0) > 0$, so $g > 1$ near $x = 0$. Since $g'(z) > 0$ for $g(z) = 1$ and $g'(z) < 0$ for $g(z) = \frac{2}{1-\iota}$, $g(z)$ cannot leave the band $[1, \frac{2}{1-\iota}]$. \square

We now state a second lemma.

Lemma 3.B.2. *The function $h(x; \iota)$ satisfies the following properties.*

1. For $\iota \in (0, 1)$, we have

$$\frac{h(x; \iota)}{x} \rightarrow \begin{cases} 2(rl)^{1/2}\iota, & x \rightarrow 0+, \\ -2(rl)^{1/2}, & x \rightarrow +\infty. \end{cases}$$

Let $0 < \bar{x}_\iota < \infty$ be the first zero of $h(x; \iota)$. We have

$$h'(x; \iota) = \begin{cases} 2(rl)^{1/2}\iota > 0, & x = 0, \\ -2(rl)^{1/2}2(1-\iota)\bar{x}_\iota g'(\bar{z}_\iota) < 0, & x = \bar{x}_\iota, \end{cases}$$

where $\bar{z}_\iota = \frac{1}{2\alpha^2}\bar{x}_\iota^2$ and

$$h''(x; \iota) < 0, \quad x \in [0, \bar{x}_\iota].$$

2. For $x \in (0, \infty)$, we have

$$\frac{\partial}{\partial \iota} h(x; \iota) > 0$$

and

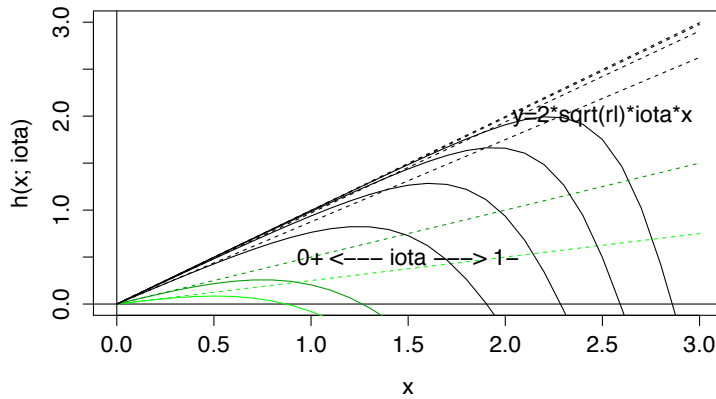
$$h(x; \iota) \rightarrow 2(rl)^{1/2}x, \quad \iota \rightarrow 1-.$$

We have

$$\bar{x}_\iota = O(\iota^{1/2}), \quad \iota \rightarrow 0+$$

and hence

$$\max_{x \in [0, \bar{x}_\iota]} h(x; \iota) \rightarrow 0, \quad \iota \rightarrow 0+.$$

Figure 3.2 – Qualitative behaviour of $h(x; \iota)$

Proof. For fixed ι . The asymptotic behaviour of h follows from Lemma 3.B.1. The second property is clear since

$$\begin{aligned} h'(x; \iota) &= 2(rl)^{1/2} \left(-2(1 - \iota)g'(z)z + (1 - (1 - \iota)g(z)) \right) \\ &= 2(rl)^{1/2} (-2(1 - \iota)g'(z)z) + \frac{h(x)}{x}, \end{aligned}$$

with $z = \frac{1}{2\alpha^2}x^2$. Finally, we get

$$h''(x; \iota) = \frac{d}{dz} h'(z; \iota) z'(x) = -2(rl)^{1/2} (1 - \iota) (3g'(z) + 2g''(z)z) z'(x).$$

Furthermore, we have

$$\begin{aligned} 3g'(z) + 2g''(z)z &= 3g'(z) + 2z \left(g'(z)(1 - (1 - \iota)g(z)) + \frac{1}{2z^2} (g(z) - 1 - zg'(z)) \right) \\ &= 2zg'(z)(1 - (1 - \iota)g(z)) + \frac{1}{z} (g(z) - 1) + 2g'(z) \end{aligned}$$

which is strictly positive term by term for $x \in [0, \bar{x}_\iota]$.

For fixed x . The limit of h as $\iota \rightarrow 1-$ follows from the fact that ${}_1F_1(a, b, z)$ is entire in a . Now we show that h is monotone in ι . Let $G := \partial_\iota {}_1F_1$, we have

$$\begin{aligned} \partial_\iota h(x; \iota) &= -2al \frac{\partial}{\partial \iota} \frac{\partial}{\partial x} \ln {}_1F_1 \left(\frac{1 - \iota}{4}; \frac{1}{2}; \frac{1}{2\alpha^2} x^2 \right) \\ &= -2al \frac{\partial}{\partial x} \frac{\partial}{\partial \iota} \ln {}_1F_1 \left(\frac{1 - \iota}{4}; \frac{1}{2}; \frac{1}{2\alpha^2} x^2 \right) \\ &= \frac{1}{2} al \frac{\partial}{\partial x} \frac{G}{{}_1F_1} \left(\frac{1 - \iota}{4}; \frac{1}{2}; \frac{1}{2\alpha^2} x^2 \right) \\ &= \frac{1}{2} al z'(x) \frac{{}_1F_1 \frac{\partial}{\partial z} G - G \frac{\partial}{\partial z} {}_1F_1}{{}_1F_1^2} \left(\frac{1 - \iota}{4}; \frac{1}{2}; z \right), \end{aligned}$$

with $z(x) = \frac{1}{2\alpha^2}x^2$. It is enough to show that the last term is positive. Using the series representation of ${}_1F_1$ and G (see Lemma 3.A.1), write

$${}_1F_1 = \sum_{k=0}^{\infty} f_k z^k, \quad G = \sum_{k=0}^{\infty} \beta_k f_k z^k,$$

with

$$\beta_k = \sum_{p=0}^{k-1} \frac{1}{p+a}, \quad a = \frac{1-\iota}{4}.$$

We get

$$\begin{aligned} {}_1F_1 \frac{\partial}{\partial x} G - G \frac{\partial}{\partial x} {}_1F_1 &= \left(\sum_{i=0}^{\infty} f_i z^i \right) \left(\sum_{j=0}^{\infty} (j+1) \beta_{j+1} f_{j+1} z^j \right) - \left(\sum_{i=0}^{\infty} (i+1) f_{i+1} z^i \right) \left(\sum_{j=0}^{\infty} \beta_j f_j z^j \right) \\ &= \sum_{k=0}^{\infty} \left(\sum_{i+j=k} f_i (j+1) \beta_{j+1} f_{j+1} \right) z^k - \sum_{k=0}^{\infty} \left(\sum_{i+j=k} (i+1) f_{i+1} \beta_j f_j \right) z^k. \end{aligned}$$

Then the coefficient of z^k is given by

$$\begin{aligned} &\sum_{i+j=k} f_i (j+1) \beta_{j+1} f_{j+1} - \sum_{i+j=k} (i+1) f_{i+1} \beta_j f_j \\ &= \sum_{i+j=k} (j+1) (\beta_{j+1} - \beta_i) f_i f_{j+1} \\ &= \sum_{1 \leq i < j+1 \leq k} \left((j+1) (\beta_{j+1} - \beta_i) f_i f_{j+1} - i (\beta_i - \beta_{j+1}) f_{j+1} f_i \right) + \sum_{j+1=k+1} (\dots) + \sum_{i=j+1} (\dots) \\ &= \sum_{1 \leq i < j+1 \leq k} (j+1-i) (\beta_{j+1} - \beta_i) f_i f_{j+1} + \sum_{j+1=k+1} (\dots) + \sum_{i=j+1} (\dots). \end{aligned}$$

This term is positive since β_k is increasing in k . Hence $\partial_\iota h > 0$. Thus $h(x; \iota)$ is increasing in ι for fixed $x \in \mathbb{R}^+$.

From the relation between g and h , \bar{z}_ι is the first solution of

$$1 - (1-\iota)g(z) = 0, \quad z > 0.$$

Moreover, we have

$$g(z) = 1 + \left(1 - \frac{1-\iota}{2}\right)z + o(z), \quad z \rightarrow 0+.$$

Then uniformly on ι , $g(z)$ is bounded from below by $1 + \frac{1}{3}z$ on $[0, z_0]$, hence

$$\bar{z}_\iota \leq \frac{3\iota}{1-\iota} = O(\iota), \quad \iota \rightarrow 0+.$$

Finally, we have

$$\max_{[0, \bar{x}_\iota]} h(x; \iota) \leq 2(r\iota)^{1/2} \bar{x}_\iota \rightarrow 0, \quad \iota \rightarrow 0+.$$

□

Proposition 3.B.1. *For any parameters $r, l, h > 0$ and $k \geq 0$, there exist $\iota \in (0, 1)$ and $0 \leq U + \xi \leq U$ such that*

$$\begin{aligned} \int_{U+\xi}^U h(x; \iota) dx &= k - h\xi, \\ h(U; \iota) &= h, \\ h(U + \xi; \iota) &= h, \\ h(x; \iota) &\in \begin{cases} (0, h), & 0 \leq x \leq U + \xi, \\ (h, \infty), & U + \xi \leq x \leq U, \end{cases} \\ h''(x; \iota) &< 0, \quad 0 \leq x \leq U. \end{aligned}$$

Moreover, (ι, ξ, U) depends continuously on (r, l, k, h) .

Proof. Existence. Let $k > 0$. Since $h(x; \iota)$ is monotone in ι and $h(x; \iota) \rightarrow 2(rl)^{1/2}x$ as $\iota \rightarrow 1-$, there exists $\underline{\iota} = \underline{\iota}(h) \geq 0$ such that

$$(\underline{\iota}, 1) = \{\iota \in (0, 1), \text{ there exists exactly two solutions } U_\iota + \xi_\iota \text{ and } U_\iota \text{ on } [0, \bar{x}_\iota]\}.$$

We have

$$h'(U_\iota + \xi_\iota; \iota) > 0, \quad h'(U_\iota) < 0,$$

so by the implicit function theorem, U_ι and $U_\iota + \xi_\iota$ depend continuously on ι . Define

$$K(\iota) = \int_{U_\iota + \xi_\iota}^{U_\iota} h(x; \iota) dx.$$

Then K is continuous in ι and

$$\lim_{\iota \rightarrow \underline{\iota}} K(\iota) = 0, \quad \lim_{\iota \rightarrow 1-} K(\iota) = \infty.$$

Hence there exists $\iota(h, k) \in (\underline{\iota}(h), 1)$ such that $K(\iota(h, k)) = k$. The remaining property of h is easily verified.

If $k = 0$, then there exists exactly one $\iota(h) \in (0, 1)$ such that the maximum of $h(x; \iota)$ is h and is attained by U_ι such that

$$h'(U_\iota; \iota) = 0.$$

Since $h''(U_\iota; \iota) < 0$, U_ι depends continuously on ι by the implicit function theorem.

Continuous dependence. Since ξ and U depend continuously on ι , it suffices to show that ι depends continuously on the parameters a, r, h, k, l . To see this, note that $\iota = \iota(a, l, r, h, k)$ is determined by

$$K(\iota; a, r, l, k, h) = k.$$

But, we have

$$\frac{\partial}{\partial \iota} K(\iota; a, r, l, k, h) > 0.$$

Thus ι depends continuously on the parameters by the implicit function theorem. \square

Proof of Lemma 3.7.2. Extend the function w in Proposition 3.B.1 to \mathbb{R} by

$$w(x) = \begin{cases} w(|x|), & |x| \leq U, \\ w(U) + h(|x| - U), & |x| > U. \end{cases}$$

Then (3.7.9)-(3.7.14) hold. By (3.7.10) and (3.7.13), we have $w \in C^1(\mathbb{R}) \cap C^2(\mathbb{R} \setminus \{U, U\})$. \square

Appendix 3.C Tightness function

For more details on the following results, see [DE11, Appendix A.3] and [Bog07].

Definition 3.C.1. A measurable real-valued function g on a metric space $g : (E, d) \rightarrow \mathbb{R} \cup \{\infty\}$ is a tightness function if

1. $\inf_{x \in E} g(x) > -\infty$.
2. $\forall M < \infty$, the level set $\{x \in E | g(x) \leq M\}$ is a relatively compact subset of (E, d) .

Lemma 3.C.1. If g is a tightness function on a Polish space E , then

1. The function $G(\mu) = \int_E g(x) \mu(dx)$ is a tightness function on $\mathcal{P}(E)$.

2. If in addition $g \geq \delta$ where δ is a positive constant and E is compact, then $G(\mu) = \int_E g(x)\mu(dx)$ is a tightness function on $\mathcal{M}(E)$.

Proof. Note that $\mathcal{M}(E)$ is a metric space, thus sequential compactness is equivalent to relative compactness (see [DE11, pp. 303] for the metric). For the first property, see [DE11, pp. 309]. For the second property, we consider the level set $\{\mu \in \mathcal{M}(E) | G(\mu) \leq M\}$ and let $\{\mu_n\}$ be any sequence in the level set. By [Bog07, Theorem 8.6.2], it is enough to show that

1. The sequence of nonnegative real numbers $\mu_n(E)$ is bounded.
2. The family $\{\mu_n\}$ is tight.

Since $g \geq \delta$, we have $\mu(E) \leq G(\mu)/\delta \leq M/\delta$. Hence the first condition is true. On the other hand, for any $\varepsilon > 0$, we consider $\mu_n/\mu_n(E) \in \mathcal{P}(E)$. Then $G(\mu_n/\mu_n(E)) \leq M/\mu_n(E) \leq M/\varepsilon$, if $\mu_n(E) > \varepsilon$. Since G is a tightness function, we deduce that $\{\mu_n | \mu_n(E) > \varepsilon\}$ is tight. Therefore $\{\mu_n\}$ is tight and the second condition follows. \square

Appendix 3.D Convergence in probability, stable convergence

Let (Ω, \mathcal{F}) be a measurable space and (E, \mathcal{E}) a Polish space where \mathcal{E} is the Borel algebra of E . Define

$$\bar{\Omega} = \Omega \times E, \quad \bar{\mathcal{F}} = \mathcal{F} \otimes \mathcal{E}.$$

Let $B_{mc}(\bar{\Omega})$ be the set of bounded measurable functions g such that $z \mapsto g(\omega, z)$ is a continuous application for any $\omega \in \Omega$. Let $M_{mc}(\bar{\Omega})$ be the set of finite positive measures on $(\bar{\Omega}, \bar{\mathcal{F}})$, equipped with the weakest topology such that

$$\mu \mapsto \int_{\bar{\Omega}} g(\omega, z)\mu(d\omega, dz),$$

is continuous for any $g \in B_{mc}(\bar{\Omega})$.

We fix a probability measure \mathbb{P} on (Ω, \mathcal{F}) . Let $\mathcal{P}^{\mathbb{P}}(\Omega \times E, \mathcal{F} \otimes \mathcal{E}) \subset M_{mc}(\bar{\Omega})$ be the set of probability measures on $\bar{\Omega}$ with marginal \mathbb{P} on Ω , equipped with the induced topology from $M_{mc}(\bar{\Omega})$. Note that $\mathcal{P}^{\mathbb{P}}(\Omega \times E, \mathcal{F} \otimes \mathcal{E})$ is a closed subset of $M_{mc}(\bar{\Omega})$. For any random variable Z defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, we define

$$\mathbb{Q}^Z(d\omega, dz) := \mathbb{P}(d\omega) \otimes \delta_{Z(\omega)}(dz) \in \mathcal{P}^{\mathbb{P}}(\Omega \times E, \mathcal{F} \otimes \mathcal{E}).$$

Definition 3.D.1 (Stable convergence). *Let $\{Z^\varepsilon, \varepsilon > 0\}$ be random variables defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with values in the Polish space (E, \mathcal{E}) . We say that Z^ε converges stably in law to $\mathbb{Q} \in \mathcal{P}^{\mathbb{P}}(\Omega \times E, \mathcal{F} \otimes \mathcal{E})$, written $Z^\varepsilon \rightarrow_{stable} \mathbb{Q}$, if $\mathbb{Q}^{Z^\varepsilon} \rightarrow \mathbb{Q}$ in $M_{mc}(\bar{\Omega})$.*

We use the following properties in our proofs.

Proposition 3.D.1. *Let $\{Z^\varepsilon, \varepsilon > 0\}$ be random variables on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with values in (E, \mathcal{E}) .*

1. *We have $Z^\varepsilon \rightarrow_{stable} \mathbb{Q} \in \mathcal{P}^{\mathbb{P}}(\Omega \times E, \mathcal{F} \otimes \mathcal{E})$ if and only if*

$$\mathbb{E}[Yf(Z^\varepsilon)] \rightarrow \mathbb{E}^{\mathbb{Q}}[Yf(z)],$$

for all bounded random variables Y on (Ω, \mathcal{F}) and all bounded continuous functions $f \in C_b(E, \mathbb{R})$.

2. *Assume that $Z^\varepsilon \rightarrow_{stable} \mathbb{Q} \in \mathcal{P}^{\mathbb{P}}(\Omega \times E, \mathcal{F} \otimes \mathcal{E})$. Then*

$$\liminf_{\varepsilon \rightarrow 0} \mathbb{E}[g(Z^\varepsilon)] \geq \mathbb{E}^{\mathbb{Q}}[g(\omega, z)], \tag{3.D.1}$$

for any g bounded from below with lower semi-continuous section $g(\omega, \cdot)$ on E .

3. Let Z be a random variable defined on $(\Omega, \mathcal{F}, \mathbb{P})$. We have

$$Z^\varepsilon \rightarrow_p Z \Leftrightarrow Z^\varepsilon \rightarrow_{stable} \mathbb{Q}^Z.$$

4. The sequence $\{\mathbb{Q}^{Z^\varepsilon}, \varepsilon > 0\}$ is relatively compact in $\mathcal{P}^{\mathbb{P}}(\Omega \times E, \mathcal{F} \otimes \mathcal{E})$ if and only if $\{Z^\varepsilon, \varepsilon > 0\}$ is relatively compact as subset of $\mathcal{P}(E)$. In particular, if E is compact, then $\mathcal{P}^{\mathbb{P}}(\Omega \times E, \mathcal{F} \otimes \mathcal{E})$ is compact.

Proof.

1. This is a direct consequence of [JM81, Proposition 2.4].

2. This is generalization of the Portmanteau theorem, see [JM81, Proposition 2.11].

3. The \Rightarrow implication is obvious. Let us prove the other. Consider $F(\omega, z) = |Z(\omega) - z| \wedge 1$. On the one hand, we have $\mathbb{E}[F(\omega, Z^\varepsilon)] \rightarrow \mathbb{E}^{\mathbb{Q}^Z}[F(\omega, z)] = 0$ by definition. On the other hand, for any $\delta \in (0, 1)$ we have

$$\mathbb{P}[|Z^\varepsilon - Z| > \delta] \leq \mathbb{E}[F(\omega, Z^\varepsilon) > \delta] \leq \frac{\mathbb{E}[F(\omega, Z^\varepsilon)]}{\delta},$$

by Markov inequality. We deduce that $Z^\varepsilon \rightarrow_p Z$.

4. See [JM81, Theorem 3.8 and Corollary 3.9]. \square

Lemma 3.D.1. Let $\{Z, Z^\varepsilon, \varepsilon > 0\}$ be positive random variables on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. If, for any random variable Y defined on $(\Omega, \mathcal{F}, \mathbb{P})$ with $c_Y \leq Y \leq C_Y$ where c_Y, C_Y are positive constants depending on Y ,

$$\liminf_{\varepsilon \rightarrow 0} \mathbb{E}[YZ^\varepsilon] \geq \mathbb{E}[YZ],$$

then

$$\liminf_{\varepsilon \rightarrow 0} Z^\varepsilon \geq_p Z.$$

Proof. Let $\delta > 0$ be any real number and, without loss of generality, let $\{Z^\varepsilon\}$ be a minimizing sequence of $\mathbb{P}[Z^\varepsilon > Z - \delta]$ as $\varepsilon \rightarrow 0$. Considering the one-point compactification $\mathbb{R}_+ \cup \{\infty\}$, we can assume that Z^ε converge stably to $\mathbb{Q} \in \mathcal{P}(\Omega \times (\mathbb{R}_+ \cup \{\infty\}))$ with canonical realization \bar{Z} . Then we have

$$\mathbb{E}[Y\bar{Z}] \geq \limsup_{\varepsilon \rightarrow 0} \mathbb{E}[YZ^\varepsilon] \geq \mathbb{E}[YZ],$$

where the first inequality comes from the fact that $z \mapsto z$ is u.s.c. on $\mathbb{R}_+ \cup \{\infty\}$ and [JM81, Prop 2.11]. Since Y is arbitrary, we conclude that $\bar{Z} \geq Z$. Then by stable convergence of Z^ε to \bar{Z} , we have $\mathbb{P}[Z^\varepsilon > Z - \delta] \rightarrow \mathbb{P}[\bar{Z} > Z - \delta] = 1$. \square

Chapter 4

Feedback strategies and applications

Contents

4.1	Introduction	137
4.2	Combined regular and impulse control	138
4.2.1	Feedback strategies	139
4.2.2	Asymptotic performance	141
4.3	Extensions to other types of control	143
4.3.1	Combined regular and singular control	143
4.3.2	When only one control is present	145
4.4	Applications	147
4.4.1	Explicit optimal strategies in dimension one	147
4.4.2	Discretization of hedging strategies	148
4.4.3	Impacts of small market frictions	149
4.5	Proofs	152
4.5.1	Proof of Theorem 4.2.1	152
4.5.2	Proof of Theorem 4.3.1	154
4.5.3	Proof of Theorem 4.3.4	155
Appendix 4.A	Convergence of integral functionals	156
Appendix 4.B	Separability of (A, B)	158
4.B.1	Diffusion with Rebirth	159
4.B.2	Diffusion with Reflection	160

4.1 Introduction

In Chapter 3, we have formulated the problem of tracking as a stochastic control problem minimizing both running deviation costs and intervention efforts. Finding optimal control policies for such systems, which have general non-Markovian dynamics and complex cost structures, is in general infeasible. Under a suitable asymptotic framework, we have established rigorously a lower bound for the best achievable asymptotic performance.

As we have shown heuristically, the lower bound is related to the time-average control of Brownian motion. It is then natural to try to construct tracking policies that are (near-)optimal by suitably adapting the solution of time-average control of Brownian Motion. We will show that this is indeed possible when the latter is available. However, closed-form solution of time-average control of Brownian motion is in general not available either, and numerical solutions are time-consuming to obtain. From a practical viewpoint, it is also irrelevant to find *the* optimal

strategy since there are already many approximations in the model.

As a continuation of Chapter 3, we study a class of feedback strategies which is easy to implement in practice and show how to determine their performance. Here, “feedback” means that the control decision is Markovian, depending only on the current state of the system. We establish in this way an upper bound for the best achievable performance and expect the two bounds to be tight.

Another main contribution of the current paper is to review some interesting asymptotic results in the literature under a unifying framework. First, we provide a new limit theorem for the discretization errors of stochastic integrals. Comparing with [HM05, Roo80, Fuk11c], we establish formally a new relation for different discretization schemes based on time or space. We also revisit some recent results on the asymptotic optimal discretization of hedging strategies ([Fuk11a, GL14a]) and relate their lower bounds to the time-average control of Brownian motion. Our approach allows more general cost functions in high dimension and we provide several new explicit examples in dimension one. The second application is the impact of market frictions such as transaction costs ([KMK15, KL13, ST13, PST15, AMKS15, LMKW14]), temporary market impact ([MMKS14, BSV15]) and illiquidity cost ([RS10, NW11]). We provide an alternative probabilistic approach for the asymptotics of small market frictions which have been extensively studied in the literature in the viscosity approach initiated by [ST13] under Markovian setting or duality method in [KMK15, KL13]. By a suitable choice of deviation cost, we establish a correspondence between the formulas in the above references and our asymptotic lower bounds, see Fig.4.2.

Notation. We denote the graph of an application $M : E \rightarrow E'$ by M^g . $\Delta = \{\gamma \in \mathbb{R}^d \mid \sum_i |\gamma^i| = 1\}$ is the simplex in \mathbb{R}^d . Let (E, d) be a complete separable metric space. $\mathcal{M}(E)$ denotes the space of Borel measures on E equipped with weak topology and $\mathcal{P}(E)$ denotes the subspace of probability measures. The hausdorff distance on the space of closed sets is given by

$$H(A, B) = \inf\{\delta, B \subseteq V_\delta(A) \text{ and } A \subseteq V_\delta(B)\},$$

where $V_\delta(\cdot)$ denotes the δ -neighbourhood of a closed set, i.e. $V_\delta(A) = \{x \in E, d(x, A) \leq \delta\}$.

The cost functions D, Q, F, P verify the following homogeneity property

$$D(\varepsilon x) = \varepsilon^{\zeta_D} D(x), \quad Q(\varepsilon u) = \varepsilon^{\zeta_Q} Q(u), \quad F(\varepsilon \xi) = \varepsilon^{\zeta_F} F(\xi), \quad P(\varepsilon \xi) = \varepsilon^{\zeta_P} P(\xi),$$

with $\zeta_D > 0$, $\zeta_Q > 1$, $\zeta_F = 0$, $\zeta_P = 1$. For example, we could take

$$D(x) = x^T \Sigma^D x, \quad Q(u) = u^T \Sigma^Q u, \quad F(\xi) = \sum_{i=1}^d F_i \mathbb{1}_{\{\xi^i \neq 0\}}, \quad P(\xi) = \sum_{i=1}^d P_i |\xi^i|,$$

with $\min_i F_i > 0$ and $\Sigma^D, \Sigma^Q \in \mathcal{S}_+^d$ such that $\zeta_D = \zeta_Q = 2$. For the rest of the notation, we refer the reader to Chapter 3.

4.2 Combined regular and impulse control

We focus on the case of combined regular and impulse control. Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ be a filtered probability space and (X_t°) an Itô diffusion such that

$$dX_t^\circ = b_t dt + \sqrt{a_t} dW_t,$$

with (W_t) a d -dimensional Brownian motion and (b_t) , (a_t) predictable processes with values in \mathbb{R}^d and \mathcal{S}_+^d respectively. The deviation (X_t) from the target is given by

$$X_t = -X_t^\circ + \int_0^t u_s ds + \sum_{j:0 < \tau_j \leq t} \xi_j,$$

where (u_t) is a progressively measurable process, $\{\tau_j\}$ a sequence of stopping times and $\{\xi_j\}$ corresponding \mathcal{F}_{τ_j} -measurable random variables. The process (u_t) represents the speed of the tracker. The stopping time τ_j represents the timing of the j -th instantaneous jump towards the target and ξ_j the size of the jump.

4.2.1 Feedback strategies

Motivated by various results in the literature and by practice, we consider a class of *feedback strategies* which take the following form. Let G be a domain in \mathbb{R}^d , $x \mapsto x + \xi(x)$ an application from the boundary ∂G to G , and u a bounded application from G to \mathbb{R}^d . When the deviation X_t is inside the domain G , only the regular control is active and the tracker maintains a speed of $u(X_t)$. When X_t touches the boundary ∂G at time τ , a jump of size $\xi(X_{\tau-})$ towards the interior of G happens and the process X_t begins with $X_{\tau-} + \xi(X_{\tau-})$ at time τ . Denote the resulting process by $X^{(u,G,\xi)}$. Let $\{\tau_j^{(u,G,\xi)}\}$ be the hitting times of $X^{(u,G,\xi)}$ and $\{Y_j^{(u,G,\xi)}\}$ the corresponding chain on the boundary ∂G where

$$Y_j^{(u,G,\xi)} = X_{\tau_j^{(u,G,\xi)}-}^{(u,G,\xi)}, \quad j = 1, 2, \dots,$$

then we have

$$X_t^{(u,G,\xi)} = -X_t^\circ + \int_0^t u(X_s^{(u,G,\xi)}) ds + \sum_{0 < \tau_j^{(u,G,\xi)} \leq t} \xi(Y_j^{(u,G,\xi)}).$$

Recall that the asymptotic framework of tracking is to consider the sequence of optimization problems indexed by $\varepsilon \rightarrow 0$ with

$$J^\varepsilon(u^\varepsilon, \tau^\varepsilon, \xi^\varepsilon) = \int_0^T (r_t D(X_t^\varepsilon) + \varepsilon^{\beta_Q} l_t Q(u_t^\varepsilon)) dt + \sum_{j:0 < \tau_j^\varepsilon \leq T} (\varepsilon^{\beta_F} k_{\tau_j^\varepsilon} F(\xi_j^\varepsilon) + \varepsilon^{\beta_P} h_{\tau_j^\varepsilon} P(\xi_j^\varepsilon)), \quad (4.2.1)$$

where

$$X_t^\varepsilon = -X_t^\circ + \int_0^t u_s^\varepsilon ds + \sum_{j:0 < \tau_j^\varepsilon \leq t} \xi_j^\varepsilon,$$

and β_F, β_P some fixed real numbers.

Under our asymptotic setting, an appropriate feedback strategy should depend on time. More precisely, let (G_t) be a random moving open bounded domain associated with jump rule (ξ_t) on the boundary such that $x + \xi_t(x) \in C^0(\partial G_t, G_t)$, and (u_t) be a random function on the closure \bar{G}_t such that $u_t \in C^0(\bar{G}_t, \mathbb{R}^d)$. The triplet (u_t, G_t, ξ_t) is continuous if $(u_t^g, \partial G_t, \xi_t^g)$ is continuous as closed set-valued processes w.r.t. the Hausdorff distance H . We say that (u_t, G_t, ξ_t) is progressively measurable, if $(u_t^g, \partial G_t, \xi_t^g)$ is progressively measurable w.r.t. (\mathcal{F}_t) as closed set-valued processes (see [Kis13] for more details).

We will consider (u_t, G_t, ξ_t) which is continuous and progressively measurable. Intuitively, we require that the data (u_t, G_t, ξ_t) is determined in a non-anticipative way, based on information up to time t , and does not vary too much in time. Note that since ∂G_t is continuous, the

topology of the domain G_t (number of holes, boundedness, etc.) remains the same.

The sequence of feed-back strategies $(X^\varepsilon, u^\varepsilon, \tau^\varepsilon, \xi^\varepsilon)$ can be constructed in the following recursive way.

1. Let $\tau_0^\varepsilon = 0, X_0^\varepsilon = 0$.
2. For $t \geq \tau_{j-1}^\varepsilon$, let X_t^ε be defined by

$$dX_t^\varepsilon = -dX_t^\circ + u_t^\varepsilon dt, \quad (4.2.2)$$

with

$$u_t^\varepsilon = \varepsilon^{-(\alpha-1)\beta} u_t(\varepsilon^{-\beta} X_t^\varepsilon).$$

3. Put

$$\tau_j^\varepsilon = \inf\{t > \tau_{j-1}^\varepsilon, \frac{1}{\varepsilon^\beta} X_t^\varepsilon \notin G_t\}, \quad \xi_j^\varepsilon = \varepsilon^\beta \xi_{\tau_j^\varepsilon}(\varepsilon^{-\beta} X_{\tau_j^\varepsilon}^\varepsilon),$$

and

$$X_{\tau_j^\varepsilon}^\varepsilon = X_{\tau_j^\varepsilon-}^\varepsilon + \xi_j^\varepsilon.$$

4. Go to step 2.

The boundary chain $\{Y_j^\varepsilon\}$ is defined by

$$Y_j^\varepsilon = \varepsilon^{-\beta} X_{\tau_j^\varepsilon-}^\varepsilon \in \partial G_{\tau_j^\varepsilon}, \quad j = 1, 2, \dots. \quad (4.2.3)$$

We assume that

Assumption 4.2.1. *The feedback strategy $(X^\varepsilon, u^\varepsilon, \tau^\varepsilon, \xi^\varepsilon)$ exists and is unique for each $\varepsilon > 0$.*

A rigorous verification of the above definition requires detailed analysis with specific conditions on (G_t, ξ_t, u_t) . We provide below a simple situation where the above assumption is verified.

Lemma 4.2.1. *Let (u_t) be a random function with uniformly Lipschitz coefficient $K > 0$, and assume that the distance to the boundary after the jump is bounded from below by $\delta > 0$, i.e.*

$$|u_t(x) - u_t(y)| \leq K|x - y|, \quad \forall x, y \in \mathbb{R}_x^d, \forall (t, \omega), \quad (4.2.4)$$

and

$$H(\xi_t(\partial G_t), \partial G_t) > \delta, \quad \forall (t, \omega). \quad (4.2.5)$$

Then Assumption 4.2.1 holds. Moreover, it continues to hold if (4.2.4)-(4.2.5) are true up to a localization procedure.

Proof. For the existence of (4.2.2) (in the strong sense), Lipschitz-type regularity on u_t is sufficient (see [Pro04, Chapter V] for more details). For the existence of X^ε on the whole horizon $[0, T]$, it suffices to show that

$$\lim_{j \rightarrow \infty} \tau_j^\varepsilon = +\infty,$$

the proof of which is the same as in [CK06, Proof of Theorem 4.1] (see also [GK12]). □

4.2.2 Asymptotic performance

In order to have well-behaved strategies, we restrict ourselves to the following class of admissible triplets (u_t, G_t, ξ_t) .

Definition 4.2.1 (Admissible Strategy). *The triplet (u_t, G_t, ξ_t) is admissible if*

1. (Potential.) *There exists $V \in C^2(\mathbb{R}^d, \mathbb{R})$ such that $\forall(t, \omega)$,*

$$V(x + \xi_t^\omega(x)) - V(x) < 0, \quad \forall x \in \partial G_t^\omega.$$

We call V a potential function of the jump rule (G_t, ξ_t) .

2. (Separability.) *For any (t, ω) , let $u = u_t^\omega$, $G = G_t^\omega$ and $\xi = \xi_t^\omega$, then there exists a unique couple $(\pi, \nu) \in \mathcal{P}(\bar{G}) \times \mathcal{M}(\partial G)$ verifying the constraints*

$$\int_{\bar{G}} A_u^a f(x, u) \pi(dx) + \int_{\partial G} B_\xi f(x) \nu(dx) = 0, \quad \forall f \in C_0^2(\mathbb{R}^d), \quad (4.2.6)$$

where

$$A_u^a f(x) = \frac{1}{2} \sum_{i,j} a_{ij} \partial_{ij}^2 f(x) + u(x) \cdot \nabla f(x), \quad B_\xi f(x) = f(x + \xi(x)) - f(x).$$

We note $\pi =: \pi^{(a,u,G,\xi)}$ and $\nu =: \nu^{(a,u,G,\xi)}$ to indicate the dependence on (a, u, G, ξ) .

Remark 4.2.1. *The existence of V ensures that all impulse controls try to constrain the process X^ε in the same direction and hence avoid the explosion of tracking efforts (last term on the left hand side of (4.2.1)) due to the cancellation of different jumps.*

Loosely speaking, π and ν are, up to a constant, stationary distributions over the horizon $(t, t + \delta^\varepsilon]$ of X^ε and Y^ε respectively. The separability of (A, B) w.r.t. (u, G, ξ) guarantees uniqueness of these stationary distributions, which is a desired property when we consider the limit under the asymptotic framework.

For a numerical computation of (π, ν) based on Monte-Carlo simulation, see [BCR13] for a possible approach.

Usually, the barrier G_t and jump rule ξ are defined by prescribing a potential function V as the following example shows.

Example 4.2.1. *Let α_t, β_t be a continuous positive adapted processes such that $\alpha_t > \beta_t$ and $V(x) = cx^2$ with $c > 0$. Define $G_t = \{x \in \mathbb{R}, V(x) < \alpha_t\}$, we have*

$$G_t = [-L_t, L_t], \quad L_t = \sqrt{\frac{\alpha_t}{c}}.$$

Take

$$\xi_t(\pm L_t) = \pm l_t, \quad l_t = \sqrt{\frac{\beta_t}{c}}.$$

Then the triplet (u_t, G_t, ξ_t) is admissible.

Indeed, the function V can serve as a potential function for the jump rule. It suffices to show the separability of (π, ν) , which can be written as, $\forall f \in C_0^2(\mathbb{R})$,

$$\int_{-L}^L \left[\frac{1}{2} a f''(x) + u(x) f'(x) \right] \pi(dx) + \nu_+ [f(l) - f(L)] + \nu_- [f(-l) - f(-L)] = 0,$$

for some $\pi \in \mathcal{P}([-L, L])$ and $\nu_{\pm} \in \mathbb{R}^+$. Let $f(x; g)$ be a solution of

$$\frac{1}{2}af''(x) + u(x)f'(x) = g(x), \quad x \in [-L, L].$$

Let $U(x) = \int u(x)dx$, we can solve explicitly

$$f(x; g) = \int_0^x \int_0^z g(y)e^{-\frac{2}{a}(U(z)-U(y))} dydz + C_1 \int_0^x e^{-\frac{1}{a}U(z)} dz + C_2, \quad x \in [-L, L],$$

which can be extended to a function in $C_0^2(\mathbb{R})$ and is denoted also by $f(x; g)$. Take $g \equiv 0$ and $g \equiv 1$ on $[-L, L]$, we obtain

$$\begin{cases} \nu_+(f(l; 0) - f(L; 0)) + \nu_-(f(-l; 0) - f(-L; 0)) = 0, \\ \nu_+(f(l; 1) - f(L; 1)) + \nu_-(f(-l; 1) - f(-L; 1)) = -1, \end{cases}$$

which admits a unique solution (ν_+, ν_-) . Hence ρ is uniquely determined. Then π is also uniquely determined since we can take $g|_{[-L, L]} \in C^0([-L, L])$ arbitrarily.

Let $\rho(dx \times d\xi) = \nu(dx) \otimes \delta_{\xi(x)}(d\xi)$, then (π, ν) is given by (3.7.18). In particular, if $u_t \equiv 0$ then (π, ρ) is explicitly given by (3.4.28)-(3.4.29), which corresponds to the stationary distribution and boundary measures of Brownian motion with rebirth.

Remark 4.2.2. In high dimension, we can define a similar strategy where $V = x^T \Sigma x$ and $\Sigma \in \mathcal{S}_+^d$. Then G_t is an ellipsoid given by $G_t = \{x \in \mathbb{R}^d | V(x) < \alpha_t\}$. In that case, we have more freedom for the choice of jump rule ξ_t . For example, the jump rule $\xi(x) = -x$ corresponds to the strategies in [GL14a, AMKS15]. However, the uniqueness of (π, ν) can't be handled as in dimension one. In Appendix 4.B, we propose an alternative approach which reduces the uniqueness of (π, ν) to the ergodicity of the corresponding diffusion with rebirth, under the hypothesis that the latter is well-defined.

Now we state one of our main results in this Chapter.

Theorem 4.2.1 (Asymptotic performance for combined regular and impulse control). Consider an admissible triplet (u_t, G_t, ξ_t) and assume Assumption 4.2.1. Let $(X^\varepsilon, u^\varepsilon, \tau^\varepsilon, \xi^\varepsilon)$ be the corresponding feedback strategy, we have

$$\frac{1}{\varepsilon \zeta_D \beta} J^\varepsilon(u^\varepsilon, \tau^\varepsilon, \xi^\varepsilon) \rightarrow_p \int_0^T c(a_t, u_t, G_t, \xi_t; r_t, l_t, k_t, h_t) dt,$$

where $c(a, u, G, \xi; r, l, k, h)$ is given by

$$\begin{aligned} c(a, u, G, \xi; r, l, k, h) &= \int_{\bar{G}} [rD(x) + lQ(u(x))] \pi^{(a, u, G, \xi)}(dx) \\ &\quad + \int_{\partial G} [kF(\xi(x)) + hP(\xi(x))] \nu^{(a, u, G, \xi)}(dx). \end{aligned}$$

Moreover, the convergence holds term by term for the cost functions D, Q, F, P respectively.

Remark 4.2.3. In the theorem above, we need Assumption 4.2.1 and 4.B.1 to hold. While these assumptions can be verified directly in dimension one, there is not documented result for Assumption 4.B.1 in high dimension yet (See however [CK15, Remark 3.2 and 5.3] for an approach based on viscosity solution).

Corollary 4.2.1. If $\{\varepsilon^{-\zeta_D \beta} J^\varepsilon, \varepsilon > 0\}$ is uniformly integrable, we have

$$\lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon \zeta_D \beta} \mathbb{E} [J^\varepsilon] = \mathbb{E} \left[\int_0^T c(a_t, u_t, G_t, \xi_t; r_t, l_t, k_t, h_t) dt \right].$$

Remark 4.2.4. *The deviation cost is easily bounded by the barriers hence uniformly integrable if the latter is integrable. For the uniform integrability of tracking cost, see Assumptions (HA₂) and (HA'₂) in [RT14] for a particular jump rule.*

Recall that we have established a lower bound of J^ε in terms of I , where $I = I(a, r, l, k, h)$ is given by

$$I = \inf_{(\mu, \rho)} \int_{\mathbb{R}_x^d \times \mathbb{R}_u^d} (rD(x) + lQ(u))\mu(dx \times du) + \int_{\mathbb{R}_x^d \times \mathbb{R}_\xi^d} (kF(\xi) + hP(\xi))\rho(dx \times d\xi),$$

with $(\mu, \rho) \in \mathcal{P}(\mathbb{R}_x^d \times \mathbb{R}_u^d) \times \mathcal{M}(\mathbb{R}_x^d \times \mathbb{R}_\xi^d)$ verifying the following constraints

$$\int_{\mathbb{R}_x^d \times \mathbb{R}_u^d} A^a f(x, u)\mu(dx \times du) + \int_{\mathbb{R}_x^d \times \mathbb{R}_\xi^d} Bf(x, \xi)\rho(dx \times d\xi) = 0, \quad \forall f \in C_0^2(\mathbb{R}_x^d).$$

If we can find (u_t, G_t, ξ_t) such that $c(u_t, G_t, \xi_t) = I_t$, then the lower bound (3.1.6) is sharp.

Corollary 4.2.2 (Definition of asymptotic optimal strategies). *Assume that r_t, l_t, k_t and h_t are adapted and there exists an admissible triplet (u_t^*, G_t^*, ξ_t^*) such that $dt \otimes d\mathbb{P}$ -a.e. (t, ω) ,*

$$c(a_t, u_t^*, G_t^*, \xi_t^*; r_t, l_t, k_t, h_t) = I(a_t, r_t, l_t, k_t, h_t).$$

Then we have

$$\lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon^{\zeta_D \beta}} J^\varepsilon(u^{*,\varepsilon}, \tau^{*,\varepsilon}, \xi^{*,\varepsilon}) \rightarrow_p \int_0^T I(a_t, r_t, l_t, k_t, h_t) dt,$$

and we say that the feed-back strategy $(u^{*,\varepsilon}, \tau^{*,\varepsilon}, \xi^{*,\varepsilon})$ based on (u_t^*, G_t^*, ξ_t^*) is asymptotically optimal.

The consequence of this section is twofold. First, we will be able to verify that the lower bounds (3.1.6) established previously in Examples 4.3.2-3.4.7 are in fact tight and can be obtained by feedback strategies when closed-form solution I_t of time-average control of Brownian motion is available. Second, we provide an asymptotic framework for the evaluation of a large class of feedback strategies which is sufficient for practical purpose. Indeed, to show that the lower bound is tight, one needs explicit information about the solution of the time-average control of Brownian motion, which is very difficult to obtain in general. Hence the lower bound will serve only as a *theoretical benchmark* of performance and in practice it is sufficient to compare simple implementable strategies.

4.3 Extensions to other types of control

4.3.1 Combined regular and singular control

Without the presence of fixed cost component, the tracking problem is to consider a strategy $(u^\varepsilon, \gamma^\varepsilon, \varphi^\varepsilon)$ with u^ε a progressively measurable process, $\gamma_t^\varepsilon \in \Delta = \{\gamma \in \mathbb{R}^d \mid \sum_i |\gamma^i| = 1\}$ and φ_t^ε non-decreasing such that

$$X_t^\varepsilon = -X_t^o + \int_0^t u_s^\varepsilon ds + \int_0^t \gamma_s^\varepsilon d\varphi_s^\varepsilon,$$

and

$$J^\varepsilon(u^\varepsilon, \gamma^\varepsilon, \varphi^\varepsilon) = \int_0^T (r_t D(X_t^\varepsilon) + \varepsilon^{\beta_Q} l_t Q(u_t^\varepsilon)) dt + \int_0^T \varepsilon^{\beta_P} h_t P(\gamma_t^\varepsilon) d\varphi_t^\varepsilon.$$

Besides the class of impulse feed-back strategies based on hitting times of moving barriers and ‘‘Markov control’’ (G_t, ξ_t, u_t) in the previous section, we can also consider another class of strategies based on continuous control of local time type.

More precisely, let G_t be a moving domain with piecewise C^2 boundary and u_t a continuous Markov control defined on \bar{G}_t as before, let Γ_t be a set-valued mapping defined on ∂G_t with *closed graph* such that $\Gamma_t(x)$ is a non-empty closed convex cone in \mathbb{R}^d with vertex at the origin 0 for each $x \in \partial G_t$. Usually $\Gamma_t(x)$ contains only a single direction on the smooth part of ∂G . If ∂G is C^2 then $\Gamma_t(x)$ can also be expressed in terms of a vector field on ∂G . We assume that the triplet (u_t, G_t, Γ_t) is continuous and progressively measurable.

The feedback strategy $(X^\varepsilon, u^\varepsilon, \gamma^\varepsilon, \varphi^\varepsilon)$ based on (u_t, G_t, Γ_t) is to keep X_t^ε in the domain $\varepsilon^\beta \bar{G}_t$ using singular control in the direction Γ_t on the boundary and Markov control u_t inside the domain, i.e.

$$X_t^\varepsilon = -X_t^o + \int_0^t u_s^\varepsilon ds + \int_0^t \gamma_s^\varepsilon d\varphi_s^\varepsilon,$$

with $\gamma_t^\varepsilon \in \Gamma_t$, φ^ε continuous non decreasing and

$$u_t^\varepsilon = \varepsilon^{-(\alpha-1)\beta} u_t(\varepsilon^{-\beta} X_t^\varepsilon),$$

such that

1. $\varepsilon^{-\beta} X_t^\varepsilon \in \bar{G}_t$,
2. $\gamma_t^\varepsilon \in \Gamma_t(\varepsilon^{-\beta} X_t^\varepsilon) \cap \Delta$, $d\varphi_t$ -almost surely,
3. $\int_0^T \mathbb{1}_{G_t}(\varepsilon^{-\beta} X_t^\varepsilon) d\varphi_t^\varepsilon = 0$.

Note that $\gamma_t^\varepsilon \in \Delta$ allows us to distinguish γ_t^ε and φ_t^ε . We assume again that

Assumption 4.3.1. *The feedback strategy $(X^\varepsilon, u^\varepsilon, \gamma^\varepsilon, \varphi^\varepsilon)$ exists and is unique for each $\varepsilon > 0$.*

While the existence of above strategy is true in dimension one (see [SW13]), it is not at all trivial in higher dimension.

Remark 4.3.1. *The existence of such strategies is closely related to the Skorohod oblique reflection problem in time-dependent domain, we refer the reader to [NÖ10] for precise conditions on (G_t, Γ_t) . In particular, using the numerations in [NÖ10], it suffices to check that (i) The domain G_t satisfies (1.2), (1.10) and a uniform exterior sphere condition in time in the sense of (1.8). Note that in general, the condition on the temporal regularity is only C^0 . (ii) The cone of reflection Γ_t verifies (1.11), (1.12) and (1.14). (iii) (1.18) and (1.19) hold (cf. [Cos92, Proposition 2.5]).*

However, the result in [NÖ10] is not sufficient here since we need to ensure the existence a strong solution for each $\varepsilon > 0$ supported on the same probability space $(\Omega, \mathbb{P}, \mathcal{F})$. Nevertheless, we can avoid this difficulty by implementing the approximation procedure in [NÖ10] which consists of considering feedback strategies $(X^{\varepsilon,d}, u^{\varepsilon,d}, \gamma^{\varepsilon,d}, \varphi^{\varepsilon,d})$ based on piecewise constant data $(u_t^{\varepsilon,d}, G_t^{\varepsilon,d}, \Gamma_t^{\varepsilon,d})$.

From now on, we restrict ourselves to (u_t, G_t, Γ_t) for which the above strategies exist and are well-behaved.

Definition 4.3.1 (Admissible Strategy). *The triplet (u_t, G_t, Γ_t) is admissible if*

1. (Potential.) *There exists $V \in C^2(\mathbb{R}^d)$ such that $\forall(t, \omega)$,*

$$\langle \nabla V(x), \gamma \rangle < 0, \quad \forall x \in \partial G_t^\omega, \gamma \in \Gamma_t^\omega(x).$$

We say that V is a potential function for (G_t, Γ_t) .

2. (Separability.) For any (t, ω) , let $u = u_t^\omega$, $G = G_t^\omega$ and $\Gamma = \Gamma_t^\omega$, then there exists a unique couple $(\pi, \rho) \in \mathcal{P}(\bar{G}) \times \mathcal{M}(\Gamma_\Delta^g)$ such that

$$\int_{\bar{G}} A_u^a f(x) \pi(dx) + \int_{\Gamma_\Delta^g} Bf(x, \gamma) \rho(dx, d\gamma) = 0, \quad \forall f \in C_0^2(\mathbb{R}^d), \quad (4.3.1)$$

where $\Gamma_\Delta^g = \{(x, \gamma) | x \in \partial G, \gamma \in \Delta \cap \Gamma(x)\}$ and

$$A_u^a f(x) = \frac{1}{2} \sum_{ij} a_{ij} \partial_{ij}^2 f(x) + u(x) \cdot \nabla f(x), \quad Bf(x, \gamma, \delta) = \gamma \cdot \nabla f(x).$$

We note $\pi =: \pi^{(a, u, G, \Gamma)}$ and $\rho =: \rho^{(a, u, G, \Gamma)}$.

Below is an example in dimension one.

Example 4.3.1. Let α_t be a continuous positive adapted process and $V(x) = cx^2$ with $c > 0$. Put $G_t = \{x, V(x) < \alpha_t\}$, then

$$G_t = [-L_t, L_t], \quad L_t = \sqrt{\frac{\alpha_t}{c}}.$$

Take Γ such that

$$\Gamma_t(\pm L_t) = \{z \nabla V(\pm L_t) | z \geq 0\},$$

then the triplet (u_t, G_t, Γ_t) is admissible.

As for Example 4.2.1, we can take V as potential function and the proof for the separability is the same hence omitted. The pair (π, ρ) is given by (3.4.35)-(3.4.36). For the case where $u_t \equiv 0$, π and ρ are explicitly given by (3.4.24)-(3.4.25).

Remark 4.3.2. In high dimension, we can characterize (π, ρ) as the stationary distribution and boundary measure of the corresponding diffusion with reflection as in [KR14] under suitable regularity conditions on (u, G, Γ) and reduce the uniqueness of (π, ρ) to the ergodicity of the reflected diffusion, see Appendix 4.B.

Now we are ready to state another main result.

Theorem 4.3.1 (Asymptotic performance for combined regular and singular control). Let (u_t, G_t, Γ_t) be an admissible triplet and assume Assumption 4.3.1, we have

$$\frac{1}{\varepsilon \zeta_{D\beta}} J^\varepsilon(u^\varepsilon, \gamma^\varepsilon, \varphi^\varepsilon) \rightarrow_p \int_0^T c(a_t, u_t, G_t, \Gamma_t; r_t, l_t, h_t) dt,$$

where $c(a, u, G, \Gamma; r, l, h)$ is given by

$$\begin{aligned} c(a, u, G, \Gamma; r, l, h) &= \int_{\bar{G}} (rD(x) + lQ(u(x))) \pi^{(a, u, G, \Gamma)}(dx) \\ &\quad + \int_{\Gamma_\Delta^g} hP(\gamma) \rho^{(a, u, G, \Gamma)}(dx \times d\gamma). \end{aligned}$$

Moreover, the convergence holds term by term for the cost functions D , Q and P respectively.

4.3.2 When only one control is present

In this section, we extend Theorem 4.2.1 and 4.3.1 to the case where only one type of control is present. The cases with singular or impulse control only are included in the previous results. Only the case with regular control needs to be treated separately.

Impulse control

In that case, a feedback strategy is determined by a time-dependent domain G_t with a jumping rule ξ_t . Let $u_t = 0$ and $l = 0$ in Theorem 4.2.1, we obtain

Theorem 4.3.2 (Asymptotic performance for impulse control). *Consider an admissible triplet (G_t, ξ_t) and assume Assumption 4.2.1. Let $(X^\varepsilon, \tau^\varepsilon, \xi^\varepsilon)$ be the corresponding feedback strategy, we have*

$$\frac{1}{\varepsilon^{\zeta_D \beta}} J^\varepsilon(\tau^\varepsilon, \xi^\varepsilon) \rightarrow_p \int_0^T c(a_t, G_t, \xi_t; r_t, k_t, h_t) dt,$$

where $c(a, G, \xi; r, k, h)$ is given by

$$\begin{aligned} c(a, G, \xi; r, k, h) &= \int_{\bar{G}} r D(x) \pi^{(a, G, \xi)}(dx) \\ &\quad + \int_{\partial G} [kF(\xi(x)) + hP(\xi(x))] \nu^{(a, G, \xi)}(dx). \end{aligned}$$

Moreover, the convergence holds term by term for the cost functions D, F, P respectively.

Singular control

Similarly, let $u_t = 0$ and $l = 0$ in Theorem 4.3.1, we obtain

Theorem 4.3.3 (Asymptotic performance for combined regular and singular control). *Let (G_t, Γ_t) be an admissible triplet and assume Assumption 4.3.1, we have*

$$\frac{1}{\varepsilon^{\zeta_D \beta}} J^\varepsilon(\gamma^\varepsilon, \varphi^\varepsilon) \rightarrow_p \int_0^T c(a_t, G_t, \Gamma_t; r_t, h_t) dt,$$

where $c(a, G, \Gamma; r, l, h)$ is given by

$$\begin{aligned} c(a, G, \Gamma; r, h) &= \int_{\bar{G}} r D(x) \pi^{(a, G, \Gamma)}(dx) \\ &\quad + \int_{\Gamma_\Delta^g} h P(\gamma) \rho^{(a, G, \Gamma)}(dx \times d\gamma). \end{aligned}$$

Moreover, the convergence holds term by term for the cost functions D and P respectively.

Regular control

Let $u_t : \mathbb{R}^d \rightarrow \mathbb{R}^d$ be continuous predictable random function. The feedback strategy $(X_t^\varepsilon, u_t^\varepsilon)$ based on the Markov control policy u_t is given by

$$X_t^\varepsilon = -X_t^\circ + u_t^\varepsilon dt, \quad u_t^\varepsilon = \varepsilon^{-\beta} u_t(\varepsilon^{-\beta} X_t^\varepsilon).$$

Under suitable conditions on u_t , we have well-defined process $(X^\varepsilon, u^\varepsilon)$. Let us just assume that

Assumption 4.3.2. *The controlled process $(X^\varepsilon, u^\varepsilon)$ exists and is unique on $[0, T]$.*

To have well-behaved controlled process X^ε , we consider the following class of admissible strategies.

Definition 4.3.2 (Admissible strategy). *The feedback control (u_t) is admissible if*

1. (Potential.) *For each (t, ω) , $u_t^\omega : \mathbb{R}_x \rightarrow \mathbb{R}_u$ is locally bounded and there exists a non-negative, inf-compact Lyapunov function $V \in C^2(\mathbb{R}_x)$ such that*

$$(A^{a_t} + u_t \cdot \nabla)V(x) \leq \theta_t - 2\Theta_t V(x), \quad \forall x \in \mathbb{R}_x,$$

with θ_t, Θ_t positive processes. Moreover, $rD + lQ \circ u_t^\omega$ are dominated by V near infinity, i.e. there exist locally bounded positive processes R_t and b_t such that

$$rD(x) + lQ \circ u_t^\omega(x) \leq b_t^\omega V(x), \quad x \in \mathbb{R}^d \setminus B(x, R_t^\omega).$$

2. (Separability.) For each (t, ω) , $u := u_t^\omega : \mathbb{R}_x^d \rightarrow \mathbb{R}_u^d$ is a stable Markov, i.e. there exists a unique $\pi \in \mathcal{P}(\mathbb{R}_x)$ such that,

$$\int_{\mathbb{R}_x} A_u^a f(x) \pi(dx) = 0, \quad \forall f \in C_0^2(\mathbb{R}_x).$$

We will denote $\pi = \pi^{(a,u)}$.

Remark 4.3.3. Here π is the unique invariant measure of $dX_t = \sqrt{a}dW_t + u(X_t)dt$ under the stable control u .

Example 4.3.2. Assume that D and Q are quadratic. Let $u_t(x) = -\Sigma_t x$ with Σ_t a continuous process with values in \mathcal{S}_+^d . Then (u_t) is an admissible Markov policy. To see this, it suffices to take $V(x) = x^T x$. The separability of π is equivalent to the ergodicity of $dX_t = \sqrt{a}dW_t - \Sigma X_t dt$. In dimension one, let $D(x) = x^2$, $Q(u) = u^2$, then we have

$$\pi(dx) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{x^2}{2\sigma^2}} dx, \quad \sigma^2 = \frac{a}{2\Sigma}.$$

Below is one of our main results.

Theorem 4.3.4 (Asymptotic performance for regular control). Let (u_t^ε) be the feed-back strategy based on the admissible Markov control policy u_t , then under the Assumptions 4.3.2, we have

$$\lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon \zeta_{D\beta}} J^\varepsilon \rightarrow_p \int_0^T c(a_t, u_t; r_t, l_t) dt,$$

where $c(a, u; r, l)$ is given by

$$c(a, u; r, l) = \int_{\mathbb{R}_x} (rC(x) + lQ \circ u(x)) \pi(dx),$$

with $\pi = \pi^{(a,u)}$.

4.4 Applications

4.4.1 Explicit optimal strategies in dimension one

Before reviewing and extending existing results in the literature, we summarize in the following theorem some explicit asymptotic optimal strategies of the tracking problem. The costs for deviation and regular control are quadratic and the costs for singular controls are affine or linear, i.e.

$$C_A(x, u) = rx^2 + lu^2, \quad C_B(x, \xi) = k + h|\xi|.$$

Theorem 4.4.1. Assume that (r_t, l_t, k_t, h_t) are progressively measurable. In dimension one, the asymptotic lower bounds for the tracking problem with combined/single controls are tight and can be achieved by the admissible feedback strategies defined in Fig.4.1 and Fig.4.2.

Proof. The existence and admissibility of the corresponding feedback strategies being easy to verify in dimension one (Example 4.2.1, 4.3.1 and 4.3.2), we can apply Theorem 4.2.1, 4.3.1-4.3.4, and Theorem 3.2.1, 3.3.1-3.3.4. \square

4.4.2 Discretization of hedging strategies

Let X° be the benchmark strategy in continuous time (e.g. delta strategy under Black-Scholes model) and the underlying asset (S_t) be an Ito diffusion with

$$dS_t = b_t^S dt + \sqrt{a_t^S} dW_t.$$

Then the deviation from the benchmark following the discretization strategy $(\tau^\varepsilon, \xi^\varepsilon)$ is given by

$$X_t^\varepsilon = -X_t^\circ + \sum_{0 < \tau_j^\varepsilon \leq t} \xi_j^\varepsilon.$$

and the corresponding d -dimensional hedging error (Z_t^ε) by

$$Z_t^\varepsilon = \int_0^t \text{Diag}(X_{t-}^\varepsilon) dS_t.$$

Our results apply directly to the problem of discrete hedging if the discretization is due to the presence of small fixed transaction cost $\varepsilon^{\beta_F} k_t$ and proportional cost $\varepsilon^{\beta_P} h_t$ with $\varepsilon \rightarrow 0$. To see this, put

$$J^\varepsilon = \langle Z^\varepsilon \rangle_T + \sum_{0 < \tau_j^\varepsilon \leq T} (\varepsilon^{\beta_F} k_{\tau_j^\varepsilon} + \varepsilon^{\beta_P} h_{\tau_j^\varepsilon} |\xi_j^\varepsilon|),$$

where

$$\langle Z^\varepsilon \rangle_t := \int_0^t (X_s^\varepsilon)^T a_s^S X_s^\varepsilon ds.$$

Here the tracking error is represented by the quadratic variation of the deviation process Z^ε . We can read the optimal strategies from Figure 4.2.

Relation with other asymptotic hedging criteria Different authors (see [Fuk11a, GL14a, RT14]) have considered the optimal discretization of a hedging strategy X_t° under various criteria and asymptotic frameworks. Our framework produces similar results and provides more flexible extensions. For example, in [Fuk11a], the problem is to minimize

$$\mathbb{E}[\langle Z^\varepsilon \rangle_T] \mathbb{E}[N_T^\varepsilon],$$

with N_T^ε the total number of rebalancing. It is shown that, under one dimensional setting,

$$\liminf_{\varepsilon \rightarrow 0} \mathbb{E}[\langle Z^\varepsilon \rangle_T] \mathbb{E}[N_T^\varepsilon] \geq \left(\frac{1}{2} \mathbb{E} \left[\frac{1}{\sqrt{6}} \int_0^T \left(\frac{2}{3} a_t^X a_t^S \right)^{1/2} dt \right]^2 \right). \quad (4.4.1)$$

We can recover above results of [Fuk11a]. To this aim, we consider instead

$$J^\varepsilon = \langle Z^\varepsilon \rangle_T + \varepsilon N_T^\varepsilon.$$

Lemma 4.4.1. *The asymptotic lower bound (4.4.1) is equivalent to*

$$\liminf_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon^{1/2}} \mathbb{E}[J^\varepsilon] \geq \mathbb{E} \left[\int_0^T \left(\frac{2}{3} a_t^X a_t^S \right)^{1/2} dt \right].$$

which is given by the third column of Figure 4.2.

Proof. Note that the sequence of ε in the above lower bound is arbitrary. Assuming there is a strategy such that

$$\liminf_{\varepsilon \rightarrow 0} \mathbb{E}[\langle Z^\varepsilon \rangle_T] \mathbb{E}[N_T^\varepsilon] < \left(\frac{1}{2} \mathbb{E}[J^* - \delta] \right)^2.$$

with $\delta > 0$, then we have $\tilde{\varepsilon}(\varepsilon) = \mathbb{E}[\langle Z^\varepsilon \rangle_T]^{1/2} \mathbb{E}[N_T^\varepsilon]^{-1/2} \rightarrow 0$ and we have for the same strategy,

$$\liminf_{\tilde{\varepsilon} \rightarrow 0} \frac{1}{\tilde{\varepsilon}^{1/2}} J^{\tilde{\varepsilon}} = \liminf_{\tilde{\varepsilon} \rightarrow 0} 2 (\mathbb{E}[\langle Z^\varepsilon \rangle_T] \mathbb{E}[N_T^\varepsilon])^{1/2} < \mathbb{E}[J^* - \delta].$$

The other direction is obvious by Cauchy-Schwartz inequality. \square

Applying Theorem 4.2.1, we obtain that the lower bound (4.4.1) is attained by the discretization scheme based on hitting times of barriers in [Fuk11a].

Weak Convergence of hedging errors We present weak convergence theorem closely related to the results in [HM05] and [Fuk11c].

Theorem 4.4.2 (Stable limit in law). *Assume that (b_t^S) and (a_t^S) are continuous on $[0, T]$ and let (Z_t^ε) be the error processes based on the moving barrier (G_t, ξ_t) and $\beta = 1$. We have*

$$\frac{1}{\varepsilon} Z_t^\varepsilon \rightarrow_{\mathcal{F}} \int_0^\cdot \text{Diag}(m_t) dS_t + \int_0^\cdot \sqrt{v_t \cdot a_t^S - m_t^T a_t^S m_t^T} dW_t^*,$$

where the convergence holds stably in law in $C[0, T]$, (W_t^*) is a Brownian motion independent of \mathcal{F} and $m_t = (m_t^1, \dots, m_t^d)$, $v_t = (v_t^{ij})$ with

$$m_t^i = \int_{G_t} x^i d\mu(a_t, G_t, \xi_t), \quad v_t^{ij} = \int_{G_t} x^i x^j d\mu(a_t, G_t, \xi_t),$$

and $v_t \cdot a_t^S := (v_t^{ij} a_{ij,t}^S)$.

Proof. After [JS13, Theorem IX.7.3], it suffices to show that

$$\begin{aligned} \frac{1}{\varepsilon} \int_0^\cdot (Z_t^\varepsilon)^T b_t^S dt &\xrightarrow{u.c.p.} \int_0^\cdot m_t^T b_t^S dt, \\ \frac{1}{\varepsilon^2} \int_0^\cdot (Z_t^\varepsilon)^T a_t^S Z_t^\varepsilon dt &\rightarrow_p \int_0^\cdot v_t \cdot a_t^S dt, \\ \frac{1}{\varepsilon} \int_0^\cdot (Z_t^\varepsilon)^T a_t^S dt &\rightarrow_p \int_0^\cdot m_t^T a_t^S dt. \end{aligned}$$

To see this, just put $r_t D(x) = x^T b_t^S$, $x^T a_t^S x$ and $x^T a_t^S$. □

Example 4.4.1. *Consider the discretization strategies based on two barriers $\underline{l}_t > 0$ and $\bar{l}_t > 0$ such that $G_t = [-\underline{l}_t, \bar{l}_t]$ and $\xi_t(\bar{l}_t) = \xi_t(-\underline{l}_t) = 0$. The corresponding (μ, ρ) is given by the stationary distribution and boundary measure of Brownian motion with rebirth (see [GK07]). We have*

$$m_t = \frac{1}{3}(l_t - \bar{l}_t), \quad v_t = \frac{1}{6}(\bar{l}_t^2 + \underline{l}_t^2 - \bar{l}_t \underline{l}_t),$$

recovering example in [Fuk11c].

Remark 4.4.1. *If we compare with the results in [Roo80] and [HM05], we observe that their results fall into above framework in the sense that, for equidistant discretization scheme, (μ, ρ) satisfies the LP constraint*

$$\int_{\mathbb{R}} \frac{1}{2} a f''(x) \mu(dx) + \int_{\mathbb{R}} (f(0) - f(x)) \rho(dx) = 0, \quad \forall f \in C_0^2(\mathbb{R}),$$

and $\rho(\mathbb{R}) = 1$. Intuitively, ρ is Gaussian distribution $\mathcal{N}(0, a)$ and μ is the stationary measure of a Brownian motion with rebirth at $t \in \mathbb{N}^*$. However, some technical issues for the tightness of occupation measure and characterization of (μ, ρ) need to be treated with care.

The quantities m and ν are the first and second moments of μ . By Ito formula, μ is the law of $\int_0^1 \sqrt{a} W_t dt = \frac{1}{2} \sqrt{a} (W_t^2 - 1)$, so m and ν are related to the third and fourth moments of a normal distribution, which is a known fact in the literature.

4.4.3 Impacts of small market frictions

As we have already mentioned, the lower bound (3.1.6) appears also in the study of impact of small market frictions in the framework of utility maximization, see [KMK15, KL13, GW15a, ST13, PST15, AMKS15, MMKS14, LMKW14]. In this section, we explain heuristically how to relate utility maximization under small market frictions to the problem of tracking. It should

be pointed out that we are just making connections between these two problems and rigorous proof of the equivalence is left for future research.

We follow the presentation in [KMK15] and consider the classical utility maximization problem

$$u(t, w_t) = \sup_{\varphi} \mathbb{E}[U(w_T^{t, w_t})],$$

with

$$w_s^{t, w_t} = w_t + \int_t^s \varphi_u dS_u,$$

where φ is the trading strategy. The market dynamics is an Itô semi-martingale

$$dS_t = b_t^S dt + \sqrt{a_t^S} dW_t.$$

In the frictionless market, we denote by φ_t^* the optimal strategy and by w_t^* the corresponding wealth process. As mentioned in [KMK15], the indirect marginal utility $u'(t, w_t^*)$ evaluated along the optimal wealth process is a martingale density, which we denote by Z_t :

$$Z_t = u'(t, w_t^*).$$

Note that S is a martingale under \mathbb{Q} with

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \frac{Z_T}{Z_0}.$$

One also defines the indirect risk tolerance process R_t by

$$R_t = -\frac{u'(t, w_t^*)}{u''(t, w_t^*)}.$$

Consider the exponential utility function as in [KL13], that is

$$U(x) = -e^{-px}, \quad p > 0.$$

Then we have

$$R_t = R = \frac{1}{p}.$$

In a market with proportional transaction costs, the portfolio dynamics is given by

$$w_s^{t, w_t, \varepsilon} = w_t^\varepsilon + \int_t^s \varphi_u^\varepsilon dS_u - \int_t^s \varepsilon h_u d\|\varphi^\varepsilon\|_u,$$

where h_t is a random weight process and φ_t^ε a process with finite variation. The control problem is then

$$u^\varepsilon(t, w_t) = \sup_{\varphi^\varepsilon} \mathbb{E}[U(w_T^{t, w_t, \varepsilon})].$$

When the cost ε is small, we can expect that φ_t^ε is close to φ_t^* and set

$$\Delta w_T^\varepsilon := w_T^{0, w_0, \varepsilon} - w_T^* = \int_0^T (\varphi_t^\varepsilon - \varphi_t^*) dS_t - \varepsilon \int_0^T h_t d\|\varphi^\varepsilon\|_t.$$

Then up to first order quantities, we have

$$\begin{aligned} u^\varepsilon - u &= \mathbb{E}[U(w_T^* + \Delta w_T^\varepsilon)] - \mathbb{E}[U(w_T^*)] \\ &\simeq \mathbb{E}[U'(w_T^*) \Delta w_T^\varepsilon + \frac{1}{2} U''(w_T^*) (\Delta w_T^\varepsilon)^2] \\ &= -u'(w_0) \mathbb{E}^\mathbb{Q}[\Delta w_T^\varepsilon + \frac{1}{2R} (\Delta w_T^\varepsilon)^2] \\ &\simeq -u'(w_0) \mathbb{E}^\mathbb{Q}[\varepsilon \int_0^T h_t d\|\varphi^\varepsilon\|_t + \frac{1}{2R} (\int_0^T (\varphi_t^\varepsilon - \varphi_t^*) dS_t)^2] \\ &= -u'(w_0) \mathbb{E}^\mathbb{Q}[\varepsilon \int_0^T h_t d\|\varphi^\varepsilon\|_t + \int_0^T \frac{a_t^S}{2R} (\varphi_t^\varepsilon - \varphi_t^*)^2 dt]. \end{aligned}$$

Similarly, in a market with fixed transaction costs εk_t , see [AMKS15], the portfolio dynamics is given by

$$w_s^{t, w_t, \varepsilon} = w_t^\varepsilon + \int_t^s \varphi_u^\varepsilon dS_u + \Psi_s - \varepsilon \sum_{t < \tau_j^\varepsilon \leq s} k_{\tau_j^\varepsilon} F(\xi_j^\varepsilon), \quad \varphi_t^\varepsilon = \sum_{0 < \tau_j^\varepsilon \leq t} \xi_j^\varepsilon,$$

and we have

$$u^\varepsilon - u \simeq -u'(w_0) \mathbb{E}^\mathbb{Q} \left[\varepsilon \sum_{0 < \tau_j^\varepsilon \leq T} k_{\tau_j^\varepsilon} F(\xi_j^\varepsilon) + \int_0^T \frac{a_t^S}{2R} (\varphi_t^\varepsilon - \varphi_t^*)^2 dt \right].$$

Finally, in a market with linear impact εl_t on price, see [RS10, MMKS14], the portfolio dynamics is given by

$$w_s^{t, w_t, \varepsilon} = w_t^\varepsilon + \int_t^s \varphi_u^\varepsilon dS_u - \varepsilon \int_t^s l_u (u_u^\varepsilon)^2 du, \quad \varphi_t^\varepsilon = \int_0^t u_t^\varepsilon dt,$$

and we have

$$u^\varepsilon - u \simeq -u'(w_0) \mathbb{E}^\mathbb{Q} \left[\varepsilon \int_0^T l_t (u_t^\varepsilon)^2 dt + \int_0^T \frac{a_t^S}{2R} (\varphi_t^\varepsilon - \varphi_t^*)^2 dt \right].$$

To sum up, utility maximization under small market frictions is heuristically equivalent to the tracking problem if the deviation penalty is set to be

$$r_t D(x) = \frac{a_t^S}{2R} x^2. \quad (4.4.2)$$

Defining the certainty equivalent wealth loss Δ^ε by

$$u^\varepsilon =: u(w_0 - \Delta^\varepsilon),$$

it follows that

$$\frac{1}{\varepsilon \beta \zeta_D} \Delta^\varepsilon \simeq \mathbb{E}^\mathbb{Q} \left[\int_0^T I_t dt \right], \quad (4.4.3)$$

see also [KMK15, Equation (3.4)] and [MMKS14, pp. 18].

Remark 4.4.2 (Higher dimension and general utility function). *For the case of higher dimension and general utility function, one should set*

$$r_t D(x) = \frac{1}{2R_t} \langle x, a_t^S x \rangle. \quad (4.4.4)$$

In other words, utility maximization under market frictions can be approximated at first order by the problem of tracking with quadratic deviation cost (4.4.4). Thus one can establish a connection between the tracking problem and the utility maximization problems in [AMKS15, MMKS14, GW15c, GW15b].

Remark 4.4.3 (General cost structures). *When there are multiple market frictions with comparable impacts, the choice of deviation penalty is the same as (4.4.4) and one only needs to adjust the cost structure. Our results apply directly in these cases, see [LMKW14, GW15b]. For example, in the case of trading with proportional cost and linear market impact, see [LMKW14], the local problem is the time-average control of Brownian motion with cost structure*

$$C_A(x, u) = rx^2 + lu^2 + h|u|.$$

Indeed, Equations (4.3)-(4.5) in [LMKW14] give rise to a verification theorem for the HJB equation of the time-average control problem of Brownian motion under this cost structure.

Remark 4.4.4 (Non-zero interest rate). *In the case of non-zero interest rate, the correspondence should be written as*

$$\frac{1}{\varepsilon^{\beta\zeta_D}}\Delta^\varepsilon \simeq \mathbb{E}^{\mathbb{Q}}\left[\int_0^T e^{-rt} I_t dt\right], \quad (4.4.5)$$

where $e^{-rt}S_t$ is a \mathbb{Q} -martingale and the tracking problem is defined by (4.4.4). For example, the right hand side of (4.4.5) is the probabilistic representation under Black-Scholes model of Equation (3.11) in [WW97].

Remark 4.4.5 (Optimal consumption over infinite horizon). *In [ST13, PST15], the authors consider the problem of optimal consumption over infinite horizon under small proportional costs. Their results can be related to the tracking problem in the same way, that is*

$$\frac{1}{\varepsilon^{\beta\zeta_D}}\Delta^\varepsilon \simeq \mathbb{E}^{\mathbb{Q}}\left[\int_0^\infty e^{-rt} I_t dt\right],$$

where $e^{-rt}S_t$ is a \mathbb{Q} -martingale and the tracking problem is defined by (4.4.4).

Remark 4.4.6 (Indifference pricing of option). *The formula (4.4.3) is used in [KMK13] for the indifference pricing of option under small proportional transaction costs. Let Δ^H and Δ^0 be the normalized CE for the problem of utility maximization with and without selling option payoff H . Let φ^* be the optimal strategy of the pure investment problem and φ^H the replicating strategy for H (we assume complete market), then the tracking targets for Δ^H and Δ^0 are $\varphi^* + \varphi^H$ and φ^* respectively. The short/long price p_\pm^ε of the option H is given by*

$$\pm p_\pm^\varepsilon \simeq \mathbb{E}^{\mathbb{Q}}[\pm H] + \varepsilon^{\beta\zeta_D}(\Delta^{\pm H} - \Delta^0).$$

The option spread due to transaction costs is given by

$$p_+^\varepsilon - p_-^\varepsilon \simeq \varepsilon^{\beta\zeta_D}(\Delta^H + \Delta^{-H} - 2\Delta^0),$$

which is always positive by convexity. If the pure investment strategy φ^* is negligible compared to φ^H , then we have

$$p_+^\varepsilon - p_-^\varepsilon \simeq 2\varepsilon^{\beta\zeta_D}\Delta^H,$$

see also [MT15, Section 6].

4.5 Proofs

4.5.1 Proof of Theorem 4.2.1

Recall that the dynamics of $\tilde{X}^{\varepsilon,t}$ is given by

$$\tilde{X}_s^{\varepsilon,t} = \tilde{X}_0^{\varepsilon,t} + \int_0^s \tilde{b}_\nu^{\varepsilon,t} d\nu + \int_0^s \sqrt{\tilde{a}_\nu^{\varepsilon,t}} d\tilde{W}_\nu^{\varepsilon,t} + \int_0^s \tilde{u}_\nu^{\varepsilon,t} d\nu + \sum_{0 < \tilde{\tau}_j^{\varepsilon,t} \leq s} \tilde{\xi}_j^\varepsilon,$$

with $(\tilde{W}_s^{\varepsilon,t})$ Brownian motion w.r.t. $\tilde{\mathcal{F}}_s^{\varepsilon,t} = \mathcal{F}_{t+\varepsilon\alpha\beta s}$,

$$\tilde{b}_s^{\varepsilon,t} = \varepsilon^{(\alpha-1)\beta} b_{t+\varepsilon\alpha\beta s}, \quad \tilde{a}_s^{\varepsilon,t} = a_{t+\varepsilon\alpha\beta s},$$

and

$$\tilde{u}_s^{\varepsilon,t} = \varepsilon^{(\alpha-1)\beta} u_{t+\varepsilon\alpha\beta s}^\varepsilon, \quad \tilde{\xi}_j^\varepsilon = \frac{1}{\varepsilon^\beta} \xi_j^\varepsilon, \quad \tilde{\tau}_j^{\varepsilon,t} = \frac{1}{\varepsilon\alpha\beta} (\tau_j^\varepsilon - t) \vee 0.$$

The corresponding local cost is

$$I_t^\varepsilon = \frac{1}{T^\varepsilon} \left(\int_0^{T^\varepsilon} (r_t D(\tilde{X}_s^{\varepsilon,t}) + l_t Q(\tilde{u}_s^{\varepsilon,t})) ds + \sum_{0 < \tilde{\tau}_j^{\varepsilon,t} \leq T^\varepsilon} (k_t F(\tilde{\xi}_j^\varepsilon) + h_t P(\tilde{\xi}_j^\varepsilon)) \right).$$

It is enough to study the term on the right hand side thanks to the following fact.

Lemma 4.5.1. *We have*

$$\lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon^{\zeta_D \beta}} J^\varepsilon = \lim_{\varepsilon \rightarrow 0} \int_0^T I_t^\varepsilon dt, \quad \text{a.e.,}$$

if the term on the right hand side exists.

Proof. see Section 3.5.1. □

Moreover, up to a localization procedure, we can assume that $u_t^g = \{(x, u_t(x)), x \in G_t\}$ is contained in a bounded ball of $\mathbb{R}_x^d \times \mathbb{R}_u^d$, and there exist positive constants δ and M such that $\delta < V(\xi_t(x)) - V(x)$, $x \in \partial G_t$, and $V(x) < M$, $x \in G_t$, for any (t, ω) . We assume also that Assumption 3.5.1 holds.

Define random measures $(\mu_t^\varepsilon, \rho_t^\varepsilon)$ by

$$\begin{aligned} \mu_t^\varepsilon &= \frac{1}{T^\varepsilon} \int_0^{T^\varepsilon} \delta_{\{(\tilde{X}_s^{\varepsilon,t}, \tilde{u}_s^{\varepsilon,t})\}} ds \in \mathcal{P}(\mathbb{R}_x^d \times \mathbb{R}_u^d), \\ \rho_t^\varepsilon &= \frac{1}{T^\varepsilon} \sum_{0 < \tilde{\tau}_j^{\varepsilon,t} \leq T^\varepsilon} \delta_{\{(\tilde{X}_{\tilde{\tau}_j^{\varepsilon,t}}^{\varepsilon,t}, \tilde{\xi}_j^\varepsilon)\}} \in \mathcal{M}(\mathbb{R}_x^d \times \mathbb{R}_\xi^d), \end{aligned}$$

and $c : \Omega \times [0, T] \times \mathcal{P}(\mathbb{R}_x^d \times \mathbb{R}_u^d) \times \mathcal{M}(\mathbb{R}_x^d \times \mathbb{R}_\xi^d) \rightarrow \mathbb{R}$ by

$$\begin{aligned} c_t(\omega, \mu, \rho) &:= \int_{\mathbb{R}_x^d \times \mathbb{R}_u^d} (r_t(\omega)D(x) + l_t(\omega)Q(u))\mu(dx \times du) \\ &\quad + \int_{\mathbb{R}_x^d \times \mathbb{R}_\xi^d} (k_t(\omega)F(\xi) + h_t(\omega)P(\xi))\rho(dx \times d\xi). \end{aligned}$$

Then $(\mu^\varepsilon, \rho^\varepsilon)$ is a sequence of stochastic processes with values in $\mathcal{P}(\mathbb{R}_x^d \times \mathbb{R}_u^d) \times \mathcal{M}(\mathbb{R}_x^d \times \mathbb{R}_\xi^d)$ and

$$\int_0^T I_t^\varepsilon dt = \int_0^T c_t(\mu_t^\varepsilon, \rho_t^\varepsilon) dt.$$

Let $(\pi_t, \nu_t) := (\pi^{(a_t, u_t, G_t, \xi_t)}, \nu^{(a_t, u_t, G_t, \xi_t)})$ be the process uniquely determined by (4.2.6), and put

$$\mu_t(dx \times du) = \pi_t(dx) \otimes \delta_{u(x)}(du), \quad \rho_t(dx \times d\xi) = \nu_t(dx) \otimes \delta_{\xi(x)}(d\xi).$$

Then we have to show

$$\int_0^T c_t(\mu_t^\varepsilon, \rho_t^\varepsilon) dt \rightarrow_p \int_0^T c_t(\mu_t, \rho_t) dt. \quad (4.5.1)$$

In view of Appendix 4.A, it suffices to prove the following lemma.

Lemma 4.5.2. 1. *For any $t \in [0, T)$, $(\mu_t^\varepsilon, \rho_t^\varepsilon)$ converges in probability to (μ_t, ρ_t) .*

2. *The sequence $\{(\mu_t^\varepsilon, \rho_t^\varepsilon)_{t \in [0, T]}, \varepsilon > 0\}$ of measure valued stochastic processes is weakly tight with respect to the cost functional c (see Definition 4.A.2).*

Remark 4.5.1. *Let $\mathbb{E} := \mathcal{P}(\mathbb{R}_x^d \times \mathbb{R}_u^d) \times \mathcal{M}(\mathbb{R}_x^d \times \mathbb{R}_\xi^d)$. The first property implies that $(\mu^\varepsilon, \rho^\varepsilon)$ converges weakly to (μ, ρ) (in fact in probability) in $\mathbb{M}_{\mathbb{E}}([0, T], \mathcal{B}([0, T]), dt)$, the space of \mathbb{E} -valued measurable functions equipped with the topology of convergence in dt -measure (see [Tsu07]).*

Proof of Lemma 4.5.2. We claim first that

$$\sup_{\varepsilon > 0} \mathbb{E} \left[\left(\frac{N_{T^\varepsilon}^\varepsilon}{T^\varepsilon} \right)^k \right] < \infty, \quad k = 1, 2, \quad (4.5.2)$$

Indeed, by Ito formula, we have

$$\begin{aligned}
\frac{N_{T^\varepsilon}^\varepsilon}{T^\varepsilon} &\leq \frac{1}{T^\varepsilon} \sum_{j=1}^{N_{T^\varepsilon}^\varepsilon} \delta^{-1} (V(\tilde{X}_{\tilde{\tau}_j^\varepsilon-}^{\varepsilon,t}) - V(\tilde{X}_{\tilde{\tau}_j^\varepsilon}^{\varepsilon,t})) \\
&= \delta^{-1} \frac{1}{T^\varepsilon} \left[-V(\tilde{X}_{T^\varepsilon}^{\varepsilon,t}) + V(\tilde{X}_{0+}^{\varepsilon,t}) + \int_0^{T^\varepsilon} \nabla V(\tilde{X}_s^{\varepsilon,t})^T \sqrt{\tilde{a}_s^{\varepsilon,t}} d\tilde{W}_s^{\varepsilon,t} \right. \\
&\quad \left. + \int_0^{T^\varepsilon} \sum_i \tilde{u}_{i,s}^{\varepsilon,t} \partial_i V(\tilde{X}_s^{\varepsilon,t}) ds + \int_0^{T^\varepsilon} \frac{1}{2} \sum_{ij} \tilde{a}_{ij,s}^{\varepsilon,t} \partial_{ij}^2 V(\tilde{X}_s^{\varepsilon,t}) ds \right] \\
&\leq \frac{M}{\delta} \left(\text{const.} + \int_0^{T^\varepsilon} \nabla V(\tilde{X}_s^{\varepsilon,t})^T \sqrt{\tilde{a}_s^{\varepsilon,t}} d\tilde{W}_s^{\varepsilon,t} \right)
\end{aligned}$$

the last term being obviously square-integrable after the localization procedure. Now we are ready to prove the two claims.

Convergence in probability. By (4.5.2) and localization, for any $t \in [0, T]$ fixed, the family $\{(\mu_t^\varepsilon, \rho_t^\varepsilon), \varepsilon > 0\}$ is tight. Let \mathbb{Q}_t be any stable limit of $(\mu_t^\varepsilon, \rho_t^\varepsilon)$. Since

$$c_t(\mu_t^\varepsilon, \rho_t^\varepsilon) \leq M \left(1 + \frac{N_{T^\varepsilon}^\varepsilon}{T^\varepsilon} \right),$$

we have $\sup_\varepsilon \mathbb{E}[c_t(\mu_t^\varepsilon, \rho_t^\varepsilon)] < \infty$ in view of (4.5.2) with $k = 1$. Then by Lemma 3.5.3, we have \mathbb{Q}_t^ω -a.e.,

$$(\mu, \rho) \in S(a_t(\omega)),$$

where we recall that $S(a)$ is defined by

$$\begin{aligned}
S(a) &= \left\{ (\mu, \bar{\rho}) \in \mathcal{P}(\mathbb{R}_x^d \times \mathbb{R}_u^d) \times \overline{\mathcal{M}(\mathbb{R}_x^d \times \mathbb{R}_\xi^d \setminus \{0_\xi\})} \right\}, \\
\bar{\rho} &= \rho + \theta_{\bar{\rho}} \delta_\infty \text{ with } \rho \in \mathcal{M}(\mathbb{R}_x^d \times \mathbb{R}_\xi^d \setminus \{0_\xi\}), \\
&\int_{\mathbb{R}_x^d \times \mathbb{R}_u^d} A^a f(x, u) \mu(dx, du) + \int_{\mathbb{R}_x^d \times \mathbb{R}_\xi^d \setminus \{0_\xi\}} Bf(x, \xi) \rho(dx, d\xi) = 0, \forall f \in C_0^2(\mathbb{R}_x^d).
\end{aligned}$$

On the other hand, let F_μ^ω and F_ρ^ω be given by

$$F_\mu^\omega = \{(x, u_t^\omega(x)) | x \in G_t^\omega\}, \quad F_\rho^\omega = \{(x, \xi_t^\omega(x)) | x \in \partial G_t^\omega\}.$$

By the continuity of (G_t, ξ_t) , \mathbb{Q}_t^ω -a.e., (μ, ρ) is supported on F_μ^ω and F_ρ^ω respectively. By the separability (4.2.6) of (G_t, ξ_t) , such couple of (μ, ρ) is unique, so we have

$$\mathbb{Q}_t = \mathbb{P}(d\omega) \otimes \delta_{\{(\mu^{(a_t, u_t, G_t, \xi_t)}(\omega), \rho^{(a_t, u_t, G_t, \xi_t)}(\omega))\}}.$$

which is the unique possible limit point and we deduce that convergence in probability holds.

Weak tightness of $(\mu^\varepsilon, \rho^\varepsilon)$ with respect to c . In view of (4.5.2) with $k = 2$, the application $(t, \omega) \mapsto c_t(\omega, \mu_t^\varepsilon(\omega), \rho_t^\varepsilon(\omega))$ is uniformly square-integrable w.r.t. $\mathbb{P} \otimes dt$, hence uniformly integrable. By Lemma 2 of [CK86] and the remark after it, $(\mu^\varepsilon, \rho^\varepsilon)$ is weakly tight w.r.t. c . \square

4.5.2 Proof of Theorem 4.3.1

The proof is exactly the same as Theorem 4.2.1 with (4.5.2) replaced by

$$\sup_{\varepsilon > 0} \mathbb{E} \left[\left(\frac{\tilde{\mathcal{Q}}_{T^\varepsilon}^\varepsilon}{T^\varepsilon} \right)^k \right] < \infty, \quad k = 1, 2,$$

which can be obtained by applying Itô formula to $V(\tilde{X}_{T^\varepsilon}^\varepsilon)$:

$$\begin{aligned} \frac{\tilde{\varphi}_{T^\varepsilon}^\varepsilon}{T^\varepsilon} &\leq \frac{1}{T^\varepsilon} \int_0^{T^\varepsilon} \delta^{-1} [-\tilde{\gamma}_t^\varepsilon \cdot \nabla V(\tilde{X}_s^{\varepsilon,t})] d\tilde{\varphi}_t^\varepsilon \\ &= \delta^{-1} \frac{1}{T^\varepsilon} \left[-V(\tilde{X}_{T^\varepsilon}^{\varepsilon,t}) + V(\tilde{X}_{0+}^{\varepsilon,t}) + \int_0^{T^\varepsilon} \nabla V(\tilde{X}_s^{\varepsilon,t}) \cdot \sqrt{\tilde{a}_s^{\varepsilon,t}} d\tilde{W}_s^{\varepsilon,t} \right. \\ &\quad \left. + \int_0^{T^\varepsilon} \sum_i \tilde{u}_{i,s}^{\varepsilon,t} \partial_i V(\tilde{X}_s^{\varepsilon,t}) ds + \int_0^{T^\varepsilon} \frac{1}{2} \sum_{ij} \tilde{a}_{ij,s}^{\varepsilon,t} \partial_{ij}^2 V(\tilde{X}_s^{\varepsilon,t}) ds \right] \\ &\leq \frac{M}{\delta} \left(\text{const.} + \int_0^{T^\varepsilon} \nabla V(\tilde{X}_s^{\varepsilon,t})^T \sqrt{\tilde{a}_s^{\varepsilon,t}} d\tilde{W}_s^{\varepsilon,t} \right). \end{aligned}$$

4.5.3 Proof of Theorem 4.3.4

The rescaled process $(\tilde{X}_s^{\varepsilon,t})$ is given by

$$d\tilde{X}_s^{\varepsilon,t} = \tilde{b}_s^{\varepsilon,t} ds + \sqrt{\tilde{a}_s^{\varepsilon,t}} d\tilde{W}_s^{\varepsilon,t} + \tilde{u}_s^{\varepsilon,t} ds,$$

with

$$\tilde{b}_s^{\varepsilon,t} = \varepsilon_{t+\varepsilon^{2\beta s}}^\beta, \quad \tilde{a}_s^{\varepsilon,t} = a_{t+\varepsilon^{2\beta s}}, \quad \tilde{u}_s^{\varepsilon,t} = \varepsilon^{(2-1)\beta} u_{t+\varepsilon^{2\beta s}}^\varepsilon.$$

The empirical occupation measure μ_t^ε is defined by

$$\mu_t^\varepsilon = \frac{1}{T^\varepsilon} \int_0^{T^\varepsilon} \delta_{\{(\tilde{X}_s^{\varepsilon,t}, \tilde{u}_s^{\varepsilon,t})\}} ds.$$

The proof is slightly different since X^ε is not constrained inside a uniformly bounded domain G_t as before.

Up to a localization procedure, we can assume that there exist $\theta, \Theta > 0$ such that

$$(A^{at} + u_t(x) \cdot \nabla)V(x) \leq \theta - 2\Theta V(x), \quad x \in \mathbb{R}_x^d.$$

We follow [ABG11, Lemma 2.5.5] and obtain that

$$\mathbb{E} \left[V(\tilde{X}_s^{\varepsilon,t}) \right] \leq \frac{\theta}{2\Theta} + V(0), \quad \forall t \in [0, T), s \in (0, T^\varepsilon]. \quad (4.5.3)$$

Let π_t^ε be the marginal of μ_t^ε on \mathbb{R}_x^d . Since $rD + lQ \circ u \leq V$ near infinity, the empirical costs $\{c_t(\mu_t^\varepsilon)\}$ are bounded by $\int (b_0 + b_1 V(x)) d\pi_t^\varepsilon$. By (4.5.3), we have

$$\sup_\varepsilon \mathbb{E}[c_t(\mu_t^\varepsilon)] \leq \sup_\varepsilon \mathbb{E} \left[\int (b_0 + b_1 V(x)) d\pi_t^\varepsilon \right] = b_0 + b_1 \frac{1}{T^\varepsilon} \int_0^{T^\varepsilon} \mathbb{E}[V(\tilde{X}_s^{\varepsilon,t})] ds < \infty. \quad (4.5.4)$$

Convergence in probability. Since u_t is admissible, u_t sends compact sets into compact sets, the tightness of μ_t^ε follows directly from the tightness of π_t^ε . By (4.5.4) we can apply Lemma 3.5.3 and the convergence in probability follows from the separability of π w.r.t. A_u^a .

Weakly tightness. For any $t \in [0, T)$ and $\varepsilon > 0$, we have that (4.5.4) holds for a sufficient large constant. Since $c_t(\mu_t^\varepsilon) \rightarrow_p c_t(\mu_t)$, we have by Fubini theorem and dominated convergence theorem

$$\limsup_{\varepsilon \rightarrow 0} \int c_t(\mu_t^\varepsilon) dt \otimes d\mathbb{P} = \int c_t(\mu_t) dt \otimes d\mathbb{P} < \infty.$$

By [CK86, Lemma 3], we obtain the weakly tightness.

Appendix 4.A Convergence of integral functionals

In this section, we provide a direct generalization of the result in [CK86], which allows us to pass from the convergence of local systems (Lemma 4.5.2) to the convergence of cost integrals (4.5.1).

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, (T, \mathcal{B}, μ) a σ -finite measure space and S a Polish space with Borel σ -field $\mathcal{B}_S = \mathcal{B}(S)$, $C(S)(C_b(S))$ the space of continuous (bounded continuous) real-valued functions on S , $\mathcal{L}_1(\mu) := \mathcal{L}_1(T, \mathcal{B}, \mu)$ the space of integrable real-valued functions with seminorm $\|x\|_1 := \int_T |x_t| \mu(dt)$ and $L_1(\mu) := L_1(T, \mathcal{B}, \mu)$ the corresponding Banach space.

Now let $c : \Omega \times T \times S \rightarrow \mathbb{R}$ be a $\mathcal{F} \otimes \mathcal{B} \otimes \mathcal{B}_S$ -measurable function with $c_t(\omega, \cdot) \in C(S)$ for all $t \in T$. Let $\{X^n, n \in \mathbb{N}\}$ be a sequence of $\mathcal{F} \otimes \mathcal{B}/\mathcal{B}_S$ -measurable functions $X^n : \Omega \times T \rightarrow S$ with

$$\int_T |c_t(\omega, X_t^n(\omega))| \mu(dt) < +\infty, \quad \omega \in \Omega, n \in \mathbb{N}.$$

Then the random integral $I(X^n, c)$ is defined by

$$I(X^n, c) := \int_T c_t(X_t^n) \mu(dt).$$

Note that the map $(\omega, t) \mapsto X_t^{n,c}(\omega) := c_t(\omega, X_t^n(\omega))$ is $\mathcal{F} \otimes \mathcal{B}$ -measurable. For all $\omega \in \Omega$, $\hat{X}^{n,c} := X^{n,c}(\omega)$ is an element of $L_1(\mu)$ and $\omega \mapsto \hat{X}^{n,c}(\omega)$ is $\mathcal{F}/\mathcal{B}(L_1(\mu))$ -measurable (see [CK86]).

Definition 4.A.1. *The processes X^n converges to X^0 in probability in finite dimension if there is $T_0 \in \mathcal{B}$ with $\mu(T \setminus T_0) = 0$ such that for any $t_1, \dots, t_k \in T_0$, $(X_{t_1}^n, \dots, X_{t_k}^n)$ converges to $(X_{t_1}^0, \dots, X_{t_k}^0)$ in probability.*

Definition 4.A.2. *A sequence $\{X^n, n \in \mathbb{N}\}$ of measurable processes is called weakly tight with respect to c if for each $\delta > 0$, there is $K \subset L_1(\mu)$ weakly compact, that is compact in the $\sigma(L_1(\mu), L_\infty(\mu))$ -topology, such that*

$$\inf_n \mathbb{P}[\hat{X}^{n,c} \in K] > 1 - \delta.$$

In particular, the family X^n is weakly tight with respect to c if one of the following condition holds:

1. The family of r.v. $\{(\omega, t) \mapsto c_t(\omega, X_t^n(\omega)), n \in \mathbb{N}^*\}$ is $\mathbb{P} \otimes \mu$ -uniformly integrable, see [CK86, Lemma 2].
2. X^n converges weakly to X^0 in finite dimension and

$$\limsup_n \int c_t(X_t^n) dt \otimes d\mathbb{P} \leq \int c_t(X_t^0) dt \otimes d\mathbb{P} < \infty.$$

See [CK86, Lemma 3].

Now we can state our result concerning the convergence in probability of the random variables $I(X^n, c)$.

Theorem 4.A.1. *Let $\{X^n, n \in \mathbb{N}\}$ be a sequence of stochastic process. If X^n converges to X^0 in probability in finite dimension and if $\{X^n, n \in \mathbb{N}\}$ is weakly tight w.r.t. c , then $I(X^n, c)$ converges to $I(X^0, c)$ in probability.*

Proof. After Lemma 3.D.1, it suffices to show that $I(X^n, c)$ converges stably to $I(X^0, c)$. Let Y be any bounded r.v. and f bounded continuous and Lipschitz, we will show that

$$\mathbb{E}[Y f(I(X^n, c))] \rightarrow \mathbb{E}[Y f(I(X^0, c))], \quad n \rightarrow \infty. \quad (4.A.1)$$

i.) Let $h \in \mathcal{L}_1^+(\mu)$. Define

$$c_t^h(\omega, x) := \max\{-h_t, \min\{h_t, c_t(\omega, x)\}\},$$

and

$$I(X^n, c^h) := \int_T c_t^h(X_t^n) \mu(dt), \quad n \in \mathbb{N}.$$

We show that $I(X^n, c^h)$ converges stably to $I(X^0, c^h)$, i.e.

$$\mathbb{E}[Y f(I(X^n, c^h))] \rightarrow \mathbb{E}[Y f(I(X^0, c^h))], \quad n \rightarrow \infty. \quad (4.A.2)$$

Since $|I(X^n, c^h)| \leq \|h\|_1$ and polynomials are dense in $C([- \|h\|_1, \|h\|_1])$, we only need to consider $f(x) = x^l$ for some $l \in \mathbb{N}$. By Fubini's theorem, we obtain

$$\begin{aligned} \mathbb{E}[Y f(I(X^n, c^h))] &= \int_{\Omega} Y(\omega) \left(\int_T c_{t_1}^h(\omega, X_{t_1}^n) \mu(dt_1) \cdots \int_T Y(\omega) c_{t_l}^h(\omega, X_{t_l}^n) \mu(dt_l) \right) \mathbb{P}(d\omega) \\ &= \int \cdots \int F_n(t_1, \dots, t_l) \mu(dt_1) \cdots \mu(dt_l) \end{aligned}$$

where

$$F_n(t_1, \dots, t_l) := \int_{\Omega} Y(\omega) c_{t_1}^h(\omega, X_{t_1}^n) \cdots c_{t_l}^h(\omega, X_{t_l}^n) \mathbb{P}(d\omega).$$

By weak convergence and the finite dimensional convergence in probability of X^n , we have $F_n(t_1, \dots, t_l) \rightarrow F_0(t_1, \dots, t_l)$ for any $t_1, \dots, t_l \in T_0$. Since

$$F_n(t_1, \dots, t_l) \leq (\sup_{\omega} |Y|) h(t_1) \cdots h(t_l),$$

we have

$$\int \cdots \int F_n(t_1, \dots, t_l) \mu(dt_1) \cdots \mu(dt_l) \rightarrow \int \cdots \int F_0(t_1, \dots, t_l) \mu(dt_1) \cdots \mu(dt_l),$$

by dominated convergence, whence (4.A.1).

ii.) By [CK86, Remark 1] and the weakly tightness of X^n w.r.t. c , for all $N \in \mathbb{N}^*$ there is a weakly compact K_N of $L_1(\mu)$ and $h_N \in \mathcal{L}_1^+(\mu)$ such that

$$\inf_n \mathbb{P}[\hat{X}^{n,c} \in K_N] \geq 1 - \frac{1}{N}, \quad (4.A.3)$$

and

$$\sup_{x \in K_N} \int_T (|x| - h_N)^+ d\mu \leq \frac{1}{N}. \quad (4.A.4)$$

We can assume w.l.o.g that $h_N \uparrow \infty$ for $N \uparrow \infty$. We have

$$\begin{aligned} |\mathbb{E}[Y f(I(X^n, c))] - \mathbb{E}[Y f(I(X^0, c))]| &\leq |\mathbb{E}[Y f(I(X^n, c))] - \mathbb{E}[Y f(I(X^n, c^{h_N}))]| \\ &\quad + |\mathbb{E}[Y f(I(X^n, c^{h_N}))] - \mathbb{E}[Y f(I(X^0, c^{h_N}))]| \\ &\quad + |\mathbb{E}[Y f(I(X^0, c^{h_N}))] - \mathbb{E}[Y f(I(X^0, c))]| \\ &=: e_1 + e_2 + e_3. \end{aligned}$$

Since Y is bounded and f is bounded Lipschitz, we have

$$\begin{aligned} e_1 &\leq (\sup_{\omega} |Y|) (2(\sup_x |f|) \mathbb{P}[|I(X^n, c) - I(X^n, c^{h_N})| \geq \frac{1}{N}] + \text{Lip}(f) \frac{1}{N}) \\ &\leq \text{const.} \frac{1}{N}, \end{aligned}$$

where the constant is independent of n . Here we use (4.A.3), (4.A.4) and

$$\begin{aligned} |I(X^n, c^{h_N}) - I(X^n, c)| &\leq \int_T |c_t^{h_N}(X_t^n) - c_t(X_t^n)| \mu(dt) \\ &= \int_T (|c_t(X_t^n)| - h_N(t))^+ \mu(dt). \end{aligned}$$

Hence $e_1 < \varepsilon$ for N large enough. By dominated convergence, $I(X^0, c^{h_N}) \rightarrow I(X^0, c)$ pointwise, hence $e_3 < \varepsilon$ for N large enough. Now fix N , by (4.A.2), we have $e_2 < \varepsilon$ for n large enough. In sum, we have $|\mathbb{E}[Yf(I(X^n, c))] - \mathbb{E}[Yf(I(X^0, c))]| \leq 3\varepsilon$ for n large enough. Since ε is arbitrary, (4.A.1) follows and we can conclude. \square

Appendix 4.B Separability of (A, B)

In this section, we describe a sufficient condition for the separability of (A, B) with respect to a given domain G and its associated jump rule ξ or reflection Γ .

Let $\pi \in \mathcal{P}(\bar{G})$ and $\nu \in \mathcal{M}(\partial G)$ be the marginal distribution of μ and ρ on \mathbb{R}^d_x , then the LP constraint (4.2.6) and (4.3.1) can be rewritten as

$$\int_{\bar{G}} A_u^a f(x) \pi(dx) + \int_{\partial G} Bf(x) \nu(dx) = 0, \quad \forall f \in C_0^2(\mathbb{R}^d_x),$$

with

$$A_u^a f(x) = \frac{1}{2} \sum_{ij} a_{ij} \partial_{ij}^2 f(x) + \langle u, \nabla f \rangle,$$

and

1. In the impulse case, $Bf(x) = f(x + \xi(x)) - f(x)$
2. In the singular case, $Bf(x) = \gamma(x) \cdot \nabla f(x)$ if we consider the case where $\Gamma(x)$ is one dimensional, i.e. $\Gamma(x) \cap \Delta = \{\gamma(x)\}$.

The basic idea is to reduce the uniqueness of (μ, ρ) to the ergodicity of the corresponding constrained martingale problem developed by [Kur90, KS01, CK15].

Let G be a bounded domain of \mathbb{R}^d with piecewise- C^2 boundary and let

$$\partial G = \cup_{i=1}^m \partial_i G,$$

be a partition into C^2 segments of the boundary ∂G . Let $A := A_u^a$ and $B_i, i = 1, \dots, m$ be the corresponding operator of diffusion and boundary jumps/reflections and \mathcal{D} the common domain of the operators.

Definition 4.B.1 (Constrained Martingale Problem). *A stochastic process X with sample path in $D_{\bar{G}}[0, \infty)$ is a solution of the constrained martingale problem for $(A, G; B_1, \partial_1 G; \dots; B_m, \partial_m G)$ if there exist a filtration (\mathcal{F}_t) and random measures Λ_i on $[0, \infty) \times \bar{\partial}_i G, i = 1, \dots, m$, such that for any $f \in \mathcal{D}$,*

$$f(X_t) - \int_0^t Af(X_s) ds - \sum_{i=1}^m \int_{[0,t] \times \bar{\partial}_i G} B_i f(x) \Lambda_i(ds \times dx)$$

is a \mathcal{F}_t -martingale.

We will assume that

Assumption 4.B.1. *The constrained martingale problem for $(A, G; B_1, \partial_1 G; \dots; B_m, \partial_m G)$ is well-posed, i.e. for any given initial distribution ν , there exists a unique distribution \mathbb{P}_ν on $D_{\bar{G}}[0, \infty)$ which solves the martingale problem.*

Note that here the uniqueness is for the distribution of the process X only.

Remark 4.B.1. *While a solution of the constrained martingale problem with given initial condition is easy to construct, uniqueness is not trivial.*

In the case of diffusion with reflection, there exists criteria (see [CK15, Theorem 5.11]) based on the comparison principle for the viscosity solution of the resolvent equation

$$\lambda u - Au = h,$$

with suitable first order boundary condition (see [CK15, Definition 2.20]).

They also suggest that the approach might be generalized to diffusion with rebirth [CK15, Remark 3.2 and 5.3]. Nevertheless, if B_ξ is given by a jump rule ξ on ∂G , one might directly show that the solution of the corresponding constrained martingale problem is indeed a diffusion with rebirth defined in the next section, following the idea in [Kur11].

Theorem 4.B.1. *Under suitable conditions (see Condition 1.2 and equation (1.16) in [KS01]), there exists a stationary solution (X, Λ_i) of the constrained martingale problem $(A, G, B_i, \partial_i G)$ such that X_t has distribution $\pi \in \mathcal{P}(\bar{G})$ and $\mathbb{E}[\Lambda_i([0, t] \times \cdot)] = t\nu_i(\cdot)$ for any $t \geq 0$ if and only if*

$$\int_{\bar{G}} A(x)\pi(dx) + \sum_{i=1}^m \int_{\partial_i G} B_i f(x)\nu_i(dx) = 0, \quad \forall f \in C_0^2(\bar{G}).$$

Proof. The \Rightarrow direction is evident. The \Leftarrow direction follows from [KS01, Theorem 1.7]. \square

4.B.1 Diffusion with Rebirth

Let $x + \xi(x) \in C^0(\partial G, G)$ be a jump rule from the boundary ∂G to the domain G . Then the boundary operator is given by

$$B_\xi f(x) = f(x + \xi(x)) - f(x).$$

The solution $(X_t^{(G, \xi)})$ of the corresponding constrained martingale problem with initial distribution π can be constructed as in [GK12, GK07, BAP09, GK02]. There is also a natural associated Markov chain $(Y_n^{(G, \xi)})$ which is the hitting position of $X_t^{(G, \xi)}$ on the boundary. The following construction is taken from [BAP09].

Let $Z^{\pi, 0}$ be a diffusion killed at the boundary with generator A and initial distribution π . Let $\{Z^{y+\xi(y), n}, y \in \partial G, n \in \mathbb{N}\}$ be a family of independent diffusion killed at ∂G with generator A and initial distribution $\delta_{y+\xi(y)}$. Denote τ_G^Z the hitting time of a diffusion Z on the boundary of G . Put $\tau_0 = 0$ and $Y_0 = Z_{\tau_0}^{\mu, 0}$. Let

$$\tau_1 = \tau_G^{Z^{\pi, 0}}, \quad Y_1 = Z_{\tau_1}^{\pi, 0},$$

and

$$\tau_{n+1} = \tau_n + \tau_G^{Z^{Y_n+\xi(Y_n), n}}, \quad Y_{n+1} = Z_{\tau_{n+1}}^{Y_n+\xi(Y_n), n}.$$

Under suitable conditions, we have $\tau_n \rightarrow +\infty$ almost surely, so

$$X_t^{G, \xi} := \sum_{n=0}^{\infty} \mathbb{1}_{\{[\tau_n, \tau_{n+1})\}}(t) Z_{t-\tau_n}^{Y_n+\xi(Y_n), n}$$

is well defined. We can verify that $X^{G, \xi}$ constructed in this way is a solution of the constrained martingale problem with initial distribution π . We call $(Y_n^{(G, \xi)})$ the boundary chain.

Proposition 4.B.1 (Separability). *Under Assumption 4.B.1, if the diffusion with rebirth $(X_t^{(G,\xi)})$ is ergodic, then there exists a unique couple $(\pi, \nu) \in \mathcal{P}(\bar{G}) \times \mathcal{M}(\partial G)$ verifying*

$$\int_{\bar{G}} A(x)\pi(dx) + \int_{\partial G} B_\xi f(x)\nu(dx) = 0, \quad \forall f \in C_0^2(\bar{G}).$$

Proof. Since $\{B_\xi f, f \in C_0^2\}$ is a separating class of $\mathcal{M}(\partial G)$, the uniqueness of ν follows from the uniqueness of π . The latter is by definition the ergodicity of $(X_t^{(G,\xi)})$. \square

Remark 4.B.2. *For the ergodicity of diffusion with rebirth, see [GK12], [BAP09] or [AKK14].*

4.B.2 Diffusion with Reflection

Let G be of C^2 boundary. Let $\gamma : \partial G \rightarrow \Delta$ be a Lipschitz continuous vector field on the boundary. Let $\Gamma(x)$ be the closed convex cone in \mathbb{R}^d with vertex at the origin spanned by $\gamma(x)$. Let $n(x)$ be the interior normal vector to the domain G at $x \in \partial G$ and assume that

$$\forall x \in \partial G, \langle n(x), \gamma(x) \rangle > 0. \quad (4.B.1)$$

The boundary operator B_γ is given by

$$B_\gamma f(x) = \langle \gamma(x), \nabla f(x) \rangle.$$

The solution of the corresponding constrained martingale problem is denoted by $X^{G,\Gamma}$, see for example [LS84].

Remark 4.B.3. *Under the condition (4.B.1), the constrained martingale problem is equivalent to the submartingale problem in [SV71], see Theorem 6.3.1 of [Kur90]. In [KR14], the authors provide a characterization of the marginal π using the submartingale formulation.*

Proposition 4.B.2 (Separability). *Under Assumption 4.B.1, if the diffusion with reflection $(X_t^{(G,\Gamma)})$ is ergodic, then there exists a unique couple $(\pi, \nu) \in \mathcal{P}(\bar{G}) \times \mathcal{M}(\partial G)$ verifying*

$$\int_{\bar{G}} A(x)\pi(dx) + \int_{\partial G} B_\gamma f(x)\nu(dx) = 0, \quad \forall f \in C_0^2(\bar{G}).$$

Proof. The uniqueness of π is equivalent to the ergodicity of $(X_t^{(G,\Gamma)})$. The uniqueness of ν is less obvious since we do not know whether $\{B_\gamma f, f \in C_0^2\}$ is a separating class. However, by [KS01, Lemma 2.4], the random measure Λ can be written as

$$\Lambda(\cdot \times [0, t]) = \int_0^t \mathbb{1}_{\{\cdot\}}(X_s^{(G,\Gamma)}) d\varphi_s,$$

for some continuous non-decreasing process φ . By Remark 4.B.3 and [SV71, Theorem 2.4], φ is the unique boundary process. Hence ν is uniquely determined as marginal of Λ . \square

Remark 4.B.4. *For the case where G is a bounded domain with piecewise smooth boundary, we can follow the arguments in [RW88, Theorem 1] to show that φ_i does not charge the non-smooth part of $\text{int}(\partial_i G)$, where the reflecting direction is not unique. It follows, see for example [TW93, Lemma 2.1] and [DW96, Lemma 4.6], that $\varphi_i, i = 1, \dots, m$ can be written as functional of $X^{(G,\Gamma)}$, hence are uniquely determined by the latter.*

Remark 4.B.5. *For the ergodicity of reflected Brownian motion in a bounded planar region, see [HLS85, Theorem 2.8, 3.8]. Close-formed expression for π can be found in [HLS85, Theorem 2.18, 3.13]. A more general characterization of the stationary distribution π in terms of adjoint PDE is given in [KR14, Theorem 3].*

	Stochastic + singular control	Stochastic + impulse control
Local Problem	$\overline{\lim}_{T \rightarrow \infty} \frac{1}{T} \int_0^T (rX_t^2 + lu_t^2) dt + \frac{1}{T} h\varphi_T$	$\overline{\lim}_{T \rightarrow \infty} \frac{1}{T} \int_0^T (rX_t^2 + lu_t^2) dt + \frac{1}{T} \sum_{0 < \tau_j \leq T} (k + h \xi_j)$
Optimal Cost	$\iota(a^2rl)^{1/2}, \iota(a, r, l, h) \in (0, 1)$	$\iota(a^2rl)^{1/2}, \iota(a, r, l, k, h) \in (0, 1)$
u^*	$u^*(x) = -\frac{1}{2}l^{-1}w'(x; a, r, l, h)$, see (3.7.12)	$u^*(x) = -\frac{1}{2}l^{-1}w'(x; a, r, l, k, h)$, see (3.7.12)
$G^* = [-x^*, x^*]$	$x^* = x^*(a, r, l, h)$	$x^* = x^*(a, r, l, k, h)$
γ^*	$\gamma^*(\pm x^*) = \mp 1$	*
ξ^*	*	$\pm x^* + \xi^*(\pm x^*) = \pm \tilde{x}^*(a, r, l, k, h)$
Global Problem	$\int_0^T (r_t X_t^2 + \varepsilon^{\beta_Q} l_t u_t^2) dt + \varepsilon^{\beta_P} \int_0^T h_t d\varphi_t$	$\int_0^T (r_t X_t^2 + \varepsilon^{\beta_Q} l_t u_t^2) dt + \sum_{0 < \tau_j \leq T} (\varepsilon^{\beta_P} k_{\tau_j} + \varepsilon^{\beta_P} h_{\tau_j} \xi_j)$
(ζ_D, ζ_Q)	(2, 2)	(2, 2)
$(\beta_Q, \beta_F, \beta_P, \beta)$	$(4, *, 3, 1)\beta$	$(4, 4, 3, 1)\beta$
Asymp. Opt.	$\int_0^T \iota(a_t, r_t, l_t, h_t)(a_t^2 r_t l_t)^{1/2} dt$	$\int_0^T \iota(a_t, r_t, l_t, k_t, h_t)(a_t^2 r_t l_t)^{1/2} dt$
u_t^*	$u_t^*(x) = -\frac{1}{2}l_t^{-1}w'(x; a_t, r_t, l_t, h_t)$	$u_t^*(x) = -\frac{1}{2}l_t^{-1}w'(x; a_t, r_t, l_t, k_t, h_t)$
$G_t^* = [-x_t^*, x_t^*]$	$x_t^* = x^*(a_t, r_t, l_t, h_t)$	$x_t^* = x^*(a_t, r_t, l_t, k_t, h_t)$
γ_t^*	$\gamma_t^*(\pm x_t^*) = \mp 1$	*
ξ_t^*	*	$\pm x_t^* + \xi_t^*(\pm x_t^*) = \pm \tilde{x}^*(a_t, r_t, l_t, k_t, h_t)$

Figure 4.1 – Explicit formula for quadratic deviation cost: combined control

	Stochastic	Singular	Impulse ($h = 0$ for left column)
Local Problem	$\overline{\lim}_{T \rightarrow \infty} \frac{1}{T} \int_0^T (rX_t^2 + lu_t^2) dt$	$\overline{\lim}_{T \rightarrow \infty} \frac{1}{T} \int_0^T rX_t^2 + h\varphi_T$	$\overline{\lim}_{T \rightarrow \infty} \frac{1}{T} \left(\int_0^T rX_t^2 + \sum_{0 < \tau_j \leq T} (k+h \xi_j) \right)$
Optimal Cost	$(a^2 r t)^{1/2}$	$\left(\left(\frac{3}{4} \right)^2 a^2 r h^2 \right)^{1/3}$	$\left(\frac{2}{3} a r k \right)^{1/2}$
u^*	$w^*(x) = -(r l^{-1})^{1/2} x$	*	*
$G^* = [-x^*, x^*]$	*	$x^* = \left(\frac{3}{4} a r^{-1} h \right)^{1/3}$	$x^*(a, h, k)$
ξ^* or γ^*	*	$\gamma^*(\pm x^*) = \mp 1$	$\pm x^* + \xi^*(\pm x^*) = 0$ $\pm x^* + \xi^*(\pm x^*) = \pm \tilde{x}^*(a, h, k)$
Global Problem	$\int_0^T (r_t(X_t)^2 dt + \varepsilon^{\beta_Q} l_t (u_t)^2) dt$	$\int_0^T r_t(X_t)^2 dt + \varepsilon^{\beta_P} \int_0^T h_t d\varphi_t$	$\int_0^T r_t(X_t)^2 dt + \sum_{0 < \tau_j \leq T} (\varepsilon^{\beta_F} k_{\tau_j} + \varepsilon^{\beta_P} h_{\tau_j} \xi_j)$
(ζ_D, ζ_Q)	$(2, 2)$	$(2, *)$	$(2, *)$
$(\beta_Q, \beta_F, \beta_P, \beta)$	$(4, *, *, 1)\beta$	$(*, *, 3, 1)\beta$	$(*, 4, *, 1)\beta$
Asymp. Opt.	$\int_0^T (a_t^2 r_t l_t)^{1/2} dt$	$\int_0^T \left(\left(\frac{3}{4} \right)^2 a_t^2 r_t h_t^2 \right)^{1/3} dt$	$\int_0^T I(a_t, r_t, k_t) dt$
u^*	$u_t^*(x) = -(r_t l_t^{-1})^{1/2} x$	*	*
$G_t^* = [-x_t^*, x_t^*]$	*	$x_t^* = \left(\frac{3}{4} a_t r_t^{-1} h_t \right)^{1/3}$	$x_t^*(a_t, r_t, k_t)$
ξ_t^* or γ_t^*	*	$\gamma_t^*(\pm x_t^*) = \mp 1$	$\pm x_t^* + \xi_t^*(\pm x_t^*) = 0$ $\pm x_t^* + \xi_t^*(\pm x_t^*) = \pm \tilde{x}_t^*(a_t, r_t, k_t)$

Figure 4.2 – Explicit formula for quadratic deviation cost: single control

Part III

**PORTFOLIO SELECTION WITH
CAPTIAL GAINS TAXES**

Chapter 5

Capital gains taxes with recursive utility and regime switching¹

Contents

5.1	Introduction	165
5.2	Model formulation	167
5.3	Preliminary analysis	169
5.3.1	First properties of the value function	169
5.3.2	Certainty equivalent wealth loss (CEWL)	171
5.3.3	The value of deferring capital gains realization	171
5.3.4	HJB equation	171
5.3.5	Trading and no-trading regions	173
5.4	Asymptotic analysis	174
5.4.1	Small parameter	174
5.4.2	Asymptotic expansion	175
5.4.3	Economic analysis: single-regime case	176
5.4.4	Economic analysis: two-regime case	177
5.5	Numerical results	178
5.5.1	Single-regime case	178
5.5.2	Two-regime Case	182
5.6	Why is the order of deferral value $O(\varepsilon^{8/3})$?	185
5.6.1	Expansion around the tax-deflated model	185
5.6.2	Strategies based on barriers and reduced model	187
5.6.3	Solution of the optimal control problem (5.6.17)-(5.6.18)	189
Appendix 5.A	Formal derivation of Asymptotic Expansion 5.4.1	191
5.A.1	Expansion at solution of tax-deflated model	191
5.A.2	An approximation to the optimal consumption strategy	193

5.1 Introduction

In this part, we turn our attention to another important topic in portfolio selection: the impact of capital gains taxes in the implementation of consumption-investment strategy. Capital gain taxes differ from transaction costs in the following aspects: 1) investors pay taxes for capital gains but receive tax rebates for capital losses and 2) the amount of capital gains or losses taxed depends on the purchase price of stock holdings, known as the tax basis, which incurs strong

1. This is a joint work with Prof. Xinfu Chen and Prof. Min Dai based on the preprint [CD13].

path-dependency.

As a consequence, much of the existing literature on capital gain taxes has been restricted to discrete-time models with small number of time steps (see [Con83, Con84, DK96, DU05, GH06]). Using the average tax basis as an approximation, [DSZ01, DSZ03] develop a binomial tree model that is able to effectively work with multi-step investment and consumption decisions. [GKT06] further extend to the multiple stocks case. It is worth pointing out that the average tax basis is indeed used in Canada. In [BST10], the authors formulate a continuous-time version of the model introduced by [DSZ01], extending the classical Merton optimal consumption-investment problem.

In this chapter, we consider the optimal consumption investment problem with capital gain taxes in a regime switching market for an agent with recursive utility of Epstein-Zin type. First, it is well-known that the implication of the expected utility is not compatible with empirical findings, causing various asset pricing puzzles like the equity puzzle, the excess volatility puzzle and the credit spread puzzle and etc. The discounted expected utility in Merton's model restricts the relationship between the agent's risk aversion and elasticity of intertemporal substitution (EIS): the EIS is equal to the reciprocal of risk aversion. By separating these two parameters, recursive utility of Epstein-Zin type provides a framework to address aforementioned asset pricing anomalies (see [DE92, BY04, BCDG11, Bha09]). And it would be interesting to understand the impact of capital gains taxes on agents of different time-preference. Second, [Ham89] shows that regime-switching model represents the stock returns better than the model with deterministic coefficients. By introducing a two-state Markov chain, one can take into account of stochastic investment opportunities and distinguish the impact of capital gains taxes on the investment strategy under different economic conditions.

Since closed-form solutions are generally unavailable for the continuous-time model, we aim to utilize asymptotic analysis to characterize the optimal strategy. However, even for the single risky asset case, the model with capital gain taxes has to involve the tax basis as an additional state variable, which makes rigorous expansions intractable. Therefore, *we only provide formal expansions that will be justified by numerical results.*

To conduct asymptotic expansions with capital gain taxes, it is crucial to find appropriate perturbation parameters because unlike the transaction cost rate, the tax rate is relatively large. [BST10] provide an interesting insight: the Merton problem with capital gain taxes reduces to a frictionless Merton problem with tax-deflated drift and volatility coefficients as the interest rate vanishes.² Since the interest rate is typically small, it can serve as one of the perturbation parameters. In light of theoretical analysis, we find that asymptotic expansions can be sought in terms of not only small interest rate or tax rate but also other parameters, including large stock volatility and risk aversion index of the investor. We will introduce a unified perturbation parameter integrating these factors (see (5.4.1)).

Because no capital gains tax is paid until capital gains are realized, investors tend to defer realization of capital gains so as to save interest. This deferral option is constrained by suboptimal risk exposure. As such, investors need to achieve a trade-off between the benefit of tax deferral and the cost of suboptimal risk exposure. This leads to a no-trading region enclosed by two trading boundaries: the optimal buy boundary and the optimal sell boundary, which determine the optimal investment strategy. When the investment opportunity set is constant, we obtain explicit approximations of the optimal trading boundaries and the value of deferral. The approximations offer qualitative and quantitative insights about the effects of various factors on

2. The insight was initially documented by Domar and Musgrave (1944).

the optimal strategy and the deferral option. We deduce that capital gains tax may have a first order effect on investor utility. On the other hand, we discover several delicate properties of the optimal trading boundaries using numerical method. For example, we find that the initial tax-adjusted optimal fraction of wealth in the risky asset is slightly higher than the Merton line, and the buy boundary typically deviates more from the Merton line than the sell boundary does.

When the investment opportunity set is not constant, we find that investors smooth the value of deferral across regimes and that in contrast to the myopic investment strategy in the absence of capital gains tax, the optimal investment strategy is affected by the investment opportunity in the other regime. While such smoothing effect is similar to that observed in the presence of transaction costs, the situation in our case is more involved. We find that the degree of the smoothing effect varies in function of the level of capital gains. In the region of low capital gains, the optimal strategy is highly sensitive to market regime and the investor follows closely to the myopic strategy in the absence of capital gains tax. In contrast, if the level of capital gains is high, then the costs of reacting to market change become significant and the investor is reluctant to rebalance the portfolio.

Finally, we find that the elasticity of intertemporal substitution (EIS) may dramatically alter the value of deferral but has little impact on the optimal investment strategy.

The remainder of the paper is organized as follows. In Section 5.2, we formulate the problem as a stochastic control problem whose value function satisfies a Hamilton-Jacobi-Bellman (HJB) equation. Some properties of the value function are given in Section 5.3. In Section 5.4, we present the main results and discuss their financial implications. Numerical analysis and justifications are given in Section 5.5. We provide a probabilistic model in Section 5.6 to provide more insight into our asymptotic analysis. All technical derivations and proofs are contained in the Appendix.

5.2 Model formulation

In this section, we present a mathematical formulation for the consumption investment problem with the average tax basis, extending the model established by [BST10] to a regime switching market and for an agent with recursive utility of Epstein-Zin type. We work on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{P})$. Here $\{\mathcal{F}_t\}_{t \geq 0}$ is the filtration generated by a standard Wiener process $\{W_t, t \geq 0\}$ with $W_0 = 0$ and a Markov process $\{i_t, t \geq 0\}$ with values in $\mathcal{I} = \{1, 2, \dots\}$ and generator matrix

$$Q = (q_{ij})_{i,j \in \mathcal{I}},$$

where $\sum_{j \in \mathcal{I}} q_{ij} = 0, \forall i \in \mathcal{I}$. Moreover, W_t and i_t are independent of each other.

Assume that there are two assets that the investor can trade without any transaction costs. The first asset (“the bond”) is a money market account growing at a continuously compounded, before-tax interest rate r_t . The second asset (“the stock”) is a risky investment and its price S_t follows

$$dS_t = \mu_t S_t dt + \sigma_t S_t dW_t.$$

The market parameters r_t, μ_t and σ_t are regime dependent: let the Markov process $\{i_t, t \geq 0\}$ represent the market regime, then r_t, μ_t and σ_t depend on i_t through

$$r_t = r_{i_t}, \quad \mu_t = \mu_{i_t}, \quad \sigma_t = \sigma_{i_t},$$

where r_i, μ_i and σ_i are market parameters under regime $i \in \mathcal{I}$.

The investor is subject to capital gain taxes. We assume that (i) the capital gain can be realized immediately after sale, (ii) there is no wash sale restriction, (iii) short selling is prohibited, and (iv) the tax basis used to evaluate capital gains is defined as the weighted average of past purchase prices.

Let X_t , Y_t , and K_t be the amount invested in the bond, the current dollar value of, and the purchase price of stock holdings, respectively. We introduce two càdlàg, nonnegative, and non-decreasing $\{\mathcal{F}_t\}_{t \geq 0}$ -adapted processes L_t and M_t with $L_0 = M_0 = 0$, where dL_t represents the *dollar amount* transferred from the bank to the stock account at time t (corresponding to a purchase of stock), while $dM_t \leq 1$ represents the *proportion of shares* transferred from the stock account to the bank at time t (corresponding to a sale of stock). Bear in mind that the average tax basis is used to evaluate capital gains. Hence, when one sells stock at time t , the purchase price K_t declines by the same proportion dM_t as the dollar value of stock holdings does, and the realized capital gain is $(Y_t - K_t)dM_t$. Then, the evolution processes of X_t, Y_t , and K_t are³

$$\begin{aligned} dX_t &= [(1 - \alpha)r_t X_t - C_t] dt - dL_t + [Y_{t-} - \alpha(Y_{t-} - K_{t-})] dM_t, \\ dY_t &= \mu_t Y_{t-} dt + \sigma_t Y_{t-} dW_t + dL_t - Y_{t-} dM_t, \\ dK_t &= dL_t - K_{t-} dM_t, \end{aligned}$$

where α and $\{C_t, t \geq t\}$ are the tax rate and the consumption streams, respectively.

We consider an investor who's preference over the random consumption streams $\{C_t, t \geq 0\}$ is described by a continuous time stochastic differential utility of Epstein-Zin type. More precisely, let $\beta \in (0, \infty)$ be the rate of time preference, $\gamma \in (0, 1) \cup (1, \infty)$ be the relative risk aversion and $\frac{1}{\kappa}$ be the EIS with $\kappa \in (0, 1) \cup (1, \infty)$, then the Epstein-Zin aggregator f is given by

$$f(C, V) = \frac{\beta}{1 - \kappa} C^{1-\kappa} ((1 - \gamma)V)^\theta - p\beta V, \quad (5.2.1)$$

where

$$p = \frac{1 - \gamma}{1 - \kappa}, \quad \theta = 1 - \frac{1}{p}. \quad (5.2.2)$$

The Epstein-Zin utility $\{V_t^C, t \geq 0\}$ over the consumption streams $\{C_t, t \geq 0\}$ is implicitly determined by⁴

$$V_t^C = \mathbb{E} \left[\int_t^\infty f(C_s, V_s^C) ds \middle| \mathcal{F}_t \right], \quad \forall t \geq 0. \quad (5.2.3)$$

When $\gamma = \kappa$, the aggregator reduces to

$$f(C, V) = \beta U(C, \gamma) - \beta V,$$

where $U(\cdot, \gamma)$ is the CRRA utility function with parameter γ :

$$U(C, \gamma) = \frac{C^{1-\gamma}}{1 - \gamma}. \quad (5.2.4)$$

Then (5.2.3) can be written as

$$V_t^C = \beta \mathbb{E} \left[\int_t^\infty e^{-\beta(s-t)} U(C_s, \gamma) ds \middle| \mathcal{F}_t \right], \quad (5.2.5)$$

which is the standard additive utility up to a multiplication of β .

3. For notational reason, the bond is assumed to be also subject to capital gain taxes. The resulting model is equivalent to that in [BST10] with $(1 - \alpha)r$ replaced by r .

4. For conditions on C_t guaranteeing the existence of V_t^C , see e.g. Duffie and Lions (1994), Hao (2015).

Denote by

$$Z_t = X_t + Y_t - \alpha(Y_t - K_t)$$

the liquidation value of the portfolio, and define the solvency region as

$$\Omega := \{(X_t, Y_t, K_t) \in \mathbb{R}^3 : Z_t \geq 0, Y_t \geq 0, K_t \geq 0\}.$$

Then a consumption investment strategy $\{(C_t, L_t, M_t), t \geq 0\}$ is said to be admissible if (i) $(X_t, Y_t, K_t) \in \Omega$ and (ii) the Epstein-Zin utility V_t^C is uniquely determined by (5.2.3). Since the utility process is subject to the trading constraint (i), we denote V_t^C by $V_t^{(C,L,M)}$. Let $\mathcal{A}_i(x, y, k)$ be the set of all admissible strategies under initial condition $X_0 = x, Y_0 = y, K_0 = k$ and $i_0 = i$, then the agent aims to find

$$\varphi_i(x, y, k) := \sup_{(C,L,M) \in \mathcal{A}_i(x,y,k)} V_0^{(C,L,M)}. \quad (5.2.6)$$

and the corresponding optimal strategy $\{(C_t^*, L_t^*, M_t^*), t \geq 0\}$. Here a consumption investment strategy $\{(C_t^*, L_t^*, M_t^*), t \geq 0\}$ is said to be an optimal strategy if

$$\varphi_i(x, y, k) = V_0^{(C^*, L^*, M^*)}.$$

5.3 Preliminary analysis

We first present some properties of the value function using similar arguments from [BST10].

5.3.1 First properties of the value function

If the tax rate $\alpha = 0$, our model reduces to the classical problem of recursively utility maximisation with regime switching. We have (see [Xia12]),

Proposition 5.3.1 (Tax-free model). *Assume that $|\mathcal{I}| = 1$ or 2, and*

$$\bar{K}_i := \frac{\beta}{\kappa} - \frac{1-\kappa}{\kappa} \left(r_i + \frac{(\mu_i - r_i)^2}{2\gamma\sigma_i^2} \right) > 0, \quad \forall i \in \mathcal{I}. \quad (5.3.1)$$

Then the following system

$$\beta^{\frac{1}{\kappa}} \bar{G}_i^{1-\frac{1}{\kappa}} - \bar{K}_i + \frac{1-\kappa}{\kappa} \frac{1}{1-\gamma} \sum_{j \in \mathcal{I}} q_{ij} \frac{\bar{G}_j^{1-\gamma}}{\bar{G}_i^{1-\gamma}} = 0, \quad i \in \mathcal{I}, \quad (5.3.2)$$

admits a unique solution $\{\bar{G}_i, i \in \mathcal{S}\}$. Denote by $\bar{\varphi}_i$ the value function (5.2.6) without capital gains taxes (i.e. $\alpha = 0$), then $\bar{\varphi}_i$ depends only on the initial wealth $z = x + y$ such that

$$\bar{\varphi}_i(z) = \frac{\bar{G}_i^{1-\gamma}}{1-\gamma} z^{1-\gamma}.$$

Moreover, the optimal consumption rate and risk asset proportion under regime $i \in \mathcal{I}$ is given by

$$\bar{c}_i = \beta^{\frac{1}{\kappa}} \bar{G}_i^{1-\frac{1}{\kappa}}, \quad \bar{\xi}_i = \frac{\mu_i - r_i}{\gamma\sigma_i^2}. \quad (5.3.3)$$

We remark that for an agent with utility of Epstein-Zin type, the optimal trading strategy $\bar{\xi}_i$ doesn't depend on κ or transition intensities q_{ij} . In contrast, the optimal consumption rate \bar{c}_i is closely related to the EIS and transition intensities. When necessary, we use superscript \bar{c}_i^q to

indicate the dependence of consumption rate in the transition densities

As in [BST10], we introduce the tax-deflated model in tax-free market for which the market parameters $r_{\alpha,i}$, $\mu_{\alpha,i}$ and $\sigma_{\alpha,i}$ are given by

$$r_{\alpha,i} = (1 - \alpha)r_i, \quad \mu_{\alpha,i} = (1 - \alpha)\mu_i, \quad \sigma_{\alpha,i} = (1 - \alpha)\sigma_i, \quad \forall i \in \mathcal{I}. \quad (5.3.4)$$

We define similarly $\bar{G}_{\alpha,i}$ and $\bar{K}_{\alpha,i}$ by

$$\beta^{\frac{1}{\kappa}} \bar{G}_{\alpha,i}^{1-\frac{1}{\kappa}} - \bar{K}_{\alpha,i} + \frac{1-\kappa}{\kappa} \frac{1}{1-\gamma} \sum_{j \in \mathcal{I}} q_{ij} \frac{\bar{G}_{\alpha,j}^{1-\gamma}}{\bar{G}_{\alpha,i}^{1-\gamma}} = 0, \quad i \in \mathcal{I}, \quad (5.3.5)$$

with

$$\bar{K}_{\alpha,i} := \frac{\beta}{\kappa} - \frac{1-\kappa}{\kappa} \left(r_{\alpha,i} + \frac{(\mu_{\alpha,i} - r_{\alpha,i})^2}{2\gamma\sigma_{\alpha,i}^2} \right). \quad (5.3.6)$$

Let $\bar{\varphi}_{\alpha,i}$ be the value function (5.2.6) under the tax-deflated model (5.3.4) without capital gains taxes, then we have

$$\bar{\varphi}_{\alpha,i}(z) = \frac{\bar{G}_{\alpha,i}^{1-\gamma}}{1-\gamma} z^{1-\gamma},$$

with

$$\bar{c}_{\alpha,i} = \beta^{\frac{1}{\kappa}} \bar{G}_{\alpha,i}^{1-\frac{1}{\kappa}}, \quad \bar{\xi}_{\alpha,i} = \frac{\mu_{\alpha,i} - r_{\alpha,i}}{\gamma\sigma_{\alpha,i}^2}. \quad (5.3.7)$$

As before, we use superscript $\bar{c}_{\alpha,i}^q$ to indicate the dependence of consumption rate in the transition densities. Note also that

$$\bar{\xi}_{\alpha,i} = \frac{1}{1-\alpha} \frac{\mu_i - r_i}{\gamma\sigma_i^2} = \frac{1}{1-\alpha} \bar{\xi}_i.$$

Proposition 5.3.2. *Denote the wealth after immediate liquidation by*

$$z = x + (1 - \alpha)y + \alpha k.$$

then for any $k \geq y$ (i.e. under capital loss),

$$\varphi_i(x, y, k) = \varphi_i(z, 0, 0) = \varphi_i((1 - \theta)z, \theta z, \theta z), \quad (5.3.8)$$

where θ is any positive constant.

Proof. This is a consequence of the monotonicity in consumption of stochastic differential utility, see [DE92, Section 5.3] and [BST10, Proposition 3.5]. \square

Proposition 5.3.2 implies that it is optimal to realize capital losses whenever they occur. Hence, when $K_t > Y_t$, the wash sale

$$(X_t, Y_t, K_t) \longrightarrow (Z_t, 0, 0) \longrightarrow ((1 - \xi_i^*)Z_t, \xi_i^*Z_t, \xi_i^*Z_t) \quad (5.3.9)$$

is an optimal strategy, where ξ_i^* is the optimal fraction of wealth in the risky asset after realizing capital losses under market regime $i \in \mathcal{I}$.

5.3.2 Certainty equivalent wealth loss (CEWL)

In [BST10], it is shown that the investor cannot take advantage of tax rebates to do better than in a tax-free market. Therefore we can define the *certainty equivalent wealth loss* (CEWL) Δ_i incurred by capital gain taxes in terms of the following equation:

$$\varphi_i(x, y, k) = \frac{\bar{G}_i^{1-\gamma}}{1-\gamma} [(1 - \Delta_i)z]^{1-\gamma}, \quad (5.3.10)$$

that is,

$$\begin{aligned} \Delta_i &\approx -\log(1 - \Delta_i) \\ &= -\frac{1}{1-\gamma} \log [(1 - \gamma)z^{-(1-\gamma)}\varphi_i] + \log \bar{G}_i. \end{aligned}$$

5.3.3 The value of deferring capital gains realization

On the other hand, consider the value function associated with the sub-optimal strategy described below: for any $t > 0$ and total wealth Z_t , one keeps liquidating the portfolio and maintains⁵

$$\frac{Y_t}{Z_t} = \bar{\xi}_{\alpha,i}, \quad \frac{C_t}{Z_t} = \bar{c}_{\alpha,i}. \quad (5.3.11)$$

The strategy can be regarded as a modified Merton strategy in the tax market. The modified Merton strategy is optimal in a restricted set of admissible strategies that the investor cannot defer realizing any capital gains. Because tax is paid only when capital gains are realized, the investor has not only an incentive but also an option to defer capital gains tax payment. We then define the value of this deferral option, denoted by $\Delta_{i,0}$, as follows:

$$\varphi_i(x, y, k) \equiv \frac{\bar{G}_{\alpha,i}^{1-\gamma}}{1-\gamma} [(1 + \Delta_{i,0})z]^{1-\gamma},$$

that is,

$$\begin{aligned} \Delta_{i,0} &\approx \log(1 + \Delta_{i,0}) \\ &= \frac{1}{1-\gamma} \log [(1 - \gamma)z^{-(1-\gamma)}\varphi_i] - \log \bar{G}_{\alpha,i} \end{aligned} \quad (5.3.12)$$

Later we will see that (5.3.12) implies a transformation which plays a critical role in our asymptotic analysis.

5.3.4 HJB equation

The value function φ_i turns out to satisfy the following HJB equation:⁶

$$\max \left\{ \mathcal{A}_i \varphi_i + \sum_{j \in \mathcal{I}} q_{ij} \varphi_j, \mathcal{T}_i \varphi_i, \mathcal{P}_i \varphi_i \right\} = 0 \quad \text{in } \Omega, \quad \forall i \in \mathcal{I}, \quad (5.3.13)$$

where

$$\begin{aligned} \mathcal{A}_i \varphi_i &:= \frac{1}{2} \sigma_i^2 y^2 \varphi_{i,yy} + \mu_i y \varphi_{i,y} + (1 - \alpha) r_i x \varphi_{i,x} + \bar{U}(\varphi_{i,x}, \varphi_i), \\ \mathcal{T}_i \varphi_i &:= [(1 - \alpha)y + \alpha k] \varphi_{i,x} - y \varphi_{i,y} - k \varphi_{i,k}, \\ \mathcal{P}_i \varphi_i &:= -\varphi_{i,x} + \varphi_{i,y} + \varphi_{i,k}, \end{aligned}$$

5. Strictly speaking, the strategy is not in the set of admissible strategies. However, we can always use a sequence of admissible strategies to approach the strategy.

6. In [BST10], the authors show rigorously that the value function is a viscosity solution of the HJB equation in a suitable sense. The well-posedness is given in [BST07]. In this paper, we don't attempt to prove such results but concentrate on the financial implications.

and

$$\begin{aligned}\bar{U}(q, \varphi) &:= \sup_{C>0} (f(C, \varphi) - Cq) \\ &= \beta \left(\frac{1}{1-\kappa} - 1 \right) \left(\frac{q}{\beta} \right)^{1-\frac{1}{\kappa}} ((1-\gamma)\varphi)^{\frac{\theta}{\kappa}} - p\beta\varphi.\end{aligned}$$

for each $q > 0$. The optimal consumption proves to be

$$C_i^* = \left(\frac{1}{\beta} \frac{\varphi_{i,x}}{((1-\gamma)\varphi_i)^\theta} \right)^{-\frac{1}{\kappa}}. \quad (5.3.14)$$

Since the Epstein-Zin utility (5.2.3) is homothetic, i.e. $V_t^{\lambda C} = \lambda^{1-\gamma} V_t^C$, we have

$$\varphi_i(\lambda x, \lambda y, \lambda k) = \lambda^{1-\gamma} \varphi_i(x, y, k), \quad \forall \lambda > 0. \quad (5.3.15)$$

Hence, there is a dimension reduction:⁷

$$\xi = \frac{(1-\alpha)y}{x + (1-\alpha)y + \alpha k}, \quad b = \frac{k}{y}, \quad (5.3.16)$$

such that

$$\varphi_i(x, y, k) = (x + (1-\alpha)y + \alpha k)^{1-\gamma} \phi_i(b, \xi), \quad (5.3.17)$$

where b is the scaled average tax basis. The combination of (5.3.12) and (5.3.17) motivates us to make the following transformation:

$$w_i = \frac{1}{1-\gamma} \log[(1-\gamma)\phi_i] - \log \bar{G}_{\alpha,i}. \quad (5.3.18)$$

It is not hard to verify that w_i satisfies

$$\begin{aligned}\max \left\{ \frac{\sigma_i^2}{2} (\mathcal{L}_i w_i + f_i) + \frac{1}{1-\gamma} \sum_{j \in \mathcal{I}} q_{ij} \frac{\bar{G}_{\alpha,j}^{1-\gamma}}{\bar{G}_{\alpha,i}^{1-\gamma}} (e^{(1-\gamma)(w_j - w_i)} - 1), \right. \\ \left. - w_{i,\xi}, \quad \xi w_{i,\xi} + (1-b)w_{i,b} \right\} = 0, \quad (5.3.19)\end{aligned}$$

in Q , where

$$\begin{aligned}\mathcal{L}_i w_i &= b^2 w_{i,bb} - 2b\xi(1-\xi)w_{i,b\xi} + \xi^2(1-\xi)^2 w_{i,\xi\xi} \\ &\quad + \left[2\left(1 - \frac{\mu_i}{\sigma_i^2}\right) + (1-\gamma)(-2\xi - \xi(1-\xi)w_{i,\xi} + bw_{i,b}) \right] bw_{i,b} \\ &\quad + \left[(-2\gamma(\xi - \bar{\xi}_i) + (1-\gamma)(\xi(1-\xi)w_{i,\xi} - bw_{i,b})) (1-\xi) + \frac{2r_i\alpha[1 + (b-1)\xi]}{\sigma_i^2} \right] \xi w_{i,\xi} \\ &\quad + \frac{2}{\sigma_i^2} \bar{c}_{\alpha,i} (U^*(e^{w_i}(1-\xi w_{i,\xi}), \kappa) - U^*(1, \kappa)), \\ f_i(b, \xi) &= -\gamma(\xi - \bar{\xi}_i)^2 + \frac{2}{\sigma_i^2} r_i \alpha \xi (1-b).\end{aligned} \quad (5.3.20)$$

and

$$U^*(q, \kappa) = \frac{\kappa}{1-\kappa} q^{1-\frac{1}{\kappa}}.$$

7. Following (5.3.15), one should use $\hat{\xi} = y/(x + (1-\alpha)y + \alpha k)$ and $\hat{b} = k/(x + (1-\alpha)y + \alpha k)$ as state variables. To derive (5.3.17), we can further make a transformation: $\xi = (1-\alpha)\hat{\xi}$ and $b = \hat{b}/\hat{\xi}$.

It is worth pointing out that the reduced equation (5.3.19) remains valid for $\gamma = 1$, which corresponds to the logarithm utility case. As such, the subsequent analysis works for all $\gamma > 0$.⁸ In what follows we identify w_i with the value of deferral $\Delta_{i,0}$. It is easy to see that we can rewrite the CEWL as

$$-\log(1 - \Delta_i) = \log \frac{\bar{G}_i}{\bar{G}_{\alpha,i}} - w_i.$$

5.3.5 Trading and no-trading regions

From now on, we assume $\min_{i \in \mathcal{I}} r_i \alpha > 0$. We define the wash-sale region \mathbf{WS}_i , the sell region \mathbf{SR}_i , the buy region \mathbf{BR}_i , and the non-trading region \mathbf{NT}_i by

$$\begin{aligned} \mathbf{WS}_i &:= \{(b, \xi) \mid b \geq 1, \xi \geq 0\}, \\ \mathbf{SR}_i &:= \{(b, \xi) \mid 0 \leq b < 1, w_{i,\xi} = 0\}, \\ \mathbf{BR}_i &:= \{(b, \xi) \mid 0 \leq b < 1, \xi w_{i,\xi} + (1 - b)w_{i,b} = 0\}, \\ \mathbf{NT}_i &:= \{(b, \xi) \mid 0 \leq b < 1, \xi w_{i,\xi} + (1 - b)w_{i,b} < 0 < w_{i,\xi}\}. \end{aligned}$$

The definition of the \mathbf{WS}_i is based on part (i) of Proposition 5.3.2. As a consequence, we are able to restrict attention to the domain $0 \leq b \leq 1$.

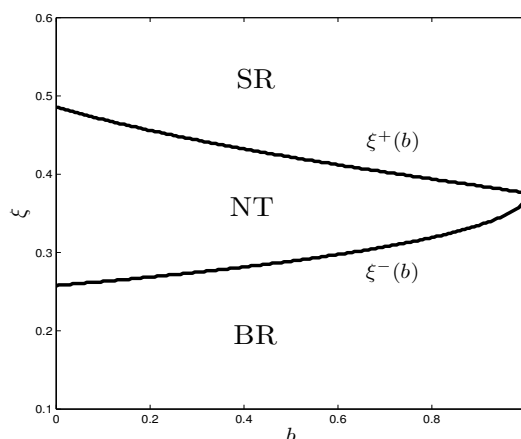


Figure 5.1 – Trading and No-trading Regions

We assume that $\phi_i \in C^1([0, \infty) \times [0, \infty))$ and there exist functions $\xi_i^\pm(\cdot) \in C^1[0, 1] \cap C[0, 1]$ such that

$$\begin{aligned} \mathbf{SR}_i &= \{(b, \xi) \mid 0 \leq b < 1, \xi \geq \xi_i^+(b)\}, \\ \mathbf{BR}_i &= \{(b, \xi) \mid 0 \leq b < 1, \xi \leq \xi_i^-(b)\}, \\ \mathbf{NT}_i &= \{(b, \xi) \mid 0 \leq b < 1, \xi_i^-(b) < \xi < \xi_i^+(b)\}. \end{aligned}$$

The assumption relies on the financial intuition that an investor whose preference is of Epstein-Zin type should sell (buy) the risky asset when the fraction of wealth in the risky asset is high

8. In fact, transformation (5.3.18) is initially inspired by an attempt to derive a unified reduced equation for all CRRA investors. A similar transformation is used in [DY09] for the Merton problem with transaction costs, where the transformation does not involve the constant $-\log \bar{G}_{\alpha,i}$.

(low) enough.⁹ These regions are depicted in Figure 1. The existence of the \mathbf{NT}_i in $b < 1$ implies that the investor can defer paying capital gain taxes.

We call the curve $\{(b, \xi_i^\pm(b)), b \in [0, 1]\}$, $i \in \mathcal{I}$, the *sell boundary/buy boundary* (the free boundary in mathematics). By analyzing the trading strategy, we shall assume that there hold the admissibility conditions

$$\xi_i^-(b) < \xi_i^+(b), \quad \forall b \in [0, 1]$$

and

$$\xi_i^+(1) = \xi_i^-(1). \quad (5.3.21)$$

Recall that there is a factor $(1 - \alpha)$ in the transform (5.3.16), Assumption (5.3.21) implies that after wash sale, the fraction in the risky asset should be re-balanced to

$$\xi_i^* = \frac{\xi_i^\pm(1)}{1 - \alpha}.$$

Coinciding with (5.3.9), ξ_i^* is the optimal fraction of wealth in the risky asset after wash sale. It can also be regarded as the optimal fraction of *initial* wealth in the risky asset.¹⁰ When we implement the above strategy, the trading boundaries in terms of risky asset proportion should be

$$\frac{\xi_i^-(b_t)}{1 - \alpha} \leq \frac{y_t}{z_t} \leq \frac{\xi_i^+(b_t)}{1 - \alpha}. \quad (5.3.22)$$

We emphasize that all of the above assumptions are supported by numerical results.

5.4 Asymptotic analysis

In this section, we present the main theoretical results derived from asymptotic analysis. All technical derivations and proofs are deferred to the Appendix. We shall perform the asymptotic analysis using the HJB equation (5.3.19).

5.4.1 Small parameter

The expression of $f_i(b, \xi)$ in (5.3.20) inspires us to introduce a small parameter

$$\varepsilon_i := \sqrt{\frac{2r_i \alpha \bar{\xi}_i}{\gamma \sigma_i^2}}, \quad (5.4.1)$$

to be used for asymptotic expansions. Then we can rewrite $f_i(b, \xi)$ as

$$f_i(b, \xi) = \gamma[\varepsilon_i^2(1 - b) - (\xi - \bar{\xi}_i)^2] + \varepsilon_i^2(1 - b)\left(1 - \frac{\xi}{\bar{\xi}_i}\right).$$

Due to the quadratic term $(\xi - \bar{\xi}_i)^2$, we expect the non-transaction zone be a narrow band containing the Merton line $\xi \equiv \bar{\xi}_i$. We will seek asymptotic expansion of the solution to (5.3.19) as $\varepsilon_i \searrow 0$ such that the last term on the right hand side becomes negligible and

$$f_i(b, \xi) \approx \gamma[\varepsilon_i^2(1 - b) - (\xi - \bar{\xi}_i)^2]. \quad (5.4.2)$$

9. However, it is difficult to prove the assumption because the value function is no longer concave [see [BST10]] for the additive utility case.

10. Apparently $\xi_i^\pm(1)$ (or ξ_i^*) plays a key role in application. Nevertheless, a rigorous mathematical proof for (5.3.21) is unavailable.

5.4.2 Asymptotic expansion

The differential operator \mathcal{L}_i in (5.3.19) is degenerate at $\xi = 1$, which results in different expansions for $\xi_i \neq 1$ and $\xi_i = 1$. Let us first investigate in this section the following case that is our primary interest.

Asymptotic Expansion 5.4.1.¹¹ Assume that $\bar{\xi}_i \neq 1, \forall i \in \mathcal{I}$. Let ε_i be as given in (5.4.1), and let $\xi_i^+(\cdot)$ and $\xi_i^-(\cdot)$ be the optimal sell and buy boundaries, respectively. Denote

$$A_i := 4/[3\bar{\xi}_i(\bar{\xi}_i - 1)^2], \quad \delta_i = (A_i\varepsilon_i)^{2/3}, \quad \theta_i = \frac{\sigma_i^2}{2}A_i^{2/3}\varepsilon_i^{8/3}. \quad (5.4.3)$$

i) We have the following approximation for the value of deferral w_i in \mathbf{NT}_i :

$$w_i(b, \xi)|_{b=1} = \frac{\gamma}{\bar{c}_{\alpha,i}} \frac{\sigma_i^2}{2} m_i A_i^{2/3} \varepsilon_i^{8/3} + o(\varepsilon_i^{8/3}), \quad (5.4.4)$$

$$\approx \frac{\gamma}{\bar{c}_{\alpha,i}} \theta_i m_i. \quad (5.4.5)$$

where $\{m_i, i \in \mathcal{I}\}$ is the solution of the following linear system:

$$\left(1 + p\kappa\left(\frac{\bar{K}_{\alpha,i}}{\bar{c}_{\alpha,i}} - 1\right)\right)(\theta_i m_i) - \sum_{j \in \mathcal{I}} \frac{\bar{G}_{\alpha,j}^{1-\gamma}}{\bar{G}_{\alpha,i}^{1-\gamma}} \frac{q_{ij}}{\bar{c}_{\alpha,j}} (\theta_j m_j) = \theta_i m_*^\delta, \quad i \in \mathcal{I}, \quad (5.4.6)$$

with $m_* : \delta \mapsto m_*^\delta$ uniquely determined by

$$\begin{cases} (1 - \delta p)^2 g''(p) + p - m_*^\delta - (pg'(p))^{2/3} = 0, & p \in [0, \frac{1}{\delta}], \\ g(0) = 0, g'(0) = 0. \end{cases} \quad (5.4.7)$$

ii) We have an approximation to the optimal trading boundaries $\xi_i^\pm(b)$, $b \in [0, 1]$:

$$\xi_i^\pm(b) \approx \bar{\xi}_i \pm \varepsilon_i \sqrt{1-b} \left(\frac{g'_i(p_i)}{\sqrt{p_i}} \right)^{1/3} \Big|_{p_i = \frac{1-b}{\delta_i}}, \quad (5.4.8)$$

iii) We have an approximation to the optimal consumption strategy for $i \in \mathcal{I}$:

$$\frac{C^*}{x + (1 - \alpha)y + \alpha k} = \bar{c}_{\alpha,i} + \frac{\kappa - 1}{\kappa} \frac{\sigma_i^2}{2} \gamma m_i A_i^{2/3} \varepsilon_i^{8/3} + o(\varepsilon_i^{8/3}). \quad (5.4.9)$$

We remark that the only unknown parameters m_i in the expansion (5.4.4) are determined by m_*^δ , which depends on the parameter δ only. Using shooting method, we obtain numerically the constant m_*^δ as a function $m_* : \delta \mapsto m_*^\delta$ in Fig.5.2. It turns out that m_*^δ is a decreasing function of δ and has a maximum of about $m_*^\delta = 0.78$ for $\delta = 0$. Moreover, we observe numerically that the solution g of (5.4.7) verifies

$$g'(p) \approx \eta_\delta p, \quad (5.4.10)$$

where η_δ is a constant depending on δ (see Fig.5.2). Hence we obtain an explicit approximation for the trading boundaries

$$\xi_i^\pm(b) \approx \bar{\xi}_i \pm \eta_{\delta_i}^{1/3} \delta_i^{-1/6} \varepsilon_i (1-b)^{2/3}. \quad (5.4.11)$$

Combining (5.3.22) and (5.3.7), Theorem 5.4.1 states that, up to first order approximation, the optimal trading strategy is to keep the proportion of risk asset (over liquidation wealth) around

$$\xi_i^* = \bar{\xi}_{\alpha,i}, \quad (5.4.12)$$

which is the "Merton line" of the tax-deflated model (5.3.4), such that

$$\left| \frac{y_t}{z_t} - \bar{\xi}_{\alpha,i} \right| \leq \frac{1}{1-\alpha} \eta_{\delta_i}^{1/3} \delta_i^{-1/6} \varepsilon_i (1-b_t)^{2/3}. \quad (5.4.13)$$

11. This argument for this result is heuristic without rigorous justification.

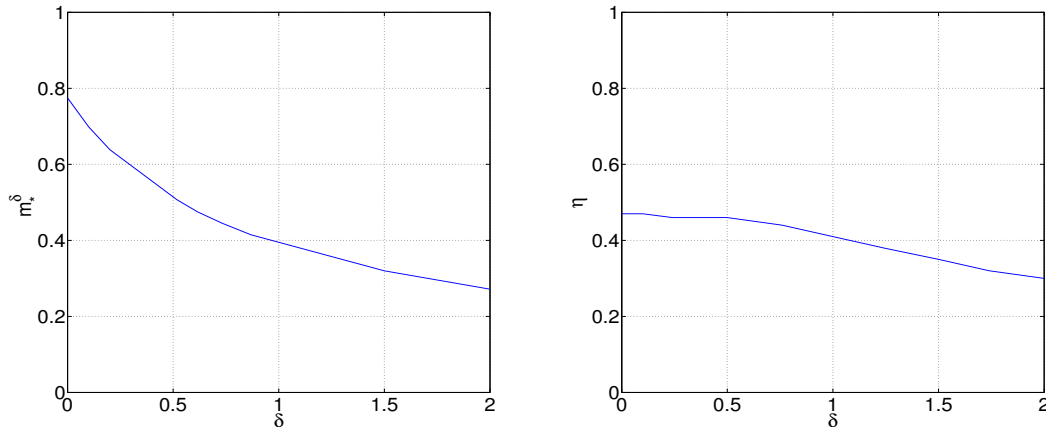


Figure 5.2 – The constants m_*^δ in (5.4.7) and η_δ in (5.4.10)

5.4.3 Economic analysis: single-regime case

We begin by analysing the case of single regime market. Fixing a market regime $i_0 \in \mathcal{I}$ and taking $q_{i_0 j} = 0, \forall j \in \mathcal{I}$ in Theorem 5.4.1, we obtain¹²

$$w(b, \xi)|_{b=1} = \frac{\gamma}{\bar{c}_\alpha} \frac{\sigma^2}{2} m_*^\delta A^{2/3} \varepsilon^{8/3} + o(\varepsilon^{8/3}), \quad (5.4.14)$$

where m_*^δ is given in Fig.5.2.

The impact of capital gains taxes is $O(\varepsilon^{8/3})$

In the classical optimal consumption investment problem, it is well-known that the utility loss due to proportional transaction costs is of order $O(\varepsilon^{2/3})$ if the proportional transaction cost coefficient is ε . It is the consequence of minimizing the utility losses due to deviation from the Merton line and the cumulated transaction costs.

In our case, the deferral value w is the result of a delicate balancing between the loss due to deviation from the Merton line and the benefit of deferring capital gains. Recall that

$$f(b, \xi) = \gamma[\varepsilon^2(1-b) - (\xi - \bar{\xi})^2].$$

The term $-(\xi - \bar{\xi})^2$ corresponds to the utility loss and $\varepsilon^2(1-b)$ corresponds to the benefit of deferring capital gains. As already pointed out in the literature, the deferral value originates from the non-zero interest $r = O(\varepsilon^2)$. Here we obtain a more precisely expression for the deferral benefit: it is actually weighted by $(1-b)$. Therefore, the investor should keep high capital gains (b small) while stay close to the Merton line ($(\xi - \bar{\xi})^2$ small). On the one hand, as ε tends to zeros, the non-transaction zone $\xi^\pm - \bar{\xi}$ should shrink to the Merton line $\xi = \bar{\xi}$, reducing utility loss of deviating from Merton line. On the other hand, a narrow non-transaction zone would incur frequent trading and push the tax basis b towards 1. Consequently, the benefit of deferring capital gains taxes $\varepsilon^2(1-b)$ decreases.

Different from the impact of transaction costs, our results show that the impact of capital gains taxes w is of order $O(\varepsilon^{8/3})$, i.e. $O(\alpha^{4/3})$. Indeed, we can show heuristically that if the non-transaction zone $\xi^\pm - \bar{\xi}$ has width of order $O(\varepsilon^{2\nu})$, then $(1-b)$ would be of order $O(\varepsilon^\nu)$.

12. For notational simplicity, we omit the index i_0 in this part.

Therefore the deviation loss $-(\xi - \bar{\xi})^2$ is about $O(\varepsilon^{4\nu})$ while the deferral benefit $\varepsilon^2(1 - b)$ is $O(\varepsilon^{2+\nu})$. To attain the optimal solution, we must have

$$4\nu \geq 2 + \nu \Rightarrow \nu \geq \frac{2}{3}.$$

At the same time, we want the deferral value to be as large as possible. Hence ν should be as large as possible, i.e. $\nu = 2/3$. The order of deferral value is

$$O(\varepsilon^{4\nu}) = O(\varepsilon^{2+\nu}) = O(\varepsilon^{8/3}).$$

In the Appendix, we provide a detailed explanation for the order $O(\varepsilon^{8/3})$ of w and a related optimal control problem for the constant m_*^δ .

Impact of EIS

We rewrite (5.4.14) as

$$\bar{c}_\alpha w \approx \frac{\sigma^2}{2} m_*^\delta A^{2/3} \varepsilon^{8/3}, \quad (5.4.15)$$

where the right hand side does not depend on the EIS. It means that the agent's EIS influence the deferral value only through the optimal consumption rate. Recall that $\bar{c}_\alpha = \bar{K}_\alpha$ under single regime model, we have

$$\frac{\partial \bar{c}_\alpha}{\partial \kappa} = -\frac{1}{\kappa^2} \left(\beta - (1 - \alpha)r - \frac{(\mu - r)^2}{2\gamma\sigma^2} \right).$$

Consequently,

1. If $\beta - (1 - \alpha)r - \frac{(\mu - r)^2}{2\gamma\sigma^2} > 0$, the deferral value w decreases when the EIS is higher;
2. If $\beta - (1 - \alpha)r - \frac{(\mu - r)^2}{2\gamma\sigma^2} < 0$, the deferral value w increases when the EIS is higher.

On the other hand, since (5.4.7) depends only on δ but not κ , the trading boundaries are independent of the EIS. This is consistent with the tax-free case (5.3.3) where κ impacts the optimal consumption rate \bar{c} but not the optimal proportion of risky asset $\bar{\xi}$.

5.4.4 Economic analysis: two-regime case

Consider the case of two regimes where $\mathcal{I} = \{1, 2\}$ and

$$Q = \begin{pmatrix} -\lambda_1 & \lambda_1 \\ \lambda_2 & -\lambda_2 \end{pmatrix},$$

with $\lambda_1, \lambda_2 > 0$. The system (5.4.6) can be rewritten as

$$\begin{pmatrix} M_{11}^\lambda & M_{12}^\lambda \\ M_{21}^\lambda & M_{22}^\lambda \end{pmatrix} \begin{pmatrix} \theta_1 m_1 \\ \theta_2 m_1 \end{pmatrix} = \begin{pmatrix} \theta_1 m_*^{\delta_1} \\ \theta_2 m_*^{\delta_2} \end{pmatrix}, \quad (5.4.16)$$

where $M^\lambda = (M_{ij}^\lambda)$ is given by

$$M_{11}^\lambda = 1 + p\kappa \left(\frac{\bar{K}_{\alpha,1}}{\bar{c}_{\alpha,1}} - 1 \right) + \frac{\lambda_1}{\bar{c}_{\alpha,1}}, \quad (5.4.17)$$

$$M_{12}^\lambda = -p\kappa \left(\frac{\bar{K}_{\alpha,1}}{\bar{c}_{\alpha,2}} - \frac{\bar{c}_{\alpha,1}}{\bar{c}_{\alpha,2}} \right) - \frac{\lambda_1}{\bar{c}_{\alpha,2}}, \quad (5.4.18)$$

$$M_{21}^\lambda = -p\kappa \left(\frac{\bar{K}_{\alpha,2}}{\bar{c}_{\alpha,1}} - \frac{\bar{c}_{\alpha,2}}{\bar{c}_{\alpha,1}} \right) - \frac{\lambda_2}{\bar{c}_{\alpha,1}}, \quad (5.4.19)$$

$$M_{22}^\lambda = 1 + p\kappa \left(\frac{\bar{K}_{\alpha,2}}{\bar{c}_{\alpha,2}} - 1 \right) + \frac{\lambda_2}{\bar{c}_{\alpha,2}}. \quad (5.4.20)$$

Value of deferral and impact of transition intensities

Using the expansion (5.4.4), we can evaluate not only the impact of capital gains taxes under different economics but also the dependence of the impact in terms of the transition intensities $\lambda = (\lambda_1, \lambda_2)$. Let $(w_1^\lambda, w_2^\lambda)$ the corresponding deferral value. By (5.4.5), we have

$$\begin{pmatrix} M_{11}^\lambda & M_{12}^\lambda \\ M_{21}^\lambda & M_{22}^\lambda \end{pmatrix} \begin{pmatrix} \bar{c}_{\alpha,1}^\lambda w_1^\lambda \\ \bar{c}_{\alpha,2}^\lambda w_2^\lambda \end{pmatrix} \approx \gamma \begin{pmatrix} \theta_1 m_*^{\delta_1} \\ \theta_2 m_*^{\delta_2} \end{pmatrix}, \quad \forall \lambda \in \mathbb{R}_+^2. \quad (5.4.21)$$

where the right hand side is independent of λ , and we use superscript λ to indicate the dependence of $\bar{c}_{\alpha,i}$, $i = 1, 2$ on λ . Denote (w_1^0, w_2^0) for $\lambda = (0, 0)$, which corresponds to the deferral values under each economics. Note that $M^\lambda = \text{Id}$ when $\lambda = (0, 0)$. Then the right hand side of (5.4.21) corresponds to $\lambda = 0$ and we obtain that

$$\begin{pmatrix} M_{11}^\lambda & M_{12}^\lambda \\ M_{21}^\lambda & M_{22}^\lambda \end{pmatrix} \begin{pmatrix} \bar{c}_{\alpha,1}^\lambda \\ \bar{c}_{\alpha,2}^\lambda \end{pmatrix} \begin{pmatrix} w_1^\lambda \\ w_2^\lambda \end{pmatrix} \approx \begin{pmatrix} \bar{c}_{\alpha,1}^0 \\ \bar{c}_{\alpha,2}^0 \end{pmatrix} \begin{pmatrix} w_1^0 \\ w_2^0 \end{pmatrix},$$

or equivalently,

$$\begin{pmatrix} w_1^\lambda \\ w_2^\lambda \end{pmatrix} \approx \begin{pmatrix} M_{11}^\lambda & 1 - M_{11}^\lambda \\ 1 - M_{22}^\lambda & M_{22}^\lambda \end{pmatrix}^{-1} \begin{pmatrix} \bar{c}_{\alpha,1}^0 / \bar{c}_{\alpha,1}^\lambda \\ \bar{c}_{\alpha,2}^0 / \bar{c}_{\alpha,2}^\lambda \end{pmatrix} \begin{pmatrix} w_1^0 \\ w_2^0 \end{pmatrix}. \quad (5.4.22)$$

We can write

$$\begin{aligned} w_1^\lambda &= \frac{1}{M_{11}^\lambda} \frac{\bar{c}_{\alpha,1}^0}{\bar{c}_{\alpha,1}^\lambda} w_1^0 + \left(1 - \frac{1}{M_{11}^\lambda}\right) w_2^\lambda, \\ w_2^\lambda &= \frac{1}{M_{22}^\lambda} \frac{\bar{c}_{\alpha,2}^0}{\bar{c}_{\alpha,2}^\lambda} w_2^0 + \left(1 - \frac{1}{M_{22}^\lambda}\right) w_1^\lambda. \end{aligned}$$

By (5.4.17) and (5.4.20), M_{11}^λ and M_{22}^λ increase as λ_1 and λ_2 increase, hence the above relation shows that w_1^λ and w_2^λ will move towards each other. This is a typical behaviour of values functions for models with regime-switching. Our asymptotic expansion captures this feature in an explicit way.

Trading boundaries

Since the optimal proportional of risky asset $\bar{\xi}_i$ doesn't depend on the EIS or the transition intensities (see (5.3.3)), the non-transaction zone \mathbf{NT}_i should always be a wedged region around the line $\xi \equiv \bar{\xi}_i$.

On the other hand, the asymptotic trading boundaries ξ_i^\pm are determined by (5.4.7), hence is approximately independent of the transition intensities between the two regimes. This is true only if ε is small enough.

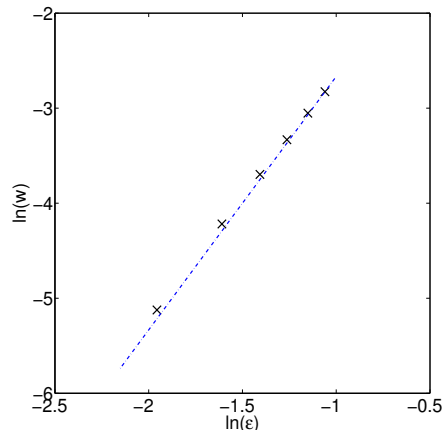
5.5 Numerical results

5.5.1 Single-regime case

In this section, we present our asymptotic results for single regime market. This allows us to analyze the joint impact of capital gains taxes and EIS on the value function and optimal trading boundaries. For comparison, we always employ the penalty method with the finite difference discretization (see [FV02]) to generate benchmark values.¹³

13. It should be emphasized that implementing the penalty method with the finite difference discretization is absolutely non-trivial. Because the resulting matrix is, in general, not an M-matrix for the present problem, one

α	8%	16%	24%	32%	40%	48%
δ	1.32	1.66	1.90	2.09	2.25	2.39
ε	0.14	0.20	0.24	0.28	0.32	0.35

Table 5.1 – Tax rates and corresponding values of δ .

Black points: numerical results for (5.5.1). Blue dashed line: the reference line $\log(\varepsilon) \mapsto \frac{8}{3} \log(\varepsilon)$.

Parameters: $\beta = 0.05$, $r = 3\%$, $\mu = 0.09$, $\sigma = 0.20$, $\gamma = 3$, $\kappa = 0.7$.

Figure 5.3 – The order of deferral value is $O(\varepsilon^{8/3})$

The value of deferring capital gains realization

After the expansion (5.4.4), (5.5.1) should be close to the line $\log(\varepsilon) \mapsto \frac{8}{3} \log(\varepsilon)$. We take tax rate from 1% to 48% (see Table 5.1). Fig.5.3 shows a very good fit for the order of w .

In this section, we examine the asymptotic expansion (5.4.14) for the option value of deferring capital gains. We compare the following function

$$\log(\varepsilon) \mapsto \log(w) - \log\left(\frac{\gamma}{\bar{c}_\alpha} A^{2/3} m_*^\delta\right). \quad (5.5.1)$$

Fig.5.4 compares the expansion (5.4.4) and the numerical results of the deferral value. We introduce

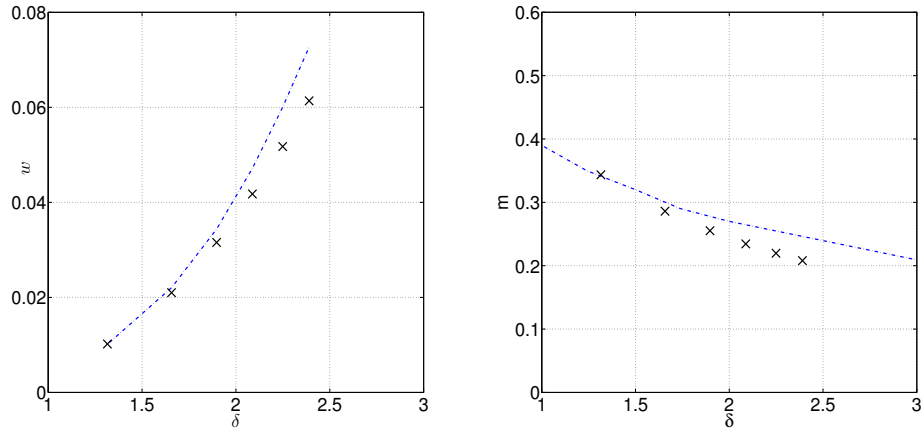
$$m := \left(\gamma \frac{\sigma^2}{2} \frac{1}{\bar{c}_\alpha} A^{2/3} \varepsilon^{8/3}\right)^{-1} w(1, \bar{\xi}), \quad (5.5.2)$$

where $w(1, \bar{\xi})$ is obtained numerically. Our formula (5.4.14) exhibits an extremely good approximation (Fig.5.4).

On the other hand, we can compare CEWL with its upper bound $\log \frac{\bar{G}}{\bar{G}_\alpha}$ by computing

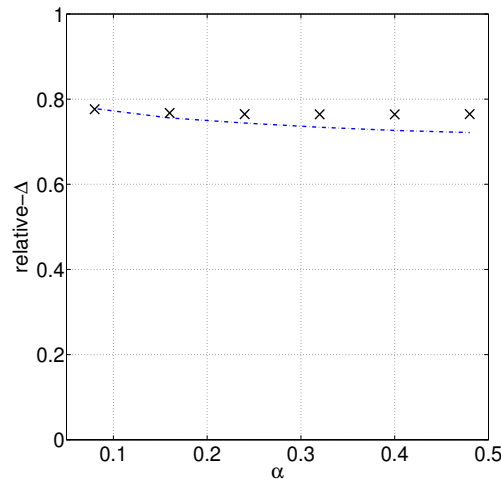
$$\text{relative-}\Delta := \frac{\Delta}{\log(\bar{G}/\bar{G}_\alpha)}. \quad (5.5.3)$$

needs to carefully choose the penalty parameter to guarantee convergence. Moreover, it is very time-consuming and often divergent for small ε . That is why we seek for asymptotic analysis.



Left panel: deferral values Right panel: black points for (5.5.2) vs. blue dashed line for (5.4.7).
Parameters: $\beta = 0.05$, $r = 3\%$, $\mu = 0.08$, $\sigma = 0.25$, $\gamma = 3$, $\kappa = 0.7$.

Figure 5.4 – Asymptotic expansion in terms of tax rate



Parameters: $\beta = 0.05$, $r = 3\%$, $\mu = 0.08$, $\sigma = 0.25$, $\gamma = 3$, $\kappa = 0.7$.

Figure 5.5 – Relative CEWL (5.5.3)

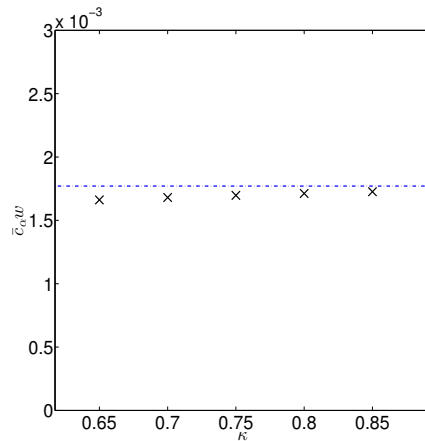
We observe that the deferral value accounts for up to 30% of the gap in [BST10](Fig.5.5). This means that, by optimally deferring the realisation of capital gains taxes, the investor can obtain up to 30% of extra benefit in terms of expected utility.

Impact of capital gains taxes and EIS

In this section we study the impact of tax for investors of different EIS. As predicted by the asymptotic expansion (5.4.14), we obtain a horizontal line in Fig.5.6. Fig.5.7 reports different monotonicity of w in terms of κ , corresponding to $\beta - (1 - \alpha)r - \frac{(\mu - r)^2}{2\gamma\sigma^2} < 0$ (> 0 respectively).

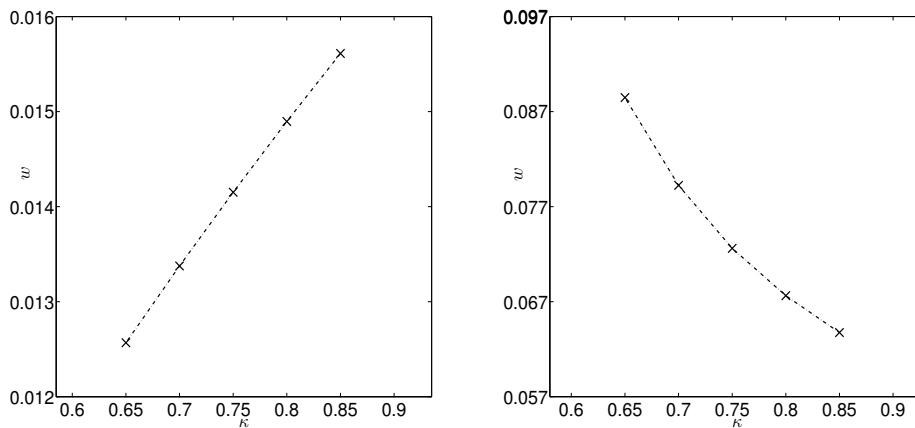
Approximations to the optimal trading boundary

The trading boundaries provided by asymptotic expansion depend only on ε and δ (through g). In particular, they are independent of EIS. This is supported by the numerical results in the



Numerical values of $\bar{c}_\alpha w$ (black points) vs. asymptotic expansion (blue dashed line). We take $\kappa \in \{0.65, 0.70, 0.75, 0.80, 0.85\}$. Parameters: $\beta = 0.1$, $\alpha = 40\%$, $r = 5\%$, $\mu = 0.15$, $\sigma = 0.35$, $\gamma = 4$.

Figure 5.6 – The value of $\bar{c}_\alpha w$ in terms of EIS



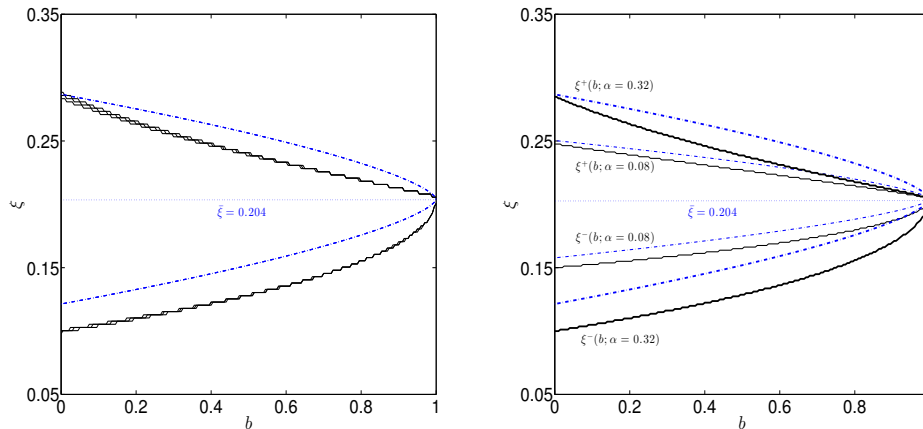
Left panel: $\beta = 0.10$, $\mu = 0.15$. Right panel: $\beta = 0.02$, $\mu = 0.10$. We take $\kappa \in \{0.65, 0.70, 0.75, 0.80, 0.85\}$. Other parameters: $\alpha = 40\%$, $r = 5\%$, $\sigma = 0.35$, $\gamma = 4$.

Figure 5.7 – Impact of tax (i.e. deferral value w) in terms of EIS

left panel of Fig.5.8. We observe also that the explicit approximation (5.4.11) is quite close to numerical results, in particular for small values of α (see right panel of Fig.5.8).

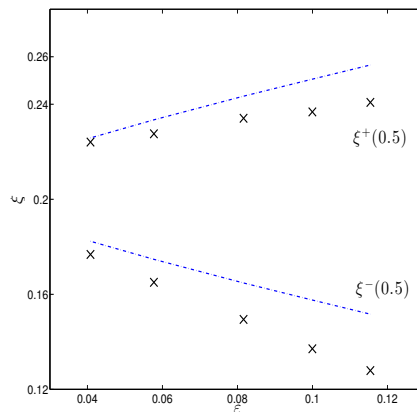
To further examine the effect of approximation (5.4.11), we plot in Fig.5.9 the optimal trading boundaries at $b = 0.5$ against ε , i.e. $\xi^\pm(0.5; \varepsilon)$, using the approximation and the penalty method respectively. We let α vary from 4% to 32%. It can be seen that the approximation is reasonably good even with high tax rate. The same pattern is observed for varying values of interest rate (right panel of Fig.5.9) and different value of $b \in [0, 1]$.

An interesting phenomenon is that the trading boundaries are not symmetric with respect to the Merton line $\xi \equiv \bar{\xi}$: the selling boundary is closer to the Merton line than the buying boundary does. This is due to the fact that selling stock does not alter tax basis b while buying stock increases tax basis (hence reduces the deferral benefit $\varepsilon^2(1 - b)$.) As the utility loss $-(\xi - \bar{\xi})^2$ is symmetric to both positive and negative deviation, the investor will tend to avoid positive



Trading boundaries: numerical (red curves) vs approximation (blue curve, (5.4.11)) where $\bar{\xi} = 0.204$. Left panel: $\kappa \in \{0.6, 0.7, 0.8\}$, $\alpha = 32\%$. Right panel: $\alpha \in \{8\%, 32\%\}$, $\kappa = 0.7$. Parameters: $\beta = 0.1$, $r = 5\%$, $\mu = 0.15$, $\sigma = 0.35$, $\gamma = 4$. We have $\eta_\delta \approx 0.4$.

Figure 5.8 – The dependence of trading boundary on EIS and tax rate



Parameters: $\alpha \in \{4\%, 8\%, 16\%, 24\%, 32\%\}$, $r = 0.05$, $\beta = 0.1$, $\kappa = 0.7$, $\mu = 0.15$, $\sigma = 0.35$, $\gamma = 4$. The red line and points represent the optimal trading boundaries computed from the penalty method with finite difference discretization and the blue line is the given by (5.4.11). The x -axis is the corresponding values of ε .

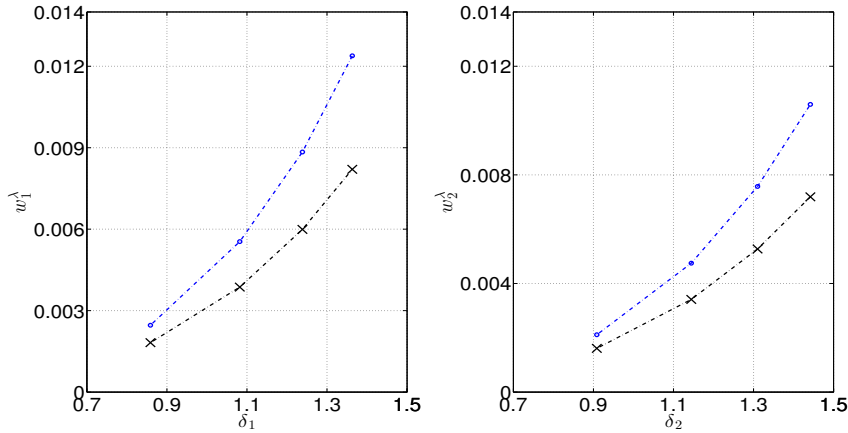
Figure 5.9 – Approximation (5.4.11) at $b = 0.5$ with varying α .

deviation while being more patient for negative deviation. Our expansion (5.4.8) is symmetric since when ε is small, such effect can be and is ignored. This limitation should be kept in mind when ε is not small enough.

5.5.2 Two-regime Case

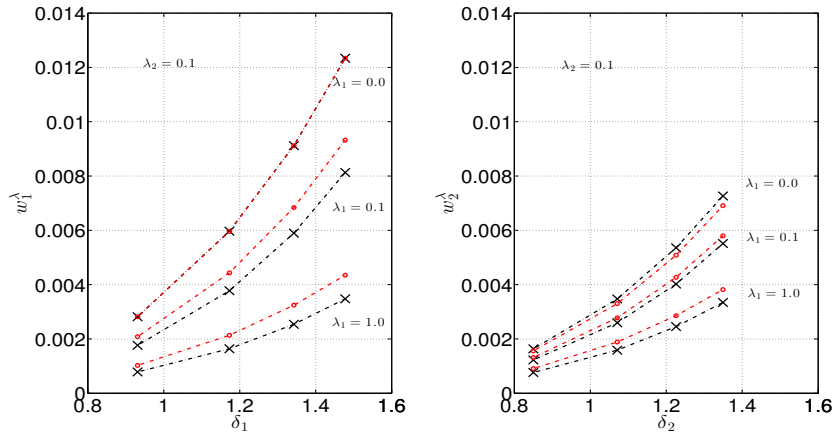
We consider two economics where for $i = 1$, the optimal strategy is $\bar{\xi}_1 = 0.19$, and the economics $i = 2$ is less risky with $\bar{\xi}_2 = 0.42$. Fig.5.10 reports the performance of the expansion (5.4.4) in two-regime case. We observe effectively a better approximation as the tax rate decreases. Now we examine the relation (5.4.22). We use deferral values under single regimes for w_1^0 and w_2^0 . Fixing λ_2 , the two deferral values decreases to the same value w_2^0 as λ_1 increases, see Fig.5.11.

The red lines are computed using w_1^0 , w_2^0 and (5.4.22). The relation (5.4.22) captures well the impact of capital gains taxes under regime-switching market.



Black curves: numerical results. Blue curves: Asymptotic expansion (5.4.4).
 Parameters: $\lambda_1 = 0.5$, $\lambda_2 = 0.2$, $\alpha \in \{10\%, 20\%, 30\%, 40\%\}$, $r_1 = 0.05$, $\sigma_1 = 0.35$, $\mu_1 = 0.15$, $r_2 = 0.01$, $\sigma_2 = 0.25$, $\mu_2 = 0.11$, $\kappa = 0.7$, $\beta = 0.10$, $\gamma = 3$.

Figure 5.10 – Asymptotic for two-regime deferral values in terms of tax rate

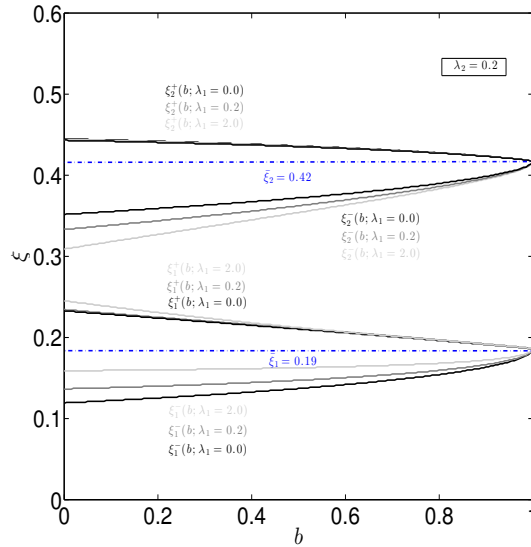


Black curves: numerical results. Red curves: Asymptotic expansion (5.4.22).
 Parameters: $\lambda_1 \in \{0, 0.1, 1\}$, $\lambda_2 = 0.1$, $\alpha \in \{10\%, 20\%, 30\%, 40\%\}$, $r_1 = 0.05$, $\sigma_1 = 0.3$, $\mu_1 = 0.1$, $r_2 = 0.01$, $\sigma_2 = 0.20$, $\mu_2 = 0.06$, $\kappa = 0.7$, $\beta = 0.10$, $\gamma = 3$.

Figure 5.11 – Impact of transition intensities on deferral values

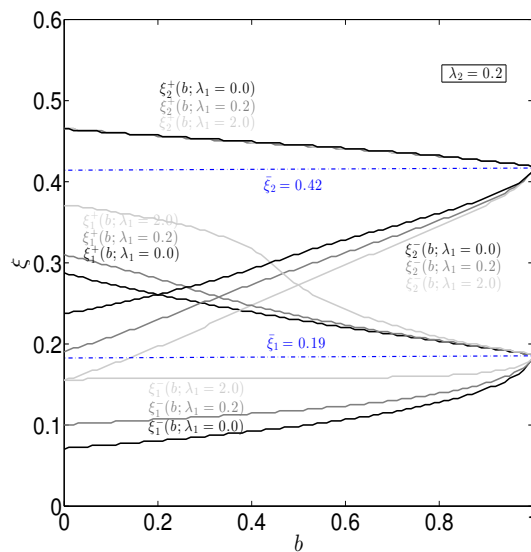
Fig.5.12 reports the trading boundaries for the case where the tax rate is small ($\alpha = 0.08$). The trading boundaries contain the tax-free strategy $\xi_i \equiv \bar{\xi}_i$ and do not change significantly when λ_1 increases from zero to 2. The expansion (5.4.8) fits well the numerical results.

If the tax rate is high ($\alpha = 0.20$), Fig.5.13 shows that $\xi_i^\pm(1) \simeq \bar{\xi}_i$ and are not sensible to the transition intensities λ_1 and λ_2 , as indicated by Theorem 5.4.1. However, second order effects begin to appear. We can observe that the trading boundaries of the two regimes move toward each other as the transition intensity λ_1 increases, in particular in the region of high capital gains (b close to 0). The only exception is the selling boundary $\xi_2^+(b)$. This is due to the



Parameters: $\lambda_1 \in \{0, 0.2, 2\}$, $\lambda_2 = 0.2$, $\alpha = 0.08$, $r_1 = 0.05$, $\sigma_1 = 0.3$, $\mu_1 = 0.1$, $r_2 = 0.01$, $\sigma_2 = 0.20$, $\mu_2 = 0.06$, $\kappa = 0.7$, $\beta = 0.10$, $\gamma = 3$.

Figure 5.12 – Impact of transition intensities on trading boundaries: small tax rate



Parameters: $\lambda_1 \in \{0, 0.2, 2\}$, $\lambda_2 = 0.2$, $\alpha = 0.2$, $r_1 = 0.05$, $\sigma_1 = 0.3$, $\mu_1 = 0.1$, $r_2 = 0.01$, $\sigma_2 = 0.20$, $\mu_2 = 0.06$, $\kappa = 0.7$, $\beta = 0.10$, $\gamma = 3$.

Figure 5.13 – Impact of transition intensities on trading boundaries

fact that, when the market regime changes, the investor has to adjust her portfolio to the new non-transaction zone. In general this will cause the increase of tax basis, except for the highest selling boundary $\xi_2^+(b)$. As a consequence, when the transition intensities increase, the investor will add an extra buffer zone for the original non-trading zone.

5.6 Why is the order of deferral value $O(\varepsilon^{8/3})$?

We present here a probabilistic explanation for the value of deferring capital gains taxes w , in particular the fact that $w = O(\varepsilon^{8/3})$. We provide also an heuristic interpretation of m_*^δ in terms of an optimal control problem. To simplify the notation, we consider the classical additive utility maximisation under a single-regime market, i.e. the model established by [BST10].

Recall that the portfolio under the strategy $\{(C_t, L_t, M_t), t \geq 0\}$ is given by

$$\begin{aligned} dX_t &= [(1 - \alpha)r_t X_t - C_t] dt - dL_t + [Y_{t-} - \alpha(Y_{t-} - K_{t-})] dM_t, \\ dY_t &= \mu_t Y_{t-} dt + \sigma_t Y_{t-} dB_t + dL_t - Y_{t-} dM_t, \\ dK_t &= dL_t - K_{t-} dM_t, \end{aligned}$$

and the expected utility is given by

$$V_0^{C,L,M} = \beta \mathbb{E} \left[\int_0^\infty e^{-\beta(s)} U(C_s, \gamma) ds \right], \quad (5.6.1)$$

Define the portfolio value after liquidation by Z_t where

$$Z_t = X_t + Y_t - \alpha(Y_t - K_t).$$

We have

$$\begin{aligned} dZ_t &= dX_t + (1 - \alpha)dY_t + \alpha dK_t \\ &= [(1 - \alpha)rZ_t - C_t] dt + (1 - \alpha)Y_t(\mu dt + \sigma dW_t - (1 - \alpha)r dt) - (1 - \alpha)\alpha r K_t dt. \end{aligned}$$

Denote

$$\xi_t = \frac{Y_t}{Z_t}, \quad b_t = \frac{K_t}{Y_t}, \quad c_t = \frac{C_t}{Z_t}, \quad (5.6.2)$$

and define the tax-deflated model by

$$r_\alpha = (1 - \alpha)r, \quad \mu_\alpha = (1 - \alpha)\mu, \quad \sigma_\alpha = (1 - \alpha)\sigma.$$

We have

$$\begin{aligned} dZ_t &= Z_t \{ [(1 - \alpha)r - c_t] dt + (1 - \alpha)\xi_t [\mu dt + \sigma dW_t - (1 - \alpha)r dt] - (1 - \alpha)\alpha r \xi_t b_t dt \} \\ &= Z_t \{ [(1 - \alpha)r - c_t] dt + (1 - \alpha)\xi_t [\mu dt + \sigma dW_t - r dt] + (1 - \alpha)\alpha r \xi_t (1 - b_t) dt \} \\ &= Z_t [(r_\alpha - c_t) dt + \xi_t (\mu_\alpha dt + \sigma_\alpha dW_t - r_\alpha dt) + \alpha r_\alpha \xi_t (1 - b_t) dt] \end{aligned} \quad (5.6.3)$$

5.6.1 Expansion around the tax-deflated model

The main idea is to consider the case where the interest rate or tax rate are small. Then the extra benefit from tax deferral is small since the terms $1 - b_t^\varepsilon$ and $\xi_t^\varepsilon - \bar{\xi}_\alpha^\varepsilon$ should be small, where $\bar{\xi}_\alpha^\varepsilon$ is the optimal investment strategy under tax-deflated model. Consider any consumption-investment strategy $\{(C_t^\varepsilon, L_t^\varepsilon, M_t^\varepsilon), t \geq 0\}$ with $\{(c_t^\varepsilon, \xi_t^\varepsilon, b_t^\varepsilon), t \geq 0\}$ determined by (5.6.2), then the corresponding liquidation wealth Z_t^ε satisfies

$$\begin{aligned} dZ_t^\varepsilon &= Z_t^\varepsilon [(r_\alpha - c_t^\varepsilon) dt + \xi_t^\varepsilon (\mu_\alpha dt - r_\alpha dt + \sigma_\alpha dW_t) + \alpha r_\alpha \xi_t^\varepsilon (1 - b_t^\varepsilon) dt], \\ Z_0^\varepsilon &= z. \end{aligned}$$

We rewrite

$$Z_t^\varepsilon = z e^{M_t^\varepsilon}, \quad (5.6.4)$$

where

$$\begin{aligned} M_t^\varepsilon &= \int_0^t (r_\alpha - c_s^\varepsilon) ds + \int_0^t \xi_s^\varepsilon (\mu_\alpha dt - r_\alpha dt + \sigma_\alpha dW_t) \\ &\quad + \int_0^t \alpha r_\alpha \xi_s^\varepsilon (1 - b_s^\varepsilon) ds - \frac{1}{2} \int_0^t (\xi_s^\varepsilon)^2 \sigma_\alpha^2 ds. \end{aligned}$$

Set the consumption rate to be the optimal rate under the tax-deflated model:

$$c_t^\varepsilon \equiv \bar{c}_\alpha.$$

Then we have

$$\begin{aligned} & -\beta t + (1 - \gamma) M_t^\varepsilon \\ &= \int_0^t (-\beta) ds + (1 - \gamma) \int_0^t (r_\alpha - \bar{c}_\alpha) ds + (1 - \gamma) \int_0^t \xi_s^\varepsilon (\mu_\alpha dt - r_\alpha dt + \sigma_\alpha dW_t) \\ &\quad - \frac{1}{2} (1 - \gamma) \int_0^t (\xi_s^\varepsilon)^2 \sigma_\alpha^2 ds + (1 - \gamma) \int_0^t \alpha r_\alpha \xi_s^\varepsilon (1 - b_s^\varepsilon) ds \\ &= \int_0^t \sigma_\alpha (1 - \gamma) \xi_s^\varepsilon dW_s - \frac{1}{2} \int_0^t (1 - \gamma)^2 (\xi_s^\varepsilon)^2 \sigma_\alpha^2 ds \\ &\quad + \int_0^t \left[-\beta + (1 - \gamma)(r_\alpha - \bar{c}_\alpha) + \frac{1}{2} (-\gamma)(1 - \gamma) \sigma_\alpha^2 (\xi_s^\varepsilon)^2 + (1 - \gamma)(\mu_\alpha - r_\alpha) \xi_s^\varepsilon \right] ds \\ &\quad + \int_0^t (1 - \gamma) \alpha r_\alpha \xi_s^\varepsilon (1 - b_s^\varepsilon) ds \\ &= \int_0^t (1 - \gamma) \sigma_\alpha \xi_s^\varepsilon dW_s - \frac{1}{2} \int_0^t (1 - \gamma)^2 (\xi_s^\varepsilon)^2 \sigma_\alpha^2 ds - \bar{c}_\alpha t \\ &\quad - \frac{1}{2} (1 - \gamma) \gamma \sigma_\alpha^2 \int_0^t (\xi_s^\varepsilon - \bar{\xi}_\alpha^\varepsilon)^2 ds + \int_0^t (1 - \gamma) \alpha r_\alpha \xi_s^\varepsilon (1 - b_s^\varepsilon) ds \end{aligned} \tag{5.6.5}$$

Now define

$$\frac{d\mathbb{Q}_\alpha^\varepsilon}{d\mathbb{P}} \Big|_{\mathcal{F}_t} = \exp \left(\int_0^t (1 - \gamma) \sigma_\alpha \xi_s^\varepsilon dW_s - \frac{1}{2} \int_0^t (1 - \gamma)^2 (\xi_s^\varepsilon)^2 \sigma_\alpha^2 ds \right).$$

Note that

$$dW_{\alpha,t}^\varepsilon := dW_t - (1 - \gamma) \sigma_\alpha \xi_t^\varepsilon dt$$

is a $\mathbb{Q}_\alpha^\varepsilon$ -Brownian motion. Consequently,

$$\begin{aligned} & \mathbb{E} \left[\int_0^\infty e^{-\beta t} U(c_t^\varepsilon Z_t^\varepsilon) dt \right] \\ & \stackrel{(5.6.4)}{=} \mathbb{E} \left[\int_0^\infty e^{-\beta t} U(\bar{c}_\alpha z e^{M_t^\varepsilon}) dt \right] \\ & \stackrel{(5.2.4)}{=} U(\bar{c}_\alpha z) \mathbb{E} \left[\int_0^\infty \exp(-\beta t + (1 - \gamma) M_t^\varepsilon) dt \right] \\ & \stackrel{(5.6.5)}{=} U(\bar{c}_\alpha z) \mathbb{E}^{\mathbb{Q}_\alpha^\varepsilon} \left[\int_0^\infty e^{-\bar{c}_\alpha t} \exp \left((1 - \gamma) \int_0^t (\alpha r_\alpha \xi_s^\varepsilon (1 - b_s^\varepsilon) - \frac{\gamma \sigma_\alpha^2}{2} (\xi_s^\varepsilon - \bar{\xi}_\alpha^\varepsilon)^2) ds \right) dt \right] \end{aligned}$$

Using $e^x \simeq 1 + x$, we obtain

$$\begin{aligned} & \mathbb{E} \left[\int_0^\infty e^{-\beta t} U(c_t^\varepsilon Z_t^\varepsilon) dt \right] \\ & \simeq U(\bar{c}_\alpha z) \mathbb{E}^{\mathbb{Q}_\alpha^\varepsilon} \left[\int_0^\infty e^{-\bar{c}_\alpha t} \left(1 + (1 - \gamma) \int_0^t (\alpha r_\alpha \xi_s^\varepsilon (1 - b_s^\varepsilon) - \frac{\gamma \sigma_\alpha^2}{2} (\xi_s^\varepsilon - \bar{\xi}_\alpha^\varepsilon)^2) ds \right) dt \right] \\ & = U(\bar{c}_\alpha z) \int_0^\infty e^{-\bar{c}_\alpha t} dt + (1 - \gamma) U(\bar{c}_\alpha z) \mathbb{E}^{\mathbb{Q}_\alpha^\varepsilon} \left[\int_0^\infty e^{-\bar{c}_\alpha t} \left(\int_0^t (\alpha r_\alpha \xi_s^\varepsilon (1 - b_s^\varepsilon) - \frac{\gamma \sigma_\alpha^2}{2} (\xi_s^\varepsilon - \bar{\xi}_\alpha^\varepsilon)^2) ds \right) dt \right] \\ & = \bar{\varphi}_\alpha(z) + (1 - \gamma) U(\bar{c}_\alpha z) \mathbb{E}^{\mathbb{Q}_\alpha^\varepsilon} \left[\int_0^\infty e^{-\bar{c}_\alpha t} \left(\int_0^t (\alpha r_\alpha \xi_s^\varepsilon (1 - b_s^\varepsilon) - \frac{\gamma \sigma_\alpha^2}{2} (\xi_s^\varepsilon - \bar{\xi}_\alpha^\varepsilon)^2) ds \right) dt \right] \\ & = \bar{\varphi}_\alpha(z) + (1 - \gamma) U(\bar{c}_\alpha z) \int_0^\infty e^{-\bar{c}_\alpha t} \mathbb{E}^{\mathbb{Q}_\alpha^\varepsilon} \left[\int_0^t (\alpha r_\alpha \xi_s^\varepsilon (1 - b_s^\varepsilon) - \frac{\gamma \sigma_\alpha^2}{2} (\xi_s^\varepsilon - \bar{\xi}_\alpha^\varepsilon)^2) ds \right] dt \end{aligned}$$

To obtain extra welfare by tax-deferral, the investor needs to maximize essentially

$$\mathbb{E}^{\mathbb{Q}_\alpha^\varepsilon} \left[\int_0^t (\alpha r_\alpha \xi_s^\varepsilon (1 - b_s^\varepsilon) - \frac{\gamma \sigma_\alpha^2}{2} (\xi_s^\varepsilon - \bar{\xi}_\alpha^\varepsilon)^2) ds \right],$$

where the first term represents the benefit of tax-deferral and the second term utility loss due to deviation from the benchmark strategy $\bar{\xi}_\alpha^\varepsilon$.

Introducing the change of variable

$$\zeta_s^\varepsilon = (1 - \alpha) \xi_s^\varepsilon,$$

we obtain that

$$\begin{aligned} & \mathbb{E} \left[\int_0^\infty e^{-\beta t} U(c_t^\varepsilon Z_t^\varepsilon) dt \right] \\ & \simeq \bar{\varphi}_\alpha(z) + \frac{1}{2} (1 - \gamma) \gamma \sigma^2 \bar{c}_\alpha \bar{\varphi}_\alpha(z) \int_0^\infty e^{-\bar{c}_\alpha t} \mathbb{E}^{\mathbb{Q}_\alpha^\varepsilon} \left[\int_0^t \left(\frac{2\alpha(1-\alpha)\xi_s^\varepsilon r}{\gamma \sigma^2} (1 - b_s^\varepsilon) - (\zeta_s^\varepsilon - \bar{\xi}^\varepsilon)^2 \right) ds \right] dt, \\ & \simeq \bar{\varphi}_\alpha(z) + \frac{1}{2} (1 - \gamma) \gamma \sigma^2 \bar{c}_\alpha \bar{\varphi}_\alpha(z) \int_0^\infty e^{-\bar{c}_\alpha t} \mathbb{E}^{\mathbb{Q}_\alpha^\varepsilon} \left[\int_0^t (\varepsilon^2 (1 - b_s^\varepsilon) - (\zeta_s^\varepsilon - \bar{\xi}^\varepsilon)^2) ds \right] dt, \end{aligned} \quad (5.6.6)$$

where we have approximated ξ_s^ε by $\bar{\xi}_\alpha^\varepsilon$ in (5.6.6) and used the fact that $\bar{\xi}^\varepsilon = (1 - \alpha) \bar{\xi}_\alpha^\varepsilon$. The constants δ , A are given in (5.4.3).

In sum, to obtain an approximation of φ^ε , the investor is led to consider the following control problem

$$J_T^\varepsilon = \mathbb{E}^{\mathbb{Q}_\alpha^\varepsilon} \left[\int_0^T (\varepsilon^2 (1 - b_s^\varepsilon) - (\zeta_s^\varepsilon - \bar{\xi}^\varepsilon)^2) ds \right], \quad (5.6.7)$$

where, applying Itô formula to (5.6.2), the dynamics of (ζ_t^ε) and (b_t^ε) are given by

$$db_t^\varepsilon = -\sigma b_t^\varepsilon dW_{\alpha,t} + \frac{1 - b_t^\varepsilon}{\zeta_t^\varepsilon} dL_t^\varepsilon + dR_t^\varepsilon + o(1)dt, \quad (5.6.8)$$

$$d\zeta_t^\varepsilon = \sigma \zeta_t^\varepsilon (1 - \zeta_t^\varepsilon) dW_{\alpha,t} + dL_t^\varepsilon - dU_t^\varepsilon + o(1)dt. \quad (5.6.9)$$

The processes L_t^ε and U_t^ε correspond to buy and sell strategies while R_t^ε is the liquidation strategy realising capital losses and keeping b_t^ε inside the interval $[0, 1]$. To maximize (5.6.7), the investor should try to keep ζ_t^ε close to $\bar{\xi}^\varepsilon$. This in turn pushes b_t^ε towards 1-, reducing the benefit of tax deferral.

5.6.2 Strategies based on barriers and reduced model

In this section, we restrict ourselves to a particular class of investment strategies based on barriers.

1. *Change of Variables.* Consider the following change of variables

$$p_t^\varepsilon = 1 - b_t^\varepsilon, \quad q_t^\varepsilon = \zeta_t^\varepsilon - \bar{\xi}^\varepsilon,$$

we have

$$dp_t^\varepsilon = -\sigma(1 - p_t^\varepsilon) dW_{\alpha,t} + \frac{p_t^\varepsilon}{\xi_t^\varepsilon + q_t^\varepsilon} dL_t^\varepsilon + dR_t^\varepsilon + o(1)dt, \quad (5.6.10)$$

$$dq_t^\varepsilon = \sigma(\bar{\xi}^\varepsilon + q_t^\varepsilon)(1 - \bar{\xi}^\varepsilon - q_t^\varepsilon) dW_{\alpha,t} + dL_t^\varepsilon - dU_t^\varepsilon + o(1)dt. \quad (5.6.11)$$

The optimization objective becomes

$$J_T^\varepsilon = \mathbb{E}^{\mathbb{Q}_\alpha^\varepsilon} \left[\int_0^T (\varepsilon^2 p_t^\varepsilon - (q_t^\varepsilon)^2) dt \right].$$

2. *Strategies Based on Barriers* l . Let $l : [0, \frac{1}{\delta}] \rightarrow \mathbb{R}_+$ be a continuous function such that $l(0) = 0$ and $l(p) > 0$ for $p \in (0, 1/\delta]$. The investment strategy $(L_t^\varepsilon, U_t^\varepsilon, R_t^\varepsilon)$ based on the barrier l consists of three non-decreasing processes keeping $(p_t^\varepsilon, q_t^\varepsilon)$ inside the domain G :

$$G := \{(p, q) \in [0, 1] \times \mathbb{R} \mid -\delta^2 l(p) \leq q \leq \delta^2 l(p), \forall p \in [0, 1/\delta]\},$$

such that

$$\int_0^T \mathbb{1}_{\{p_t^\varepsilon \neq 0\}} dR_t^\varepsilon + \int_0^T \mathbb{1}_{\{q_t^\varepsilon \neq l(p_t^\varepsilon)\}} dU_t^\varepsilon + \int_0^T \mathbb{1}_{\{q_t^\varepsilon \neq -l(p_t^\varepsilon)\}} dL_t^\varepsilon = 0.$$

3. *Change of Scaling*. Now consider the following change of scaling

$$\tilde{p}_t^\varepsilon = \frac{1}{\delta} p_{\delta^2 t}^\varepsilon, \quad \tilde{q}_t^\varepsilon = \frac{1}{\delta^2} q_{\delta^2 t}^\varepsilon. \quad (5.6.12)$$

The optimisation problem becomes

$$\frac{J_T^\varepsilon}{T} = I^\delta A^{2/3} \varepsilon^{8/3}, \quad (5.6.13)$$

with I^δ given by

$$I^\delta = \frac{1}{T^\varepsilon} \mathbb{E}^{\mathbb{Q}_\alpha^\varepsilon} \left[\int_0^{T^\varepsilon} (\tilde{p}_t^\varepsilon - A^2 (\tilde{q}_t^\varepsilon)^2) dt \right], \quad (5.6.14)$$

Here, $T^\varepsilon = \delta^{-2} T \rightarrow \infty$ as $\varepsilon \rightarrow 0$, and

$$d\tilde{p}_t^\varepsilon = \sigma(1 - \delta \tilde{p}_t^\varepsilon) d\tilde{W}_{\alpha, t}^\varepsilon - (1 + o(1)) \frac{\tilde{p}_t^\varepsilon}{\bar{\xi}} dL_{\delta^2 t}^\varepsilon + d\tilde{R}_{\delta^2 t}^\varepsilon - o(\delta) dt, \quad (5.6.15)$$

$$d\tilde{q}_t^\varepsilon = \frac{1}{\delta} \sigma(\bar{\xi} + o(1))(1 - \bar{\xi} + o(1)) d\tilde{W}_{\alpha, t}^\varepsilon + \frac{1}{\delta^2} (dL_{\delta^2 t}^\varepsilon - dU_{\delta^2 t}^\varepsilon) + \frac{o(1)}{\delta^2} dt. \quad (5.6.16)$$

Note that $o(\cdot)$ is uniform for all $\omega \in \Omega$.

4. *Reduced Model*. The above system is a two-time scales system with slow variable \tilde{p}_t^ε and fast variable \tilde{q}_t^ε , which can be approximated by averaging the fast variable \tilde{q}_t^ε . Recall that, for a reflected Brownian motion σW_t inside the interval $[-l, l]$, i.e.

$$dX_t = \sigma dW_t + dL_t - dU_t,$$

the stationary measure is uniform on $[-l, l]$ and we have

$$\lim_{t \rightarrow \infty} \frac{L_t}{t} = \lim_{t \rightarrow \infty} \frac{U_t}{t} = \frac{\sigma^2}{4l}.$$

Assume that \tilde{q}_t^ε achieves its unique stationary state on the interval $[-l_t, l_t]$ with $l_t = l(\tilde{p}_t^\varepsilon)$. We obtain therefore,

$$\frac{L_{\delta^2(t+\Delta t)}^\varepsilon - L_{\delta^2 t}^\varepsilon}{\Delta t} \simeq \frac{\sigma^2 \bar{\xi}^2 (1 - \bar{\xi})^2}{4l_t} = \frac{\sigma^2}{3Al_t},$$

and

$$\frac{1}{\Delta t} \int_t^{t+\Delta t} \tilde{q}_s^\varepsilon ds \simeq \int_{-l_t}^{l_t} q^2 \frac{dq}{2l_t} = \frac{1}{3} (l_t)^2.$$

Hence I^δ can be approximated by the following single variable control problem

$$\bar{I}^\delta = \sup_{(l_t)} \liminf_{T^\delta \rightarrow \infty} \frac{1}{T^\delta} \mathbb{E}^{\mathbb{Q}_\alpha^\varepsilon} \left[\int_0^{T^\delta} (\tilde{p}_t^\varepsilon - \frac{1}{3} A^2 (l_t)^2) dt \right], \quad (5.6.17)$$

with

$$d\bar{p}_t^\varepsilon = \sigma(1 - \delta\bar{p}_t^\varepsilon)d\bar{W}_{\alpha,t}^\delta - \frac{\sigma^2}{3A\bar{l}_t^\varepsilon}\bar{p}_t^\varepsilon dt + d\bar{R}_t^\varepsilon. \quad (5.6.18)$$

5. *Approximation of Value Function.* Plugging (5.6.17) into (5.6.13) and (5.6.6), we obtain the following lower bound for the Merton problem with taxes:

$$\begin{aligned} \bar{\varphi}(x, y, k) &\simeq V_0^{C^\varepsilon, L^\varepsilon, M^\varepsilon}(x, y, k) \\ &\simeq \bar{\varphi}_\alpha(z) + \frac{1}{2}(1 - \gamma)\gamma\sigma^2\bar{c}_\alpha\bar{\varphi}_\alpha(z) \int_0^\infty e^{-\bar{c}_\alpha t} (\bar{I}^\delta A^{2/3} \varepsilon^{8/3} t) dt \\ &\simeq \bar{\varphi}_\alpha(z) \left(1 + \frac{(1 - \gamma)\gamma\sigma^2}{2\bar{c}_\alpha} \bar{I}^\delta A^{2/3} \varepsilon^{8/3}\right). \end{aligned} \quad (5.6.19)$$

5.6.3 Solution of the optimal control problem (5.6.17)-(5.6.18)

In this section, we provide a verification theorem for the problem of time-average control

$$I^\delta = \sup_{(l_t)} \liminf_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left[\int_0^T (p_t - \frac{1}{3} A^2 l_t^2) dt \right], \quad (5.6.20)$$

with

$$dp_t = \sigma(1 - \delta p_t)dW_t - \frac{\sigma^2}{3A l_t} p_t dt + dR_t, \quad (5.6.21)$$

where (l_t) is an adapted *positive* process and (R_t) a non-decreasing process restricting p_t inside $[0, 1/\delta)$. More precisely, we have

$$\int_0^\cdot \mathbb{1}_{\{p_t \neq 0\}} dR_t = 0.$$

Lemma 5.6.1 (Verification). *Assume that there exists a unique pair of $g : [0, 1/\delta) \rightarrow \mathbb{R}_+$ and $m_*^\delta > 0$ such that*

$$(1 - \delta p)^2 g''(p) = (p g'(p))^{2/3} - p + m_*^\delta,$$

with $g(0) = 0$, $g'(0) = 0$ and that the limit $g(\frac{1}{\delta}^-)$ exists. Then, the optimal solution of (5.6.20)-(5.6.21) is given by

$$I^\delta = m_*^\delta,$$

and the corresponding optimal strategy is given by

$$dp_t^* = \sigma(1 - \delta p_t^*)dW_t - \frac{\sigma^2}{3A l_t^*} p_t^* dt + dR_t^*,$$

with

$$l_t^* = \frac{1}{A} (p_t^* g'(p_t^*))^{1/3}.$$

Proof. Define $G(p) = \frac{\sigma^2}{2} g(p)$, then we have

$$\frac{\sigma^2}{2} (1 - \delta p)^2 G''(p) + p - \left(\frac{\sigma^2}{2} p G'(p)\right)^{2/3} - m_*^\delta = 0.$$

Let (l_t) be any admissible strategy and (p_t) the corresponding controlled process. Applying Itô formula to $G(p_T)$, we obtain

$$G(p_T) = G(p_0) + \int_0^T G'(p_t) dp_t + \frac{1}{2} \int_0^T G''(p_t) d\langle p \rangle_t.$$

Since $G'(0) = 0$, we obtain

$$G(p_T) = G(p_0) + \int_0^T \left(\frac{\sigma^2}{2} (1 - \delta p_t)^2 G''(p_t) - p_t G'(p_t) \frac{\sigma^2}{3A l_t} \right) dt + \int_0^T \sigma (1 - \delta p_t) G'(p_t) dW_t.$$

Consequently, the averaged cost over the horizon T is given by

$$\begin{aligned} & \frac{1}{T} \int_0^T \left(p_t - \frac{1}{3} A^2 l_t^2 \right) dt \\ &= \frac{1}{T} \int_0^T \left(\frac{\sigma^2}{2} (1 - \delta p_t)^2 G''(p_t) - p_t G'(p_t) + p_t - \frac{1}{3} A^2 l_t^2 \frac{\sigma^2}{3A l_t} \right) dt \\ & \quad + \frac{1}{T} \left(\int_0^T \sigma (1 - \delta p_t) G'(p_t) dW_t - G(p_T) + G(p_0) \right) \\ &\leq \frac{1}{T} \int_0^T \left(\frac{\sigma^2}{2} (1 - \delta p_t)^2 G''(p_t) + p_t - \left(\frac{\sigma^2}{2} p_t G'(p_t) \right)^{2/3} \right) dt \\ & \quad + \frac{1}{T} \left(\int_0^T \sigma (1 - \delta p_t) G'(p_t) dW_t - G(p_T) + G(p_0) \right) \\ &= m_*^\delta + \frac{1}{T} \left(\int_0^T \sigma (1 - \delta p_t) G'(p_t) dW_t - G(p_T) + G(p_0) \right). \end{aligned}$$

Since G is bounded, the conclusion follows directly by taking the expectation. \square

Appendix 5.A Formal derivation of Asymptotic Expansion 5.4.1

We note that all the computation is heuristic and rigorous proof is still missing. Let us first rewrite $\mathcal{L}_i w_i$ as follows:

$$\begin{aligned}\mathcal{L}_i w_i &= b^2 w_{i,bb} - 2b\xi(1-\xi)w_{i,b\xi} + \xi^2(1-\xi)^2 w_{i,\xi\xi} + c_i b w_{i,b} + d_i \xi w_{i,\xi} - K_{i,1} w_i, \\ c_i &= 2\left(1 - \frac{\mu_i}{\sigma_i^2}\right) + (1-\gamma)\left(-2\xi - \xi(1-\xi)w_{i,\xi} + b w_{i,b}\right), \\ d_i &= \left[-2\gamma(\xi - \bar{\xi}_i) + (1-\gamma)\left(\xi(1-\xi)w_{i,\xi} - b w_{i,b}\right)\right](1-\xi) + \frac{2r_i\alpha[1 + (b-1)\xi]}{\sigma_i^2} + K_{i,2}, \\ K_{i,1} &= -\frac{2}{\sigma_i^2} \bar{c}_{\alpha,i} \frac{U^*(q, \kappa) - U^*(1, \kappa)}{q-1} \Big|_{q=e^{w_i(1-\xi w_{i,\xi})}} \frac{e^{w_i} - 1}{w_i} = \frac{2}{\sigma_i^2} \bar{c}_{\alpha,i} + O(|w_i| + |\xi w_{i,\xi}|), \\ K_{i,2} &= -\frac{2}{\sigma_i^2} \bar{c}_{\alpha,i} \frac{U^*(q, \kappa) - U^*(1, \kappa)}{q-1} \Big|_{q=e^{w_i(1-\xi w_{i,\xi})}} e^{w_i} = \frac{2}{\sigma_i^2} \bar{c}_{\alpha,i} + O(|w_i| + |\xi w_{i,\xi}|),\end{aligned}$$

where $O(|w_i| + |\xi w_{i,\xi}|)$ represents the same order of $|w_i| + |\xi w_{i,\xi}|$, and we have used $U_q^*(1, \kappa) = -1$ when evaluating $K_{i,1}, K_{i,2}$.

5.A.1 Expansion at solution of tax-defaulted model

Near the tip $(1, \xi_i^\pm(1))$, we use the stretched variable

$$q_i = \frac{\xi - \bar{\xi}_i}{\delta_i^2}, \quad p_i = \frac{1-b}{\delta_i}, \quad \delta_i := (A_i \varepsilon_i)^{2/3},$$

where A_i is as given in (5.4.3). Define

$$q_i^\pm(p_i) := \frac{\xi_i^\pm(b) - \bar{\xi}_i}{\delta_i^2} \Big|_{b=1-\delta_i p_i}.$$

We write the solution in the form

$$w_i = \gamma \varepsilon_i^2 \left\{ \delta_i \frac{\sigma_i^2}{2} \frac{1}{\bar{c}_{\alpha,i}} m_i + \delta_i^3 g_i^\delta(p_i) + \delta_i^5 v_i^\delta(p_i, q_i) \right\}, \quad (5.A.1)$$

where m_i is a positive constant to be determined, $g_i^\delta(0) = 0$, and $v_{i,\delta}(p_i, q_i) = 0$ for $q_i \geq q_i^+(p_i)$. Then

$$\begin{aligned}\frac{\mathcal{L}_i w_i + f_i}{\gamma \varepsilon_i^2 \delta_i} &= a_i v_{i,qq}^\delta + (1 - \delta_i p_i)^2 g_{i,pp}^\delta - m_i + p_i - A_i^2 q_i^2 + O(\delta_i), \\ \frac{\xi w_{i,\xi} + (1-b)w_{i,b}}{\gamma \varepsilon_i^2 \delta_i^3} &= \bar{\xi}_i v_{i,q}^\delta - p_i g_{i,p}^\delta + O(\delta_i),\end{aligned}$$

where $\zeta_i = 2(1 - \mu_i/\sigma_i^2) - 2\bar{\xi}_i(1-\gamma)$, and

$$\begin{aligned}&\frac{2}{\sigma_i^2} \frac{1}{\gamma \varepsilon_i^2 \delta_i} \frac{1}{1-\gamma} \sum_{j \in \mathcal{I}} q_{ij} \frac{\bar{G}_{\alpha,j}^{1-\gamma}}{\bar{G}_{\alpha,i}^{1-\gamma}} (e^{(1-\gamma)(w^{(j)} - w_i)} - 1) \\ &= \frac{1}{\theta_i} \sum_{j \in \mathcal{I}} q_{ij} \frac{\bar{G}_{\alpha,j}^{1-\gamma}}{\bar{G}_{\alpha,i}^{1-\gamma}} \left(\theta_j \frac{m_j}{\bar{c}_{\alpha,j}} - \theta_i \frac{m_i}{\bar{c}_{\alpha,i}} \right) + O(\delta_j, j \in \mathcal{I}).\end{aligned}$$

Sending $\vee_{j \in \mathcal{I}} \delta_j \searrow 0$, we see that $(v_i, g_i) := \lim_{\delta \searrow 0} (v_i^\delta, g_i^\delta)$ is the solution of the limit problem for the tip

$$\begin{cases} \max\{a_i v_{i,qq} - A_i^2 q_i^2 + (1 - \delta_i p_i)^2 g_i'' + p_i - m_i^{\delta_i}, -v_{i,q}, \bar{\xi}_i v_{i,q} - p_i g_i'\} = 0 \\ v_i(p_i, \infty) = 0, \\ g_i(0) = 0, g_i'(0) = 0 \end{cases}$$

in $p_i \in [0, 1/\delta_i]$, $q_j \in \mathbb{R}$, where $m_*^{\delta_i}$ is defined by

$$\begin{aligned} \theta_i m_*^{\delta_i} &= \theta_i m_i + \sum_{j \in \mathcal{I}} q_{ij} \frac{\bar{G}_{\alpha,j}^{1-\gamma}}{\bar{G}_{\alpha,i}^{1-\gamma}} \left(\theta_j \frac{m_j}{\bar{c}_{\alpha,j}} - \theta_i \frac{m_i}{\bar{c}_{\alpha,i}} \right) \\ &= \left(1 + p\kappa \left(\frac{\bar{K}_{\alpha,i}}{\bar{c}_{\alpha,i}} - 1 \right) \right) (\theta_i m_i) - \sum_{j \in \mathcal{I}} \frac{q_{ij}}{\bar{c}_{\alpha,j}} \frac{\bar{G}_{\alpha,j}^{1-\gamma}}{\bar{G}_{\alpha,i}^{1-\gamma}} (\theta_j m_j) \end{aligned}$$

i.e. the linear system (5.4.6).

Denote by $q_i^\pm(p_i)$ the resulting free boundary. One finds that $v_{i,q} = 0$ for $q_i \geq q_i^+(p_i)$ is the global minimum and $v_{i,q} = p_i g'_i / \bar{\xi}_i$ for $q_i \leq q_i^-(p_i)$ is the global maximum. Thus,

$$v_{i,qq}(p_i, q_i^\pm(p_i)) = 0, \quad q_i^\pm(p_i) = \pm \frac{\sqrt{\Phi_i(p_i)}}{A_i}, \quad \Phi_i(p_i) := p_i + (1 - \delta_i p_i) g''_i - m_*^{\delta_i}.$$

The variational inequality can be written as

$$\begin{aligned} v_i &\equiv 0 & \forall q_i &\geq q_i^+(p_i), \\ a_i v_{i,qq}(p_i, q_i) &= A_i^2 q_i^2 - \Phi_i(p_i) & \forall q_i &\in (q_i^-(p_i), q_i^+(p_i)), \\ \bar{\xi}_i v_{i,q}(p_i, q_i) &= p_i g'_i(p_i) & \forall q_i &\leq q_i^-(p_i). \end{aligned}$$

It then follows that when $q_i \in [q_i^-(p_i), q_i^+(p_i)]$,

$$\begin{aligned} v_{i,q}(p_i, q_i) &= \frac{A_i^2}{a_i} \left\{ \frac{q_i^3 - q_i^+(p_i)^3}{3} - q_i^+(p_i)^2 (q_i - q_i^+(p_i)) \right\}, \\ v_i(p_i, q_i) &= \frac{A_i^2}{a_i} \left\{ \frac{q_i^4 - q_i^+(p_i)^4}{12} - \frac{q_i^+(p_i)^3}{3} (q_i - q_i^+(p_i)) - \frac{q_i^+(p_i)^2}{2} (q_i - q_i^+(p_i))^2 \right\}. \end{aligned}$$

The boundary condition $\bar{\xi}_i v_{i,q} = p_i g'_i(p_i)$ at $q_i = q_i^-(p_i) = -q_i^+(p_i)$ gives

$$p_i g_{i,p}(p_i) = \frac{4\bar{\xi}_i}{3a_i} A_i^2 q_i^+(p_i)^3 = (\Phi_i(p_i))^{3/2} = (p_i + (1 - \delta_i p_i)^2 g''_i(p_i) - m_*^{\delta_i})^{3/2},$$

since $A_i = 4\bar{\xi}_i / (3a_i)$. The equation for g_i becomes

$$\begin{cases} (1 - \delta_i p_i)^2 g''_i(p_i) + p_i - m_*^{\delta_i} - (p_i g'_i(p_i))^{2/3} = 0 & \forall p_i \in [0, 1/\delta_i], \\ g_i(0) = 0, \quad g'_i(0) = 0. \end{cases} \quad (5.A.2)$$

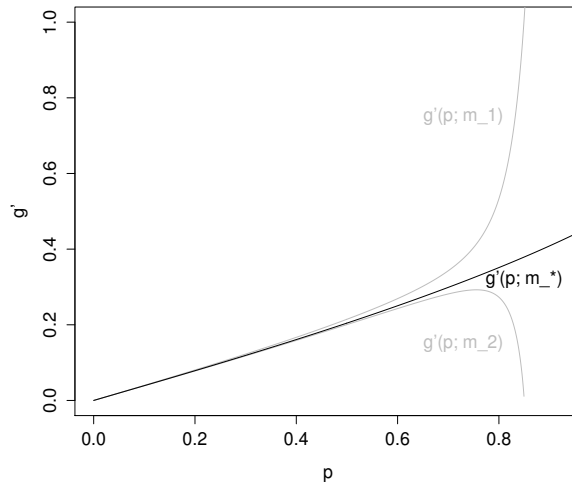
At the leading order, the free boundary is given by $q_i = \pm \sqrt{\Phi_i(p_i)} / A_i$ with

$$\Phi_i(p_i) = (1 - \delta_i p_i)^2 g''_i + p - m_*^{\delta_i} = (p_i g'_i(p_i))^{2/3} = p_i \left(\frac{g'_i(p_i)}{\sqrt{p_i}} \right)^{2/3}.$$

In the original variable, this gives the approximation

$$\begin{aligned} \xi_i^\pm &\approx \bar{\xi}_i \pm \frac{\delta_i^2 \sqrt{\Phi_i(p_i)}}{A_i} = \bar{\xi}_i \pm \frac{\delta_i^2 \sqrt{p_i}}{A_i} \left(\frac{g'_i(p_i)}{\sqrt{p_i}} \right)^{1/3} \\ &= \bar{\xi}_i \pm \varepsilon_i \sqrt{1-b} \left(\frac{g'_i(p_i)}{\sqrt{p_i}} \right)^{1/3} \Big|_{p_i = \frac{1-b}{(A_i \varepsilon_i)^{2/3}}}. \end{aligned}$$

Using a shooting method with $m_*^{\delta_i}$, $i \in \mathcal{I}$ being the shooting parameter, one can numerically show that, at least for the cases $|\mathcal{I}| \leq 2$, there exists unique values for $m_*^{\delta_i}$, $i \in \mathcal{I}$ such that $g_i(p_i)$ is defined on $[0, 1/\delta_i]$ for all $i \in \mathcal{I}$.



Left panel: $\delta = 1, m_1 = 0.40, m_2 = 0.39, m_* = 0.392799884305488372$.

Figure 5.14 – Numerical evidence for the existence of unique solution m_*^δ for (5.A.2)

5.A.2 An approximation to the optimal consumption strategy

By (5.3.14), (5.3.17), and (5.3.18), we can express the optimal consumption as

$$\begin{aligned} \frac{C^*}{x + (1 - \alpha)y + \alpha k} &= \bar{c}_{\alpha,i} e^{(\kappa-1)w_i/\kappa} (1 - \xi w_{i,\xi})^{-1/\kappa} \\ &= \bar{c}_{\alpha,i} + \frac{\kappa - 1}{\kappa} \frac{\sigma_i^2}{2} \gamma m_i A_i^{2/3} \varepsilon_i^{8/3} + o(\varepsilon_i^{8/3}) \end{aligned} \quad (5.A.3)$$

by noting that $\xi w_{i,\xi} = O(\varepsilon_i^4)$ (see (5.A.1)).

Chapter 6

Joint impact of capital gains taxes and transaction costs

Contents

6.1	Main result	196
6.1.1	HJB equation	196
6.1.2	Benchmark model	196
6.1.3	Asymptotic expansion	197
6.2	Probabilistic interpretation	198
6.3	Homothetic case	200
6.4	Asymptotic development	201
6.4.1	Fast variables	201
6.4.2	Corrector equations	203

We present in this chapter a homogenization result for portfolio selection under proportional transaction costs and linear capital gains taxes, which generalizes both [ST13] and [CD13]. The same problem has been considered under different settings in [Lel99, BCP05, CP99].

For simplicity, we consider the model in [BST07]. However, we note that the same result can be obtained for more general models and cost structures (e.g. stochastic volatility models with fixed and proportional transaction costs), as long as the structure of linear tax rate is maintained. The dynamics of risky asset is given by the Black-Scholes model

$$dS_t = S_t(\mu dt + \sigma dW_t),$$

and we use the following notations:

- $\lambda_B, \lambda_S \in [0, 1)$ the proportional transaction costs coefficients.
- $\alpha \in [0, 1)$ the tax rate.
- X_t the wealth in non-risky asset.
- Y_t the wealth in risky asset.
- K_t the weighted cumulated investment in risky asset.
- C_t the consumption process.
- dL_t the amount invested in risky asset at time t .
- dM_t the proportion of risky asset sold at time t .

The dynamics of portfolio for the strategy (C_t, L_t, M_t) is given by

$$dX_t = ((1 - \alpha)rX_t - C_t)dt - (1 + \lambda_B)dL_t + (1 - \lambda_S)[(1 - \alpha)Y_t - \alpha K_t]dM_t,$$

$$\begin{aligned} dY_t &= Y_t(\mu dt + \sigma dW_t) + dL_t - Y_t - dM_t, \\ dK_t &= dL_t - K_t - dM_t. \end{aligned}$$

Fixing any general utility function U , the problem of optimal consumption and investment under both transaction costs and capital gains taxes is defined by

$$\varphi(x, y, k) = \sup_{(C, L, M) \in \mathcal{A}} \mathbb{E} \left[\int_0^\infty e^{-\beta t} U(C_t) dt \right], \quad (6.0.1)$$

where \mathcal{A} is the set of all admissible strategies.

6.1 Main result

6.1.1 HJB equation

The HJB equation for φ is formally given as

$$\min \left\{ - \left(\frac{1}{2} \sigma^2 y^2 \partial_{yy}^2 \varphi + \mu y \partial_y \varphi + (1 - \alpha) r x \partial_x \varphi - \beta \varphi + \tilde{U}(\partial_x \varphi) \right), \quad (6.1.1) \right.$$

$$\left. (1 + \lambda_B) \partial_x \varphi - \partial_y \varphi - \partial_k \varphi, \quad (6.1.2) \right.$$

$$\left. - (1 - \lambda_S) [(1 - \alpha)y + \alpha k] \partial_x \varphi - y \partial_y \varphi - k \partial_k \varphi, \quad (y, k) \neq (0, 0), \quad (6.1.3) \right.$$

$$\left. - (1 - \lambda_S) \partial_x \varphi + \frac{1}{1 - \alpha} \partial_y \varphi, \quad y = k = 0, \quad (6.1.4) \right.$$

$$\left. \inf_{y_+} \left(\varphi(x, y, k) - \varphi(x + (1 - \alpha)y + \alpha k - \lambda_S y - \lambda_B y_+, y_+, y_+) \right) \right\} = 0. \quad (6.1.5)$$

Comparing to [BST07], we have added (6.1.5) which corresponds to the liquidation of the portfolio. The condition (6.1.2) is the buy boundary condition and (6.1.3)-(6.1.4) are the sell boundary conditions.

We consider an asymptotic development for small transaction costs λ_B , λ_S and small interest rate r . More precisely, we assume that

$$\lambda_B^\delta = \lambda_B \delta^6, \quad \lambda_S^\delta = \lambda_S \delta^6, \quad r^\delta = r \delta^3, \quad \delta > 0. \quad (6.1.6)$$

Then the above HJB equation becomes¹

$$\min \left\{ - \left(\frac{1}{2} \sigma^2 y^2 \partial_{yy}^2 \varphi^\delta + \mu y \partial_y \varphi^\delta + (1 - \alpha) r \delta^3 x \partial_x \varphi^\delta - \beta \varphi^\delta + \tilde{U}(\partial_x \varphi^\delta) \right), \quad (6.1.1^\delta) \right.$$

$$\left. (1 + \lambda_B \delta^6) \partial_x \varphi^\delta - \partial_y \varphi^\delta - \partial_k \varphi^\delta, \quad (6.1.2^\delta) \right.$$

$$\left. - (1 - \lambda_S \delta^6) [(1 - \alpha)y + \alpha k] \partial_x \varphi^\delta + y \partial_y \varphi^\delta + k \partial_k \varphi^\delta, \quad (y, k) \neq (0, 0), \quad (6.1.3^\delta) \right.$$

$$\left. \inf_{y_+} \left(\varphi^\delta(x, y, k) - \varphi^\delta(x + (1 - \alpha)y + \alpha k - \lambda_S \delta^6 y - \lambda_B \delta^6 y_+, y_+, y_+) \right) \right\} = 0. \quad (6.1.5^\delta)$$

The asymptotic development can be deduced heuristically from probabilistic arguments or directly from (6.1.2^δ).

6.1.2 Benchmark model

Without transaction costs and capital gains taxes, the solution of the optimal consumption and investment problem is given by

$$\frac{1}{2} \sigma^2 \bar{y}^2 \partial_{zz}^2 \bar{\varphi}(z) + \bar{y}(\mu - r) \partial_z \bar{\varphi}(z) + r z \partial_z \bar{\varphi}(z) - \beta \bar{\varphi}(z) + \tilde{U}(\partial_z \bar{\varphi}(z)) = 0,$$

1. Since we will essentially look at the zone around Merton line, the boundary condition (6.1.4) is omitted.

where the optimal investment $\bar{y}(z)$ verifies

$$\sigma^2 \bar{y} \partial_{zz}^2 \bar{\varphi}(z) = -(\mu - r) \partial_z \bar{\varphi}(z).$$

The optimal consumption rate is given by

$$\bar{C}(z) = -\tilde{U}'(\partial_z \bar{\varphi}(z)) = (U')^{-1}(\partial_z \bar{\varphi}(z)).$$

By similar probabilistic argument as in Section 5.6, the right benchmark turns out to be the following “tax-deflated” model

$$\frac{1}{2} \sigma_\alpha^2 (\bar{y}_\alpha^\delta)^2 \partial_{zz}^2 \bar{\varphi}_\alpha^\delta(z) + \bar{y}_\alpha^\delta (\mu_\alpha - r_\alpha^\delta) \partial_z \bar{\varphi}_\alpha^\delta(z) + r_\alpha^\delta z \partial_z \bar{\varphi}_\alpha^\delta(z) - \beta \bar{\varphi}_\alpha^\delta(z) + \tilde{U}(\partial_z \bar{\varphi}_\alpha^\delta(z)) = 0, \quad (6.1.7)$$

where

$$\mu_\alpha = (1 - \alpha)\mu, \quad r_\alpha^\delta = (1 - \alpha)r^\delta, \quad \sigma_\alpha = (1 - \alpha)\sigma. \quad (6.1.8)$$

Here we use $\bar{\cdot}_\alpha^\delta$ to indicate the optimal quantities under the frictionless tax-deflated model. In contrary to asymptotic development in the presence of only transaction costs, here the benchmark depends also on the small parameter δ . This helps eliminating unrelated terms from the underlying control problem. Denote the corresponding optimal solution by \bar{y}_α^δ and \bar{C}_α^δ such that

$$\sigma_\alpha^2 \bar{y}_\alpha^\delta \partial_{zz}^2 \bar{\varphi}_\alpha^\delta(z) = -(\mu_\alpha - r_\alpha^\delta) \partial_z \bar{\varphi}_\alpha^\delta(z), \quad (6.1.9)$$

and

$$\bar{C}_\alpha^\delta(z) = -\tilde{U}'(\partial_z \bar{\varphi}_\alpha^\delta(z)) = (U')^{-1}(\partial_z \bar{\varphi}_\alpha^\delta(z)). \quad (6.1.10)$$

6.1.3 Asymptotic expansion

The main result of this chapter is the following.

Asymptotic Expansion 6.1.1. ² Replacing the transaction cost coefficients λ_B , λ_S and the interest rate r by

$$\lambda_B^\delta = \lambda_B \delta^6, \quad \lambda_S^\delta = \lambda_S \delta^6, \quad r^\delta = r \delta^3, \quad \delta > 0,$$

and let φ^δ be the corresponding value function defined by (6.0.1). Then we have

$$\varphi^\delta = \bar{\varphi}_\alpha^\delta(z) + u_\alpha^\delta(z) \delta^4 + o(\delta^6),$$

where u_α^δ is given by

$$-\frac{1}{2} \sigma_\alpha^2 (\bar{y}_\alpha^\delta(z))^2 \partial_{zz}^2 u_\alpha^\delta(z) - [(\mu_\alpha - r_\alpha^\delta) \bar{y}_\alpha^\delta(z) + r_\alpha^\delta z - \bar{C}_\alpha^\delta(z)] \partial_z u_\alpha^\delta(z) + \beta u_\alpha^\delta(z) - a(z) = 0. \quad (6.1.11)$$

The constant $a(z)$ (depending on z) is uniquely determined by

$$\begin{aligned} \min \left\{ -\frac{1}{2} \sigma^2 (1 - \delta p)^2 \partial_{pp}^2 G_\alpha^\delta(z, p) - (\alpha r_\alpha \partial_z \bar{\varphi}_\alpha^\delta(z) \bar{y}_\alpha^\delta(z) p - b(z, p)) + a(z), \right. \\ \left. G_\alpha^\delta(z, p) + (\lambda_S + \lambda_B) \partial_z \bar{\varphi}_\alpha^\delta(z) \bar{y}_\alpha^\delta(z) - G_\alpha^\delta(z, 0) \right\} = 0, \end{aligned} \quad (6.1.12)$$

and the constant $b(z, p)$ (depending on p and z) is uniquely determined by

$$\begin{aligned} \min \left\{ \frac{1}{2} \sigma^2 (\bar{y}_\alpha^\delta(z))^2 [1 - (1 - \alpha) \partial_z \bar{y}_\alpha^\delta(z)]^2 \partial_{qq}^2 w_\alpha^\delta(z, p, q) + \frac{1}{2} \sigma_\alpha^2 (-\partial_{zz}^2 \bar{\varphi}_\alpha^\delta(z)) q^2 - b(z, p), \right. \\ \left. \frac{p}{\bar{y}_\alpha^\delta(z)} \partial_p G_\alpha^\delta(z, p) + \lambda_B \partial_z \bar{\varphi}_\alpha^\delta(z) - \partial_q w_\alpha^\delta(z, p, q), \right. \\ \left. - \lambda_S \partial_z \bar{\varphi}_\alpha^\delta(z) - \partial_q w_\alpha^\delta(z, p, q) \right\} = 0. \end{aligned} \quad (6.1.13)$$

2. We provide only formal derivation of the corrector equations.

6.2 Probabilistic interpretation

It is convenient to consider the certainty equivalent wealth gain, so we postulate

$$\begin{aligned}\varphi^\delta &= \bar{\varphi}_\alpha^\delta(z + \tilde{u}_\alpha^\delta(z)\delta^4 + \tilde{G}_\alpha^\delta(z, p)\delta^6 - \tilde{w}_\alpha^\delta(z, p, q)\delta^8 + o(\delta^8)) \\ &= \bar{\varphi}_\alpha^\delta(z) + \partial_z \bar{\varphi}_\alpha^\delta \tilde{u}_\alpha^\delta(z)\delta^4 + \partial_z \bar{\varphi}_\alpha^\delta(z) \tilde{G}_\alpha^\delta(z, p)\delta^6 - \partial_z \bar{\varphi}_\alpha^\delta(z) \tilde{w}_\alpha^\delta(z, p, q)\delta^8 + O((\tilde{u}_\alpha^\delta)^2 \delta^8) + o(\delta^8)\end{aligned}$$

It is straightforward to see that

$$\begin{aligned}\min \left\{ -\frac{1}{2}\sigma^2(1-\delta p)^2 \partial_{pp}^2 \tilde{G}_\alpha^\delta(z, p) - \alpha r_\alpha \bar{y}_\alpha^\delta(z)p + \tilde{a}(z) + \tilde{b}(z, p), \right. \\ \left. \tilde{G}_\alpha^\delta(z, p) + (\lambda_S + \lambda_B) \bar{y}_\alpha^\delta(z) - \tilde{G}_\alpha^\delta(z, 0) \right\} = 0, \quad (6.2.1)\end{aligned}$$

$$\begin{aligned}\min \left\{ \frac{1}{2}\sigma^2(\bar{y}_\alpha^\delta)^2 [1 - (1-\alpha)\partial_z \bar{y}_\alpha^\delta(z)]^2 \partial_{qq}^2 \tilde{w}_\alpha^\delta(z, p, q) + \frac{1}{2} \frac{\sigma_\alpha^2}{\eta_\alpha^\delta(z)} q^2 - \tilde{b}(z, p), \right. \\ \left. \frac{p}{\bar{y}_\alpha^\delta(z)} \partial_p \tilde{G}_\alpha^\delta(z, p) + \lambda_B - \partial_q \tilde{w}_\alpha^\delta(z, p, q), \right. \\ \left. -\lambda_S - \partial_q \tilde{w}_\alpha^\delta(z, p, q) \right\} = 0, \quad (6.2.2)\end{aligned}$$

with

$$\tilde{a}(z) \partial_z \bar{\varphi}_\alpha^\delta(z) = a(z), \quad \tilde{b}(z, p) \partial_z \bar{\varphi}_\alpha^\delta(z) = b(z, p).$$

First corrector equation The dynamic of the optimal wealth process $\bar{Z}_{\alpha,t}^\delta$ satisfies

$$d\bar{Z}_{\alpha,t}^\delta = (r_\alpha^\delta \bar{Z}_{\alpha,t}^\delta - \bar{C}_\alpha^\delta(\bar{Z}_{\alpha,t}^\delta))dt + \bar{y}_\alpha^\delta(\bar{Z}_{\alpha,t}^\delta)((\mu_\alpha - r_\alpha^\delta)dt + \sigma_\alpha dW_t), \quad \bar{Z}_{\alpha,0}^\delta = z.$$

Denoting the generator of $\bar{Z}_{\alpha,t}^\delta$ by $\mathcal{L}_\alpha^\delta$, then (6.1.11) can be written as

$$-\mathcal{L}_\alpha^\delta u_\alpha^\delta(z) + \beta u_\alpha^\delta(z) = \partial_z \bar{\varphi}_\alpha^\delta(z) \tilde{a}(z).$$

Therefore, we obtain the following probabilistic representation of $\tilde{u}_\alpha^\delta(z)$ (see [ST13, Equation (1.2)]):

$$u_\alpha^\delta(z) = \mathbb{E} \left[\int_0^\infty e^{-\beta t} \partial_z \bar{\varphi}_\alpha^\delta(\bar{Z}_{\alpha,t}^\delta) \tilde{a}(\bar{Z}_{\alpha,t}^\delta) dt \right].$$

Define the risk-neutral measure \mathbb{Q}_α^δ under the frictionless tax-deflated model by

$$\left. \frac{d\mathbb{Q}_\alpha^\delta}{d\mathbb{P}} \right|_{\mathcal{F}_t} = \frac{e^{-(\beta-r_\alpha^\delta)t} \partial_z \bar{\varphi}_\alpha^\delta(\bar{Z}_{\alpha,t}^\delta)}{\partial_z \bar{\varphi}_\alpha^\delta(z)} = \mathcal{E} \left(- \int_0^t \frac{\bar{y}_\alpha^\delta(\bar{Z}_{\alpha,s}^\delta)}{\eta_\alpha^\delta(\bar{Z}_{\alpha,s}^\delta)} \sigma_\alpha dW_s \right) = \mathcal{E} \left(- \int_0^t (\mu_\alpha - r_\alpha^\delta) \sigma_\alpha^{-1} dW_s \right),$$

where we have used (6.1.7) and (6.1.9) for the second and third equality. We obtain therefore

$$u_\alpha^\delta(z) = \partial_z \bar{\varphi}_\alpha^\delta(z) \mathbb{E}^{\mathbb{Q}_\alpha^\delta} \left[\int_0^\infty e^{-r_\alpha^\delta t} \tilde{a}(\bar{Z}_{\alpha,t}^\delta) dt \right],$$

and

$$d\bar{Z}_{\alpha,t}^\delta = (r_\alpha^\delta \bar{Z}_{\alpha,t}^\delta - \bar{C}_\alpha^\delta(\bar{Z}_{\alpha,t}^\delta))dt + \bar{y}_\alpha^\delta(\bar{Z}_{\alpha,t}^\delta) \sigma_\alpha dW_t^{\mathbb{Q}_\alpha^\delta},$$

where $W_t^{\mathbb{Q}_\alpha^\delta}$ is a \mathbb{Q}_α^δ -Brownian motion by Girsanov theorem. In particular, the equivalent wealth gain with respect to the frictionless tax-deflated model is given by

$$\tilde{u}_\alpha^\delta(z) = \mathbb{E}^{\mathbb{Q}_\alpha^\delta} \left[\int_0^\infty e^{-r_\alpha^\delta t} \tilde{a}(\bar{Z}_{\alpha,t}^\delta) dt \right],$$

see also [HMK15, Theorem 4.1] and Remark 4.4.5.

Second corrector equation Now following [ST13, Equation (4.4), Section 4.1], we solve explicitly (6.2.2) and obtain

$$\tilde{b}(z, p) = \frac{1}{2} \frac{\sigma_\alpha^2}{\eta_\alpha^\delta(z)} \left[\frac{3 \sigma^2 (\bar{y}_\alpha^\delta(z))^2 [1 - (1 - \alpha) \partial_z \bar{y}_\alpha^\delta]^2}{\sigma_\alpha^2 / \eta_\alpha^\delta(z)} \left(\lambda_S + \lambda_B + \frac{1}{\bar{y}_\alpha^\delta} p \partial_p \tilde{G}_\alpha^\delta(z, p) \right) \right]^{2/3}$$

Note that we can also write

$$\begin{aligned} \tilde{b}(z, p) &= \inf_{l>0} \left(\lambda_S + \lambda_B + \frac{1}{\bar{y}_\alpha^\delta} p \partial_p \tilde{G}_\alpha^\delta(z, p) \right) \frac{\sigma^2 (\bar{y}_\alpha^\delta(z))^2 [1 - (1 - \alpha) \partial_z \bar{y}_\alpha^\delta]^2}{4l} + \frac{\sigma_\alpha^2}{2\eta_\alpha^\delta(z)} \frac{l^2}{3} \\ &= \inf_{l>0} \left(\lambda_S + \lambda_B + \frac{1}{\bar{y}_\alpha^\delta} p \partial_p \tilde{G}_\alpha^\delta(z, p) \right) \frac{a^{\bar{y}_\alpha^\delta}(z)}{4l} + r(z) \frac{l^2}{3}, \end{aligned}$$

where

$$a^{\bar{y}_\alpha^\delta}(z) = \sigma^2 (\bar{y}_\alpha^\delta(z))^2 [1 - (1 - \alpha) \partial_z \bar{y}_\alpha^\delta]^2, \quad r(z) = \frac{\sigma_\alpha^2}{2\eta_\alpha^\delta(z)}.$$

Plugging above expression into (6.2.1) and using similar arguments as in Section 5.6.3, it follows that $\tilde{a}(z)$ can be represented by the following time-average control problem :

$$\tilde{a}(z) = \sup_{(l_t)} \lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left[\int_0^T \left(\alpha r_\alpha \bar{y}_\alpha^\delta(z) p_t - \frac{a^{\bar{y}_\alpha^\delta}(z) (\lambda_S + \lambda_B)}{4l_t} - r(z) \frac{l_t^2}{3} \right) dt - \sum_{\tau_j \leq T} (\lambda_S + \lambda_B) \bar{y}_\alpha^\delta(z) \right], \quad (6.2.3)$$

where

$$dp_t = \sigma(1 - \delta p_t) dW_t - \frac{a^{\bar{y}_\alpha^\delta}(z)}{4\bar{y}_\alpha^\delta l_t} p_t dt - \sum_{\tau_j \leq t} p_{\tau_j-}.$$

The four terms in (6.2.3) are interpreted sequentially as i) certainty equivalent wealth profit of the tax deferral value with respect to the frictionless tax-deflated model, ii) proportional transaction costs incurred when restricting $y_t - \bar{y}_\alpha^\delta(\bar{Z}_{\alpha,t}^\delta)$ inside $[-l_t, l_t]$, iii) certainty equivalent wealth loss due to deviation from the benchmark position $\bar{y}_\alpha^\delta(\bar{Z}_{\alpha,t}^\delta)$, iv) transaction costs incurred by wash-sale (realization of capital losses).

Third corrector equation The interpretation of (6.2.2) is already given in [ST13, Remark 3.3], that is

$$-\tilde{b}(z, p) = \inf_{(L,U)} \overline{\lim}_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left[\int_0^T r(z) q^2 + h^-(z, p) L_T + h^+(z, p) U_T \right],$$

where

$$dq_t = \sqrt{a^{\bar{y}_\alpha^\delta}(z)} dW_t + L_t - U_t,$$

and

$$h^-(z, p) = \lambda_B + p \frac{1}{\bar{y}_\alpha^\delta(z)} \partial_p \tilde{G}_\alpha^\delta(z, p), \quad h^+(z, p) = \lambda_S.$$

This coincides with the heuristics in Section 4.4.3. Indeed, using the following change of variable

$$\tilde{q} = \frac{q}{S},$$

we obtain

$$-\tilde{b}(z, p) = \inf_{(\tilde{L}, \tilde{U})} \overline{\lim}_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left[\int_0^T \tilde{r}(z, S) \tilde{q}^2 + \tilde{h}^-(z, p, S) \tilde{L}_T + \tilde{h}^+(z, p, S) \tilde{U}_T \right],$$

where

$$d\tilde{q}_t = \sqrt{\tilde{a}^{\bar{y}_\alpha^\delta}(z)} dW_t + \tilde{L}_t - \tilde{U}_t,$$

with

$$\tilde{r}(z, S) = \frac{\sigma_\alpha^2 S^2}{2\eta_\alpha^\delta(z)}, \quad \tilde{a}^{\bar{y}_\alpha^\delta}(z, S) = \sigma^2 (\bar{y}_\alpha^\delta(z))^2 [1 - (1 - \alpha) \partial_z \bar{y}_\alpha^\delta]^2 / S^2,$$

and

$$\tilde{h}^-(z, p, S) = \lambda_B S + p \frac{S}{\bar{y}_\alpha^\delta} \partial_p \tilde{G}_\alpha^\delta(z, p), \quad \tilde{h}^+(z, p, S) = \lambda_S S.$$

We verify easily that $r(\bar{Z}_{\alpha,t}^\delta, S_t)$, $h^\pm(\bar{Z}_{\alpha,t}^\delta, p_t, S_t)$ and $\tilde{a}^{\bar{y}_\alpha^\delta}(\bar{Z}_{\alpha,t}^\delta, S_t)$ correspond to r_t, h_t and a_t in 4.4.3, with the target φ_t^* being $\bar{y}_\alpha^\delta(\bar{Z}_{\alpha,t}^\delta)/S_t$.

Remark 6.2.1. *From the corrector equation (6.2.2), we observe that the impact of capital gains taxes on the non-transaction zone is asymptotically equivalent to adding an extra proportional cost*

$$\frac{p}{\bar{y}_\alpha^\delta(z)} \partial_p \tilde{G}_\alpha^\delta(z, p),$$

to the cost coefficient λ_B of the buy side.

Remark 6.2.2. *Based on the system of corrector equations, one can easily derive the trading boundaries in the presence of both capital gains taxes and transaction costs. The non-transaction zone is similar to [Lel99, Figure 3]. However, the corrector equations contain much more quantitative information for economic analysis. It would be interesting to compare numerical computation with our asymptotic analysis.*

6.3 Homothetic case

We illustrate our result in the case of classical CRRA utility function

$$U(C) = \frac{C^{1-\gamma}}{1-\gamma}, \quad C > 0,$$

for $\gamma > 0$ and $\gamma \neq 1$. We have

$$\bar{\varphi}(z) = \frac{\bar{K}^{-\gamma}}{1-\gamma} z^{1-\gamma}, \quad \bar{y}(z) = \bar{\xi} z, \quad \bar{C}(z) = \bar{K} z,$$

where

$$\bar{K} = \frac{\beta}{\gamma} - \frac{1-\gamma}{\gamma} \left[r + \frac{(\mu-r)^2}{2\gamma\sigma^2} \right], \quad \bar{\xi} = \frac{\mu-r}{\gamma\sigma^2}.$$

Moreover, we have

$$\eta(z) := -\frac{\partial_z \bar{\varphi}(z)}{\partial_{zz}^2 \bar{\varphi}(z)} = \frac{z}{\gamma}.$$

Without capital gains taxes ($\alpha = 0$) In the absence of capital gains taxes, \tilde{G}_α^δ is just a constant and we have

$$\tilde{a}(z) \equiv \tilde{b}(z, p).$$

The corrector equations (6.2.1)-(6.2.2) reduce to (3.12) in [ST13].

Without transaction costs ($\lambda_B = \lambda_S = 0$) Define ρ by

$$\tilde{b}(z, p) = \left[\rho^3 + \frac{\sigma^2}{2} \frac{1}{\sqrt{2}A(\gamma\sigma^2 z)^{-1/2}} p \partial_p \tilde{G}_\alpha^\delta(z, p) \right]^{2/3}, \quad A = \frac{4\bar{\xi}^\delta}{3(\bar{\xi}^\delta)^2(1-\bar{\xi}^\delta)^2}.$$

Then \tilde{G}_α^δ solves

$$-\frac{1}{2}\sigma^2(1-\delta p)^2\partial_{pp}^2\tilde{G}_\alpha^\delta(z,p) - \alpha r_\alpha \bar{y}_\alpha^\delta(z)p + \tilde{a}(z) + \left[\rho^3 + \frac{\sigma^2}{2} \frac{1}{\sqrt{2A}(\gamma\sigma^2z)^{-1/2}} p \partial_p \tilde{G}_\alpha^\delta(z,p)\right]^{2/3} = 0.$$

By a change of variables, we can show that

$$\tilde{a}(z) = \left(\sqrt{2A}(\gamma\sigma^2z)^{-1/2}\right)^{2/3} \left(\alpha r_\alpha \bar{y}_\alpha^\delta(z)\right)^{1/3} m^\delta(\tilde{\rho}), \quad \tilde{\rho} = \frac{\rho}{\sqrt{\alpha r_\alpha \bar{y}_\alpha^\delta}},$$

where $m^\delta(\rho)$ solves

$$-\frac{1}{2}(1-\delta p)^2 G''(p) - p + m^\delta(\rho) + (\rho^3 + \frac{1}{2}pG'(p))^{2/3} = 0.$$

For $\lambda_S = \lambda_B = 0$, we have $\rho = 0$ and consequently

$$\partial_z \bar{\varphi}_\alpha^\delta(z) \tilde{u}_\alpha^\delta = \frac{1}{\bar{K}_\alpha^\delta} \frac{(1-\gamma)\bar{\varphi}_\alpha^\delta(z)}{z} \tilde{a}(z) = \bar{\varphi}_\alpha^\delta(z) \frac{1-\gamma}{\bar{K}_\alpha^\delta} \frac{\gamma\sigma^2}{2} \left(\frac{2\alpha r_\alpha \bar{y}_\alpha^\delta}{\gamma\sigma^2}\right)^{4/3} m^\delta(0) A^{2/3},$$

which is exactly the result in Chapter 5 in a single regime market (see also [CD13, Equation (A.13)]).

6.4 Asymptotic development

6.4.1 Fast variables

We postulate

$$\varphi^\delta(x, y, k) = \bar{\varphi}_\alpha^\delta(z) + \delta^4 u_\alpha^\delta(z) + \delta^6 G_\alpha^\delta(z, p) - \delta^8 w_\alpha^\delta(z, p, q) + o(\delta^8), \quad (6.4.1)$$

where

$$z = x + (1-\alpha)y + \alpha k, \quad (6.4.2)$$

$$p = \frac{1-\frac{k}{y}}{\delta}, \quad (6.4.3)$$

$$q = \frac{y - \bar{y}_\alpha^\delta(z)}{\delta^2}. \quad (6.4.4)$$

The Jacobian of the above change of variable is

$$J = \begin{pmatrix} 1 & 1-\alpha & \alpha \\ 0 & \frac{1-\delta p}{\delta(\bar{y}_\alpha^\delta(z) + \delta^2 q)} & \frac{-1}{\delta(\bar{y}_\alpha^\delta(z) + \delta^2 q)} \\ \frac{-\partial_z \bar{y}_\alpha^\delta(z)}{\delta^2} & \frac{1-(1-\alpha)\partial_z \bar{y}_\alpha^\delta(z)}{\delta^2} & \frac{-\alpha \partial_z \bar{y}_\alpha^\delta(z)}{\delta^2} \end{pmatrix}.$$

By chain rule, we have

$$\begin{aligned} \partial_x \varphi^\delta &= \partial_z \varphi^\delta + \frac{-\partial_z \bar{y}_\alpha^\delta}{\delta^2} \partial_q \varphi^\delta, \\ \partial_y \varphi^\delta &= (1-\alpha)\partial_z \varphi^\delta + \frac{1-\delta p}{\delta(\bar{y}_\alpha^\delta(z) + \delta^2 q)} \partial_p \varphi^\delta + \frac{1-(1-\alpha)\partial_z \bar{y}_\alpha^\delta(z)}{\delta^2} \partial_q \varphi^\delta, \\ \partial_k \varphi^\delta &= \alpha \partial_z \varphi^\delta + \frac{-1}{\delta(\bar{y}_\alpha^\delta(z) + \delta^2 q)} \partial_p \varphi^\delta + \frac{-\alpha \partial_z \bar{y}_\alpha^\delta(z)}{\delta^2} \partial_q \varphi^\delta. \end{aligned}$$

In particular, we have

$$\partial_x \varphi^\delta - \partial_y \varphi^\delta - \partial_k \varphi^\delta = \frac{p}{(\bar{y}_\alpha^\delta(z) + \delta^2 q)} \partial_p \varphi^\delta + \frac{1}{\delta^2} \partial_q \varphi^\delta.$$

and

$$-\alpha \partial_y \varphi^\delta + (1 - \alpha) \partial_k \varphi^\delta = \frac{-1 + \alpha p \delta}{\delta(\bar{y}_\alpha^\delta(z) + \delta^2 q)} \partial_p \varphi^\delta - \frac{\alpha}{\delta^2} \partial_q \varphi^\delta.$$

Plugging (6.4.1) into above equations, we obtain

$$\begin{aligned} \partial_x \varphi^\delta &= \left(\partial_z \bar{\varphi}_\alpha^\delta(z) + \delta^4 \partial_z u_\alpha^\delta(z) + \delta^6 \partial_z G_\alpha^\delta(z, p) + o(\delta^6) \right) \\ &\quad + \frac{-\partial_z \bar{y}_\alpha^\delta}{\delta^2} \left(-\delta^8 \partial_q w_\alpha^\delta(z, p, q) + o(\delta^8) \right) \\ &= \partial_z \bar{\varphi}_\alpha^\delta(z) + \delta^4 \partial_z u_\alpha^\delta(z) + \delta^6 \left(\partial_z G_\alpha^\delta(z, p) + \partial_z \bar{y}_\alpha^\delta \partial_q w_\alpha^\delta(z, p, q) \right) + o(\delta^6), \end{aligned}$$

$$\begin{aligned} \partial_y \varphi^\delta &= (1 - \alpha) \left(\partial_z \bar{\varphi}_\alpha^\delta(z) + \delta^4 \partial_z u_\alpha^\delta(z) + \delta^6 \partial_z G_\alpha^\delta(z, p) + o(\delta^6) \right) \\ &\quad + \frac{1 - \delta p}{\delta(\bar{y}_\alpha^\delta(z) + \delta^2 q)} \left(\delta^6 \partial_p G_\alpha^\delta(z, p) + o(\delta^8) \right) \\ &\quad + \frac{1 - (1 - \alpha) \partial_z \bar{y}_\alpha^\delta(z)}{\delta^2} \left(-\delta^8 \partial_q w_\alpha^\delta(z, p, q) + o(\delta^7) \right) \\ &= (1 - \alpha) \partial_z \bar{\varphi}_\alpha^\delta(z) + \delta^4 (1 - \alpha) \partial_z u_\alpha^\delta(z) + \delta^5 \frac{1}{\bar{y}_\alpha^\delta(z) + \delta^2 q} \partial_p G_\alpha^\delta(z, p) \\ &\quad + \delta^6 \left((1 - \alpha) \partial_z G_\alpha^\delta(z, p) - \frac{p}{\bar{y}_\alpha^\delta(z) + \delta^2 q} \partial_p G_\alpha^\delta(z, p) - [1 - (1 - \alpha) \partial_z \bar{y}_\alpha^\delta(z)] \partial_q w_\alpha^\delta(z, p, q) \right) \\ &\quad + o(\delta^6), \end{aligned}$$

$$\begin{aligned} \partial_k \varphi^\delta &= \alpha \left(\partial_z \bar{\varphi}_\alpha^\delta(z) + \delta^4 \partial_z u_\alpha^\delta(z) + \delta^6 \partial_z G_\alpha^\delta(z, p) + o(\delta^6) \right) \\ &\quad + \frac{-1}{\delta(\bar{y}_\alpha^\delta(z) + \delta^2 q)} \left(\delta^6 \partial_p G_\alpha^\delta(z, p) + o(\delta^8) \right) \\ &\quad + \frac{-\alpha \partial_z \bar{y}_\alpha^\delta(z)}{\delta^2} \left(-\delta^8 \partial_q w_\alpha^\delta(z, p, q) + o(\delta^7) \right), \\ &= \alpha \partial_z \bar{\varphi}_\alpha^\delta(z) + \delta^4 \alpha \partial_z u_\alpha^\delta(z) - \delta^5 \frac{1}{\bar{y}_\alpha^\delta(z) + \delta^2 q} \partial_p G_\alpha^\delta(z, p) \\ &\quad + \delta^6 \left(\alpha \partial_z G_\alpha^\delta(z, p) + \alpha \partial_z \bar{y}_\alpha^\delta(z) \partial_q w_\alpha^\delta(z, p, q) \right) + o(\delta^6) \end{aligned}$$

Similarly, we obtain $\partial_{yy}^2 \varphi^\delta$ up to $o(\delta^4)$ as

$$\begin{aligned} \partial_{yy}^2 \varphi^\delta &= \frac{\partial z}{\partial y} \left(\frac{\partial z}{\partial y} \partial_{zz}^2 \varphi^\delta + \frac{\partial p}{\partial y} \partial_{zp}^2 \varphi^\delta + \frac{\partial q}{\partial y} \partial_{zq}^2 \varphi^\delta \right) + \partial_z \varphi^\delta \frac{\partial^2 z}{\partial y^2} \\ &\quad + \frac{\partial p}{\partial y} \left(\frac{\partial z}{\partial y} \partial_{pz}^2 \varphi^\delta + \frac{\partial p}{\partial y} \partial_{pp}^2 \varphi^\delta + \frac{\partial q}{\partial y} \partial_{pq}^2 \varphi^\delta \right) + \partial_z \varphi^\delta \frac{\partial^2 p}{\partial y^2} \\ &\quad + \frac{\partial q}{\partial y} \left(\frac{\partial z}{\partial y} \partial_{qz}^2 \varphi^\delta + \frac{\partial p}{\partial y} \partial_{qp}^2 \varphi^\delta + \frac{\partial q}{\partial y} \partial_{qq}^2 \varphi^\delta \right) + \partial_z \varphi^\delta \frac{\partial^2 q}{\partial y^2} \\ &= \left((1 - \alpha)^2 \partial_{zz}^2 \bar{\varphi}_\alpha^\delta(z) + \delta^4 (1 - \alpha)^2 \partial_{zz}^2 u_\alpha^\delta(z) + O(\delta^5) \right) \\ &\quad + \frac{1 - \delta p}{\delta(\bar{y}_\alpha^\delta(z) + \delta^2 q)} \left(\frac{1 - \delta p}{\delta(\bar{y}_\alpha^\delta(z) + \delta^2 q)} \delta^6 \partial_{pp}^2 G_\alpha^\delta(z, p) + O(\delta^5) \right) + O(\delta^6) O\left(\frac{1}{\delta}\right) \end{aligned}$$

$$\begin{aligned}
& + \frac{1 - (1 - \alpha)\partial_z \bar{y}_\alpha^\delta(z)}{\delta^2} \left(\frac{1 - (1 - \alpha)\partial_z \bar{y}_\alpha^\delta(z)}{\delta^2} \partial_{qq}^2 w_\alpha^\delta(z, p, q) \delta^8 + O(\delta^5) \right) + O(\delta^8) O\left(\frac{1}{\delta^2}\right) \\
& = (1 - \alpha)^2 \partial_{zz}^2 \bar{\varphi}_\alpha^\delta(z) \\
& + \delta^4 \left((1 - \alpha)^2 \partial_{zz}^2 u_\alpha^\delta(z) + \frac{(1 - \delta p)^2}{(\bar{y}_\alpha^\delta + \delta^2 q)^2} \partial_{pp}^2 G_\alpha^\delta(z, p) + [1 - (1 - \alpha)\partial_z \bar{y}_\alpha^\delta(z)]^2 \partial_{qq}^2 w_\alpha^\delta(z, p, q) \right) \\
& + o(\delta^4)
\end{aligned}$$

6.4.2 Corrector equations

Using the expression in the previous section, we obtain

$$\begin{aligned}
(6.1.1) & = -\frac{1}{2} \sigma^2 y^2 \left((1 - \alpha)^2 \partial_{zz}^2 \bar{\varphi}_\alpha^\delta(z) + \delta^4 \left((1 - \alpha)^2 \partial_{zz}^2 u_\alpha^\delta(z) \right. \right. \\
& \quad \left. \left. + \frac{(1 - \delta p)^2}{(\bar{y}_\alpha^\delta + \delta^2 q)^2} \partial_{pp}^2 G_\alpha^\delta(z, p) + [1 - (1 - \alpha)\partial_z \bar{y}_\alpha^\delta(z)]^2 \partial_{qq}^2 w_\alpha^\delta(z, p, q) \right) \right) \\
& - \mu y \left((1 - \alpha)\partial_z \bar{\varphi}_\alpha^\delta(z) + \delta^4 (1 - \alpha)\partial_z u_\alpha^\delta(z) \right) \\
& - (1 - \alpha) r \delta^3 x \left(\partial_z \bar{\varphi}_\alpha^\delta(z) + \delta^4 \partial_z u_\alpha^\delta(z) \right) \\
& + \beta \left(\bar{\varphi}_\alpha^\delta(z) + \delta^4 u_\alpha^\delta(z) \right) \\
& - \tilde{U} \left(\partial_z \bar{\varphi}_\alpha^\delta(z) + \delta^4 \partial_z u_\alpha^\delta(z) + o(\delta^4) \right) + o(\delta^4) \\
& = -\frac{1}{2} \sigma^2 y^2 (1 - \alpha)^2 \partial_{zz}^2 \bar{\varphi}_\alpha^\delta(z) - \mu y (1 - \alpha) \partial_z \bar{\varphi}_\alpha^\delta(z) \\
& \quad - (1 - \alpha) r \delta^3 x \partial_z \bar{\varphi}_\alpha^\delta(z) + \beta \bar{\varphi}_\alpha^\delta(z) - \tilde{U}(\partial_z \bar{\varphi}_\alpha^\delta(z)) \\
& + \delta^4 \left(-\frac{1}{2} \sigma^2 y^2 (1 - \alpha)^2 \partial_{zz}^2 u_\alpha^\delta(z) - \mu y (1 - \alpha) \partial_z u_\alpha^\delta(z) \right. \\
& \quad \left. - (1 - \alpha) r \delta^3 x \partial_z u_\alpha^\delta(z) + \beta u_\alpha^\delta(z) - \tilde{U}'(\partial_z \bar{\varphi}_\alpha^\delta(z)) \partial_z u_\alpha^\delta(z) \right) \\
& + \delta^4 \left(-\frac{1}{2} \sigma^2 (1 - \delta p)^2 \partial_{pp}^2 G_\alpha^\delta(z, p) \right) \\
& + \delta^4 \left(\frac{1}{2} \sigma^2 y^2 [1 - (1 - \alpha)\partial_z \bar{y}_\alpha^\delta(z)]^2 \partial_{qq}^2 w_\alpha^\delta(z, p, q) \right) + o(\delta^4) \\
& \stackrel{(6.1.8)}{=} -\frac{1}{2} \sigma_\alpha^2 y^2 \partial_{zz}^2 \bar{\varphi}_\alpha^\delta(z) - \mu_\alpha y \partial_z \bar{\varphi}_\alpha^\delta(z) - r_\alpha^\delta x \partial_z \bar{\varphi}_\alpha^\delta(z) + \beta \bar{\varphi}_\alpha^\delta(z) - \tilde{U}(\partial_z \bar{\varphi}_\alpha^\delta(z)) \\
& + \delta^4 \left(-\frac{1}{2} \sigma_\alpha^2 y^2 \partial_{zz}^2 u_\alpha^\delta(z) - \mu_\alpha y \partial_z u_\alpha^\delta(z) - r_\alpha^\delta x \partial_z u_\alpha^\delta(z) + \beta u_\alpha^\delta(z) - \tilde{U}'(\partial_z \bar{\varphi}_\alpha^\delta(z)) \partial_z u_\alpha^\delta(z) \right) \\
& + \delta^4 \left(-\frac{1}{2} \sigma^2 (1 - \delta p)^2 \partial_{pp}^2 G_\alpha^\delta(z, p) \right) \\
& + \delta^4 \left(\frac{1}{2} \sigma^2 y^2 [1 - (1 - \alpha)\partial_z \bar{y}_\alpha^\delta(z)]^2 \partial_{qq}^2 w_\alpha^\delta(z, p, q) \right) + o(\delta^4) \\
& \stackrel{(6.1.7), (6.1.10)}{=} -\frac{1}{2} \sigma_\alpha^2 y^2 \partial_{zz}^2 \bar{\varphi}_\alpha^\delta(z) - \mu_\alpha y \partial_z \bar{\varphi}_\alpha^\delta(z) - r_\alpha^\delta x \partial_z \bar{\varphi}_\alpha^\delta(z) \\
& \quad + \frac{1}{2} \sigma_\alpha^2 (\bar{y}_\alpha^\delta)^2 \partial_{zz}^2 \bar{\varphi}_\alpha^\delta(z) + \bar{y}_\alpha^\delta (\mu_\alpha - r_\alpha^\delta) \partial_z \bar{\varphi}_\alpha^\delta(z) + r_\alpha^\delta z \partial_z \bar{\varphi}_\alpha^\delta(z) \\
& + \delta^4 \left(-\frac{1}{2} \sigma_\alpha^2 y^2 \partial_{zz}^2 u_\alpha^\delta(z) - \mu_\alpha y \partial_z u_\alpha^\delta(z) - r_\alpha^\delta x \partial_z u_\alpha^\delta(z) + \beta u_\alpha^\delta(z) + \bar{C}_\alpha^\delta(z) \partial_z u_\alpha^\delta(z) \right) \\
& + \delta^4 \left(-\frac{1}{2} \sigma^2 (1 - \delta p)^2 \partial_{pp}^2 G_\alpha^\delta(z, p) \right) \\
& + \delta^4 \left(\frac{1}{2} \sigma^2 y^2 [1 - (1 - \alpha)\partial_z \bar{y}_\alpha^\delta(z)]^2 \partial_{qq}^2 w_\alpha^\delta(z, p, q) \right) + o(\delta^4) \\
& \stackrel{(6.1.9)}{=} \delta^4 \left(-\frac{1}{2} \sigma_\alpha^2 \partial_{zz}^2 \bar{\varphi}_\alpha^\delta(z) q^2 - \alpha r_\alpha \partial_z \bar{\varphi}_\alpha^\delta(z) y p \right)
\end{aligned}$$

$$\begin{aligned}
& + \delta^4 \left(-\frac{1}{2} \sigma_\alpha^2 y^2 \partial_{zz}^2 u_\alpha^\delta(z) - \mu_\alpha y \partial_z u_\alpha^\delta(z) - r_\alpha^\delta x \partial_z u_\alpha^\delta(z) + \beta u_\alpha^\delta(z) + \bar{C}_\alpha^\delta(z) \partial_z u_\alpha^\delta(z) \right) \\
& + \delta^4 \left(-\frac{1}{2} \sigma^2 (1 - \delta p)^2 \partial_{pp}^2 G_\alpha^\delta(z, p) \right) \\
& + \delta^4 \left(\frac{1}{2} \sigma^2 y^2 [1 - (1 - \alpha) \partial_z \bar{y}_\alpha^\delta(z)]^2 \partial_{qq}^2 w_\alpha^\delta(z, p, q) \right) + o(\delta^4)
\end{aligned}$$

Hence,

$$\begin{aligned}
(6.1.1) & = \delta^4 \left(-\frac{1}{2} \sigma_\alpha^2 (\bar{y}_\alpha^\delta(z))^2 \partial_{zz}^2 u_\alpha^\delta(z) - (\mu_\alpha - r_\alpha^\delta) \bar{y}_\alpha^\delta(z) \partial_z u_\alpha^\delta(z) \right. \\
& \quad \left. - r_\alpha^\delta z \partial_z u_\alpha^\delta(z) + \beta u_\alpha^\delta(z) + \bar{C}_\alpha^\delta(z) \partial_z u_\alpha^\delta(z) \right) \\
& + \delta^4 \left(-\frac{1}{2} \sigma^2 (1 - \delta p)^2 \partial_{pp}^2 G_\alpha^\delta(z, p) - \alpha r_\alpha \partial_z \bar{\varphi}_\alpha^\delta(z) \bar{y}_\alpha^\delta(z) p \right) \\
& + \delta^4 \left(\frac{1}{2} \sigma^2 (\bar{y}_\alpha^\delta(z))^2 [1 - (1 - \alpha) \partial_z \bar{y}_\alpha^\delta(z)]^2 \partial_{qq}^2 w_\alpha^\delta(z, p, q) - \frac{1}{2} \sigma_\alpha^2 \partial_{zz}^2 \bar{\varphi}_\alpha^\delta(z) q^2 \right) \\
& + o(\delta^4)
\end{aligned}$$

Similarly,

$$\begin{aligned}
(6.1.2) & = (\partial_x \varphi^\delta - \partial_y \varphi^\delta - \partial_k \varphi^\delta) + \lambda_B \partial_x \varphi^\delta \delta^6 \\
& = \frac{p}{\bar{y}_\alpha^\delta(z) + \delta^2 q} \partial_p \varphi^\delta + \frac{1}{\delta^2} \partial_q \varphi^\delta + \lambda_B \partial_x \varphi^\delta \delta^6 \\
& = \delta^6 \left(\frac{p}{\bar{y}_\alpha^\delta(z)} \partial_p G_\alpha^\delta(z, p) - \partial_q w_\alpha^\delta(z, p, q) + \lambda_B \partial_z \bar{\varphi}_\alpha^\delta(z) \right) + o(\delta^6).
\end{aligned}$$

$$\begin{aligned}
(6.1.3) & = -[(1 - \alpha)y + \alpha k] \left((1 - \lambda_S \delta^6) \partial_x \varphi^\delta - \partial_y \varphi^\delta - \partial_k \varphi^\delta \right. \\
& \quad \left. + \frac{-\alpha(y - k)}{(1 - \alpha)y + \alpha k} \partial_y \varphi^\delta + \frac{(1 - \alpha)(y - k)}{(1 - \alpha)y + \alpha k} \partial_k \varphi^\delta \right) \\
& = -[(1 - \alpha)y + \alpha k] \left((1 - \lambda_S \delta^6) \partial_x \varphi^\delta - \partial_y \varphi^\delta - \partial_k \varphi^\delta \right. \\
& \quad \left. + \frac{\delta p}{1 + \delta p} \left(\frac{-1 + \alpha p \delta}{\delta(\bar{y}_\alpha^\delta(z) + \delta^2 q)} \partial_p \varphi^\delta - \frac{\alpha}{\delta^2} \partial_q \varphi^\delta \right) \right) \\
& = -[(1 - \alpha)y + \alpha k] \left(\delta^6 \left(\frac{p}{\bar{y}_\alpha^\delta(z)} \partial_p G_\alpha^\delta(z, p) - \partial_q w_\alpha^\delta(z, p, q) - \lambda_S \partial_z \bar{\varphi}_\alpha^\delta(z) \right) + o(\delta^6) \right. \\
& \quad \left. - \frac{p}{\bar{y}_\alpha^\delta(z)} \partial_p G_\alpha^\delta(z, p) \delta^6 + o(\delta^6) \right) \\
& = -[(1 - \alpha)y + \alpha k] (-\partial_q w_\alpha^\delta(z, p, q) - \lambda_S \partial_z \bar{\varphi}_\alpha^\delta(z)) \delta^6 + o(\delta^6).
\end{aligned}$$

$$\begin{aligned}
(6.1.5) & = \inf_{q_+} \left(\varphi^\delta(z, p, q) - \varphi^\delta(z - \lambda_S (\bar{y}_\alpha^\delta(z) + \delta^2 q)) \delta^6 - \lambda_B (\bar{y}_\alpha^\delta(z) + \delta^2 q_+) \delta^6 \right) \\
& = \inf_{q_+} \left(G_\alpha^\delta(z, p) \delta^6 + \partial_z \bar{\varphi}_\alpha^\delta(z) (\lambda_S + \lambda_B) \bar{y}_\alpha^\delta(z) \delta^6 - G_\alpha^\delta(z, 0) \delta^6 + o(\delta^6) \right) \\
& = \left(G_\alpha^\delta(z, p) + \partial_z \bar{\varphi}_\alpha^\delta(z) (\lambda_S + \lambda_B) \bar{y}_\alpha^\delta(z) - G_\alpha^\delta(z, 0) \right) \delta^6 + o(\delta^6)
\end{aligned}$$

Finally, introducing $a(z)$ and $b(z, p)$ for the separation of variables, we obtain

$$-\frac{1}{2} \sigma_\alpha^2 (\bar{y}_\alpha^\delta)^2 \partial_{zz}^2 u_\alpha^\delta(z) - (\mu_\alpha - r_\alpha^\delta) \bar{y}_\alpha^\delta \partial_z u_\alpha^\delta(z) - r_\alpha^\delta z \partial_z u_\alpha^\delta(z) + \beta u_\alpha^\delta(z) + \bar{C}_\alpha^\delta(z) \partial_z u_\alpha^\delta(z) - a(z) = 0,$$

$$\begin{aligned}
& \min \left\{ -\frac{1}{2}\sigma^2(1-\delta p)^2\partial_{pp}^2 G_\alpha^\delta(z,p) - \alpha r_\alpha \partial_z \bar{\varphi}_\alpha^\delta(z) \bar{y}_\alpha^\delta(z) p + a(z) + b(z,p), \right. \\
& \qquad \qquad \qquad \left. G_\alpha^\delta(z,p) + \partial_z \bar{\varphi}_\alpha^\delta(z) (\lambda_S + \lambda_B) \bar{y}_\alpha^\delta(z) - G_\alpha^\delta(z,0) \right\} = 0, \\
\min \left\{ \frac{1}{2}\sigma^2(\bar{y}_\alpha^\delta)^2 [1 - (1-\alpha)\partial_z \bar{y}_\alpha^\delta(z)]^2 \partial_{qq}^2 w_\alpha^\delta(z,p,q) - \frac{1}{2}\sigma_\alpha^2 \partial_{zz}^2 \bar{\varphi}_\alpha^\delta(z) q^2 - b(z,p), \right. \\
& \qquad \qquad \qquad \frac{p}{\bar{y}_\alpha^\delta(z)} \partial_p G_\alpha^\delta(z,p) + \lambda_B \partial_z \bar{\varphi}_\alpha^\delta(z) - \partial_q w_\alpha^\delta(z,p,q), \\
& \qquad \qquad \qquad \left. -\lambda_S \partial_z \bar{\varphi}_\alpha^\delta(z) - \partial_q w_\alpha^\delta(z,p,q) \right\} = 0.
\end{aligned}$$

Bibliography

- [ABG11] A. Arapostathis, V. S. Borkar, and M. K. Ghosh. *Ergodic control of diffusion processes*. Cambridge University Press, 2011.
- [ADG98] H. Ahn, M. Dayal, and E. R. Grannan. Option replication with transaction costs: general diffusion limits. *The Annals of Applied Probability*, 8(3):676–707, 1998.
- [AG08] L. U. Ancarani and G. Gasaneo. Derivatives of any order of the confluent hypergeometric function ${}_1F_1(a, b, z)$ with respect to the parameters a or b . *Journal of Mathematical Physics*, 49(6), 2008.
- [AKK14] W. Arendt, S. Kunkel, and M. Kunze. Diffusion with nonlocal boundary conditions. *arXiv preprint arXiv:1409.5689*, 2014.
- [AMKS15] A. Altarovici, J. Muhle-Karbe, and H. M. Soner. Asymptotics for fixed transaction costs. *Finance and Stochastics*, 19(2):363–414, 2015.
- [APW97] C. Atkinson, S. R. Pliska, and P. Wilmott. Portfolio management with transaction costs. *Proc. R. Soc. Lond. A*, 453(1958):551–562, 1997.
- [AS72] M. Abramowitz and I. A. Stegun. *Handbook of mathematical functions*, volume 1. Dover New York, 1972.
- [AW95] C. Atkinson and P. Wilmott. Portfolio management with transaction costs: an asymptotic analysis of the Morton and Pliska model. *Mathematical Finance*, 5(4):357–367, oct 1995.
- [BAP09] I. Ben-Ari and R. G. Pinsky. Ergodic behavior of diffusions with random jumps from the boundary. *Stochastic processes and their applications*, 119(3):864–881, 2009.
- [BCDG11] L. Benzoni, P. Collin-Dufresne, and R. S. Goldstein. Explaining asset pricing puzzles associated with the 1987 market crash. *Journal of Financial Economics*, 101(3):552–573, 2011.
- [BCP05] C. Buescu, A. Cadenillas, and S. R. Pliska. A note on the effects of taxes and transaction costs on optimal investment, 2005.
- [BCR13] A. Budhiraja, J. Chen, and S. Rubenthaler. A numerical scheme for invariant distributions of constrained diffusions. *Mathematics of Operations Research*, 39(2):262–289, 2013.
- [BG88] V. S. Borkar and M. K. Ghosh. Ergodic control of multidimensional diffusions I: The existence results. *SIAM Journal on Control and Optimization*, 26(1):112–126, 1988.
- [BG06] A. Budhiraja and A. P. Ghosh. Diffusion approximations for controlled stochastic networks: An asymptotic bound for the value function. *Annals of Applied Probability*, 16(4):1962–2006, 2006.

- [BG12] A. Budhiraja and A. P. Ghosh. Controlled stochastic networks in heavy traffic: Convergence of value functions. *The Annals of Applied Probability*, 22(2):734–791, 2012.
- [BGL11] A. Budhiraja, A. P. Ghosh, and Chihoon Lee. Ergodic rate control problem for single class queueing networks. *SIAM Journal on Control and Optimization*, 49(4):1570–1606, 2011.
- [Bha09] H. S. Bhamra. The levered equity risk premium and credit spreads: A unified framework. *Review of Financial Studies*, 2009.
- [Bic12] M. Bichuch. Asymptotic analysis for optimal investment in finite time with transaction costs. *SIAM Journal on Financial Mathematics*, 3(1):433–458, 2012.
- [Bic14] M. Bichuch. Pricing a contingent claim liability with transaction costs using asymptotic analysis for optimal investment. *Finance and Stochastics*, 18(3):651–694, 2014.
- [Bil09] P. Billingsley. *Convergence of probability measures*. Wiley Series in Probability and Statistics. Wiley, 2009.
- [BKL00] D. Bertsimas, L. Kogan, and A. W. Lo. When is time continuous? *Journal of Financial Economics*, 55(2):173–204, 2000.
- [Bog07] V. I. Bogachev. *Measure theory*, volume 2. Springer Science & Business Media, 2007.
- [BS98] G. Barles and H. M. Soner. Option pricing with transaction costs and a nonlinear Black-Scholes equation. *Finance and Stochastics*, 2(4):369–397, 1998.
- [BST07] I. Ben Tahar, H. M. Soner, and N. Touzi. The dynamic programming equation for the problem of optimal investment under capital gains taxes. *SIAM Journal on Control and Optimization*, 46(5):1779–1801, 2007.
- [BST10] I. Ben Tahar, H. M. Soner, and N. Touzi. Merton problem with taxes: characterization, computation, and approximation. *SIAM Journal on Financial Mathematics*, 1(1):366–395, jan 2010.
- [BSV15] P. Bank, H. M. Soner, and M. Voss. Hedging with Transient Price Impact, 2015.
- [BY04] R. Bansal and A. Yaron. Risks for the long run : a potential resolution. *The Journal of Finance*, LIX(4):1481–1509, 2004.
- [CD13] Xinfu Chen and Min Dai. Asymptotics for Merton problem with small capital gain tax and interest rate, 2013.
- [CK86] H. Cremers and D. Kadelka. On weak convergence of integral functionals of stochastic processes with applications to processes taking paths in L_E^p . *Stochastic processes and their applications*, 21(2):305–317, 1986.
- [CK06] C. Costantini and T. G. Kurtz. Diffusion approximation for transport processes with general reflection boundary conditions. *Mathematical Models and Methods in Applied Sciences*, 16(05):717–762, may 2006.
- [CK15] C. Costantini and T. G. Kurtz. Viscosity methods giving uniqueness for martingale problems. *Electronic Journal of Probability*, 20(67):1–27, 2015.
- [Con83] G. M. Constantinides. Capital market equilibrium with personal tax. *Econometrica: Journal of the Econometric Society*, pages 611–636, 1983.

- [Con84] G. M. Constantinides. Optimal stock trading with personal taxes: Implications for prices and the abnormal January returns. *Journal of Financial Economics*, 13(1):65–89, 1984.
- [Cos92] C. Costantini. The Skorohod oblique reflection problem in domains with corners and application to stochastic differential equations. *Probability theory and related fields*, 91(1):43–70, 1992.
- [CP99] A. Cadenillas and S. R. Pliska. Optimal trading of a security when there are taxes and transaction costs, 1999.
- [CZ00] A. Cadenillas and F. Zapatero. Classical and impulse stochastic control of the exchange rate using interest rates and reserves. *Mathematical Finance*, 10(2):141–156, 2000.
- [DE92] D. Duffie and L.G. Epstein. Stochastic differential utility. *Econometrica: Journal of the Econometric Society*, pages 353–394, 1992.
- [DE11] P. Dupuis and Richard S Ellis. *A weak convergence approach to the theory of large deviations*, volume 902. John Wiley & Sons, 2011.
- [DK96] P. Dybvig and H.-K. Koo. Investment with taxes, 1996.
- [DK10] E. Denis and Y. Kabanov. Mean square error for the Leland-Lott hedging strategy: convex pay-offs. *Finance and Stochastics*, 14(4):625–667, jul 2010.
- [DPZ93] M. H. A. Davis, V. G. Panas, and T. Zariphopoulou. European option pricing with transaction costs. *SIAM Journal on Control and Optimization*, 31(2):470–493, mar 1993.
- [DSZ01] R. M. Dammon, C. S. Spatt, and H. H. Zhang. Optimal consumption and investment with capital gains taxes. *The Review of Financial Studies*, 14(3):583–616, 2001.
- [DSZ03] R. M. Dammon, C. S. Spatt, and H. H. Zhang. Capital gains taxes and portfolio rebalancing. Technical report, Tepper School of Business, 2003.
- [DU05] V. Demiguel and R. Uppal. Portfolio investment with the exact tax basis via non-linear programming. *Management science*, 51(2):277–290, 2005.
- [DW96] J. G. Dai and R. J. Williams. Existence and uniqueness of semimartingale reflecting Brownian motions in convex polyhedrons. *Theory of Probability & Its Applications*, 40(1):1–40, 1996.
- [DY09] Min Dai and Fahuai Yi. Finite-horizon optimal investment with transaction costs: A parabolic double obstacle problem. *Journal of Differential Equations*, 246(4):1445–1469, feb 2009.
- [DY13] J. G. Dai and Dacheng Yao. Brownian inventory models with convex holding cost, Part 1: Average-optimal controls. *Stochastic Systems*, 3(2):442–499, 2013.
- [EJPS98] N. El Karoui, M. Jeanblanc-Picqué, and S. Shreve. Robustness of the Black and Scholes formula. *Mathematical Finance*, 8(2):93–126, 1998.
- [Fuk11a] M. Fukasawa. Asymptotically efficient discrete hedging. In *Stochastic Analysis with Financial Applications*, pages 331–346. Springer Basel, 2011.
- [Fuk11b] M. Fukasawa. Conservative delta hedging under transaction costs. *Recent Advances in Financial Engineering*, pages 1–18, mar 2011.

- [Fuk11c] M. Fukasawa. Discretization error of stochastic integrals. *The Annals of Applied Probability*, 21(4):1436–1465, apr 2011.
- [Fuk14] M. Fukasawa. Efficient discretization of stochastic integrals. *Finance and Stochastics*, 18(1):175–208, oct 2014.
- [FV02] P. A. Forsyth and K. R. Vetzal. Quadratic convergence for valuing American options using a penalty method. *SIAM Journal on Scientific Computing*, 23(6):2095–2122, 2002.
- [Gei02] S. Geiss. Quantitative approximation of certain stochastic integrals. *Stochastics: An International Journal of Probability and Stochastic Processes*, 73(3-4):241–270, 2002.
- [GH06] L. Garlappi and J. Huang. Are stocks desirable in tax-deferred accounts? *Journal of Public Economics*, 2006.
- [GK02] I. Grigorescu and Min Kang. Brownian motion on the figure eight. *Journal of Theoretical Probability*, 15(3):817–844, 2002.
- [GK07] I. Grigorescu and Min Kang. Ergodic properties of multidimensional Brownian motion with rebirth. *Electronic Journal of Probability*, 12(48):1299–1322, 2007.
- [GK12] I. Grigorescu and Min Kang. Markov Processes with Redistribution. 2012.
- [GKT06] M. F. Gallmeyer, R. Kaniel, and S. Tompaidis. Tax management strategies with multiple risky assets. *Journal of Financial Economics*, 2006.
- [GL14a] E. Gobet and N. Landon. Almost sure optimal hedging strategy. *The Annals of Applied Probability*, 24(4):1652–1690, aug 2014.
- [GL14b] E. Gobet and N. Landon. Optimization of joint \mathbb{P} -variations of Brownian semimartingales. *Electronic Communications in Probability*, 19(0):1–14, 2014.
- [GS96] E. R. Grannan and G. H. Swindle. Minimizing transaction costs of option hedging strategies. *Mathematical finance*, 6(4):341–364, 1996.
- [GT01] E. Gobet and E. Temam. Discrete time hedging errors for options with irregular payoffs. *Finance and Stochastics*, 5(3):357–367, 2001.
- [GT09] S. Geiss and A. Toivola. Weak convergence of error processes in discretizations of stochastic integrals and Besov spaces. *Bernoulli*, 15(4):925–954, nov 2009.
- [GW13] P. W. Glynn and R. J. Wang. Central limit theorems and large deviations for additive functionals of reflecting diffusion processes. *arXiv preprint arXiv:1307.1574*, 2013.
- [GW15a] P. Guasoni and M. Weber. Dynamic trading volume. *Mathematical Finance*, pages n/a–n/a, 2015.
- [GW15b] P. Guasoni and M. Weber. Nonlinear Price Impact and Portfolio Choice, 2015.
- [GW15c] P. Guasoni and M. Weber. Rebalancing Multiple Assets with Mutual Price Impact, 2015.
- [Ham89] J. D. Hamilton. A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica: Journal of the Econometric Society*, 1989.

- [HLS85] J.M. Harrison, H.J. Landau, and L. A. Shepp. The stationary distribution of reflected Brownian motion in a planar region. *The Annals of Probability*, 13(3):744–757, 1985.
- [HM05] T. Hayashi and P. A. Mykland. Evaluating hedging errors: an asymptotic approach. *Mathematical finance*, 15(2):309–343, 2005.
- [HMK15] M. Herdegen and J. Muhle-Karbe. Sensitivity of optimal consumption streams, 2015.
- [HN89] S. D. Hodges and A. Neuberger. Optimal replication of contingent claims under transaction costs. *Review of futures markets*, 8:222–239, 1989.
- [HSZ14] K. Helmes, R. H. Stockbridge, and C. Zhu. Impulse control of standard Brownian motion: long-term average criterion. *System Modeling and Optimization*, 2014.
- [Hyn12] R. Hynd. The eigenvalue problem of singular ergodic control. *Communications on pure and applied mathematics*, LXV:649–682, 2012.
- [Jak97] A. Jakubowski. A non-Skorohod topology on the Skorohod space. *Electronic Journal of Probability*, 2(4):1–21, 1997.
- [JM81] J. Jacod and Jean Mémin. Sur un type de convergence intermédiaire entre la convergence en loi et la convergence en probabilité. In *Séminaire de Probabilités XV 1979/80*, pages 529–546. Springer, 1981.
- [JS10] K. Janeček and S. Shreve. Futures trading with transaction costs. *Illinois Journal of Mathematics*, 54(4):1239–1284, 2010.
- [JS13] J. Jacod and A. N. Shiryaev. *Limit theorems for stochastic processes*, volume 288. Springer Science & Business Media, 2013.
- [JZ06a] A. Jack and M. Zervos. A singular control problem with an expected and a pathwise ergodic performance criterion. *International Journal of Stochastic Analysis*, 2006, 2006.
- [JZ06b] A. Jack and M. Zervos. Impulse and absolutely continuous ergodic control of one-dimensional Itô diffusions. In Y. Kabanov, R. Lipster, and J. Stoyanov, editors, *From Stochastic Calculus to Mathematical Finance: the Shiryaev Festschrift*, pages 295–314. Springer, Berlin, 2006.
- [JZ06c] A. Jack and M. Zervos. Impulse control of one-dimensional Itô diffusions with an expected and a pathwise ergodic criterion. *Applied mathematics and optimization*, 54(1):71–93, 2006.
- [Kar83] I. Karatzas. A class of singular stochastic control problems. *Advances in Applied Probability*, 15(2):225–254, 1983.
- [Kis13] M. Kisielewicz. *Stochastic Differential Inclusions and Applications*. Springer, 2013.
- [KL13] J. Kallsen and Shen Li. Portfolio Optimization under Small Transaction Costs: a Convex Duality Approach. *arXiv preprint arXiv:1309.3479*, 2013.
- [KM93] H. J. Kushner and L. F. Martins. Limit theorems for pathwise average cost per unit time problems for controlled queues in heavy traffic. *Stochastics: An International Journal of Probability and Stochastic Processes*, 42(1):25–51, 1993.

- [KMK13] J. Kallsen and J. Muhle-Karbe. Option pricing and hedging with small transaction costs. *Mathematical Finance*, 1(4):1–20, 2013.
- [KMK15] J. Kallsen and J. Muhle-Karbe. The general structure of optimal investment and consumption with small transaction costs. *Mathematical Finance*, pages n/a–n/a, 2015.
- [Kor99] R. Korn. Some applications of impulse control in mathematical finance. *Mathematical Methods of Operations Research*, 50(3):493–518, 1999.
- [KR14] Weining Kang and K. Ramanan. Characterization of stationary distributions of reflected diffusions. *The Annals of Applied Probability*, 24(4):1329–1374, 2014.
- [KS98] T. G. Kurtz and R. H. Stockbridge. Existence of Markov controls and characterization of optimal Markov controls. *SIAM Journal on Control and Optimization*, 36(2):609–653, 1998.
- [KS99] T. G. Kurtz and R. H. Stockbridge. Martingale problems and linear programs for singular control. In *37th annual Allerton Conference on Communication Control and Computing*, 1999.
- [KS01] T. G. Kurtz and R. H. Stockbridge. Stationary solutions and forward equations for controlled and singular martingale problems. *Electronic Journal of Probability*, 6(14):1–52, 2001.
- [KS09] Y. Kabanov and M. Safarian. *Markets with transaction costs: mathematical theory*. Springer, 2009.
- [KS15] T. G. Kurtz and R. H. Stockbridge. Linear programming formulations of singular stochastic control problems. *Personal communication*, 2015.
- [Kur90] T. G. Kurtz. Martingale problems for constrained Markov processes. pages 151–168. Springer-Verlag, New York, 1990.
- [Kur91] T. G. Kurtz. Random time changes and convergence in distribution under the Meyer-Zheng conditions. *The Annals of probability*, pages 1010–1034, 1991.
- [Kur11] T. G. Kurtz. Equivalence of stochastic equations and martingale problems. In *Stochastic Analysis 2010*, pages 113–130. Springer, 2011.
- [Kus01] H. J. Kushner. *Heavy traffic analysis of controlled queueing and communication networks*, volume 47. Springer, 2001.
- [Kus14] H. J. Kushner. A partial history of the early development of continuous-time non-linear stochastic systems theory. *Automatica*, 50(2):303–334, 2014.
- [Lan13] N. Landon. *Stratégies de couverture presque optimale : théorie et applications*. PhD thesis, Ecole Polytechnique, 2013.
- [Lel85] H. E. Leland. Option pricing and replication with transactions costs. *The journal of finance*, 40:1283–1301, 1985.
- [Lel99] H. E. Leland. Optimal portfolio management with transactions costs and capital gains taxes, 1999.
- [LMKW14] Ren Liu, J. Muhle-Karbe, and M. Weber. Rebalancing with linear and quadratic costs, 2014.

- [LR13] C. Lindberg and H. Rootzén. Error distributions for random grid approximations of multidimensional stochastic integrals. *The Annals of Applied Probability*, 23(2):834–857, 2013.
- [LS84] P.-L. Lions and A.-S. Sznitman. Stochastic differential equations with reflecting boundary conditions. *Communications on Pure and Applied Mathematics*, 37(4):511–537, 1984.
- [LS97] S. Levental and A. V. Skorohod. On the possibility of hedging options in the presence of transaction costs. *The Annals of Applied Probability*, 7(2):410–443, 1997.
- [LZ02] A. E. B. Lim and Xun Yu Zhou. Mean-variance portfolio selection with random parameters in a complete market. *Mathematics of Operations Research*, 27(1):101–120, 2002.
- [MMKS14] L. Moreau, J. Muhle-Karbe, and H. M. Soner. Trading with small price impact. *Swiss Finance Institute Research Paper*, feb 2014.
- [MØ97] G. Mundaca and B. Øksendal. Optimal stochastic intervention control with application to the exchange rate. *Journal of Mathematical Economics*, 29(2):225–243, 1997.
- [MRT92] J. L. Menaldi, M. Robin, and M. I. Taksar. Singular ergodic control for multidimensional Gaussian processes. *Mathematics of Control, Signals and Systems*, 5(1):93–114, 1992.
- [MT15] C. Ménassé and P. Tankov. Asymptotic indifference pricing in exponential Lévy models, 2015.
- [NÖ10] K. Nyström and T. Önskog. The Skorohod oblique reflection problem in time-dependent domains. *The Annals of Probability*, 38(6):2170–2223, 2010.
- [NW11] F. Naujokat and N. Westray. Curve following in illiquid markets. *Mathematics and Financial Economics*, 4(4):299–335, 2011.
- [OS05] B. Oksendal and A. Sulem. *Applied stochastic control of jump diffusions*, volume 498. Springer, Berlin, 2005.
- [Pes07] G. Peskir. A change-of-variable formula with local time on surfaces. In *Séminaire de probabilités XL*, pages 70–96. Springer Berlin Heidelberg, 2007.
- [Pro04] P. E. Protter. *Stochastic integration and differential equations*, volume 21. Springer, 2004.
- [PS04] S. R. Pliska and K. Suzuki. Optimal tracking for asset allocation with fixed and proportional transaction costs. *Quantitative Finance*, 4(2):233–243, 2004.
- [PST15] D. Possamaï, H. M. Soner, and N. Touzi. Homogenization and asymptotics for small transaction costs: the multidimensional case. *Communications in Partial Differential Equations*, pages 1–42, dec 2015.
- [Rog04] L. C. G. Rogers. Why is the effect of proportional transaction costs $O(\delta^{2/3})$. In *Mathematics of Finance: Proceedings of an AMS-IMS-SIAM Joint Summer Research Conference on Mathematics of Finance, June 22-26, 2003, Snowbird, Utah*, volume 351, pages 303–308. Providence, RI: American Mathematical Society, 2004.
- [Roo80] H. Rootzén. Limit distributions for the error in approximations of stochastic integrals. *The Annals of Probability*, pages 241–251, 1980.

- [RR10] C. Y. Robert and M. Rosenbaum. On the microstructural hedging error. *SIAM Journal on Financial Mathematics*, 1(1):427–453, 2010.
- [RS10] L. C. G. Rogers and S. Singh. The cost of illiquidity and its effects on hedging. *Mathematical Finance*, 20(4):597–615, 2010.
- [RT14] M. Rosenbaum and P. Tankov. Asymptotically optimal discretization of hedging strategies with jumps. *The Annals of Applied Probability*, 24(3):1002–1048, jun 2014.
- [RW88] M. I. Reiman and R. J. Williams. A boundary property of semimartingale reflecting Brownian motions. *Probability Theory and Related Fields*, 97, 1988.
- [RY99] D. Revuz and M. Yor. *Continuous martingales and Brownian motion*, volume 293. Springer Science & Business Media, 1999.
- [Sep13] A. Sepp. When You Hedge Discretely: Optimization of Sharpe Ratio for Delta-Hedging Strategy under Discrete Hedging and Transaction Costs. *Journal of Investment Strategies*, 3(1):19–59, 2013.
- [SSC95] H. M. Soner, S. Shreve, and J. Cvitanic. There is no nontrivial hedging portfolio for option pricing with transaction costs. *The Annals of Applied Probability*, 5(2):327–355, may 1995.
- [ST13] H. M. Soner and N. Touzi. Homogenization and asymptotics for small transaction costs. *SIAM Journal on Control and Optimization*, 51(4):2893–2921, dec 2013.
- [SV71] D. W. Stroock and S. R. S. Varadhan. Diffusion processes with boundary conditions. *Communications on Pure and Applied Mathematics*, 24(2):147–225, 1971.
- [SW13] L. Słomiński and T. Wojciechowski. Stochastic differential equations with time-dependent reflecting barriers. *Stochastics An International Journal of Probability and Stochastic Processes*, 85(1):27–47, feb 2013.
- [Tan79] H. Tanaka. Stochastic differential equations with reflecting boundary condition in convex regions. *Hiroshima Math. J.*, 1979.
- [Tof96] K. B. Toft. On the mean-variance tradeoff in option replication with transactions costs. *Journal of Financial and Quantitative Analysis*, 1996.
- [Tsu07] H. Tsukahara. On the convergence of measurable processes and prediction processes. *Illinois Journal of Mathematics*, 51(4):1231–1242, 2007.
- [TV09] P. Tankov and E. Voltchkova. Asymptotic analysis of hedging errors in models with jumps. *Stochastic Processes and their Applications*, 119(6):2004–2027, jun 2009.
- [TW93] L. M. Taylor and R. J. Williams. Existence and uniqueness of semimartingale reflecting Brownian motions in an orthant. *Probability Theory and Related Fields*, 96(3):283–317, 1993.
- [WHW94] P. Wilmott, T Hoggard, and A. E. Whalley. Hedging option portfolios in the presence of transaction costs. *Advances in Futures and Options Research*, 1994.
- [WW97] A. E. Whalley and P. Wilmott. An asymptotic analysis of an optimal hedging model for option pricing with transaction costs. *Mathematical Finance*, 7(3):307–324, jul 1997.

-
- [WW99] A. E. Whalley and P. Wilmott. Optimal hedging of options with small but arbitrary transaction cost structure. *European Journal of Applied Mathematics*, 10(02):117–139, 1999.
- [Xia12] Huina Xiao. *Continuous-time recursive utility maximization*. PhD thesis, National University of Singapore, 2012.
- [Zha99] Ruotao Zhang. *Couverture approchée des options Européennes*. PhD thesis, Ecole des Ponts ParisTech, 1999.
- [ZL00] Xun Yu Zhou and Duan Li. Continuous-time mean-variance portfolio selection: A stochastic LQ framework. *Applied Mathematics and Optimization*, 42(1):19–33, 2000.