NONLINEAR PDES AND AN APPLICATION TO HIGH-FREQUENCY TRADING

BY VLADIMIR FEDOROVICH LUBYSHEV

A dissertation submitted to the

Graduate School—New Brunswick

Rutgers, The State University of New Jersey

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Graduate Program in Mathematics

Written under the direction of

Professor Paul M. N. Feehan

and approved by

New Brunswick, New Jersey

May, 2015

ABSTRACT OF THE DISSERTATION

Nonlinear PDEs and an application to high-frequency trading

by Vladimir Fedorovich Lubyshev

Dissertation Director: Professor Paul M. N. Feehan

This dissertation is concerned with nonlinear partial differential equations and their financial applications. We establish a multiplicity result for positive solutions to a class of nonlinear Dirichlet problems and study an optimal trading strategy that is characterized as a solution of a stochastic control problem and the associated quasi-variational Hamilton-Jacobi-Bellman inequality.

Acknowledgements

I would like to thank my advisor, Professor Paul Feehan, for his guidance and support throughout my study as well as giving me a lot of freedom in choosing the research topic.

I would like to thank Professor Daniel Ocone for very helpful consultations about different aspects of stochastic control.

Many thanks to Professor Yanyan Li, Professor Ovidiu Savin, and Professor Erhan Bayraktar for agreeing to be on my dissertation committee.

I am thankful to the professors at Rutgers University for sharing their knowledge through wonderful lectures and personal discussions.

Table of Contents

Chapter 1

Introduction

This dissertation is concerned with nonlinear partial differential equations as well as their financial applications through stochastic control theory.

Nonlinear partial differential equations arise in many areas of modern mathematics, physics, and social sciences. In particular, they have been at the center of many important developments in differential geometry (e.g., Yamabe problem [68], a proof of Calabi conjecture [69]). In physics, they naturally appear, say, in a theory of electromagnetism in the form of nonlinear Maxwell equations (cf. [10] and references therein). And in finance, many optimization problems can be reduced to solving a nonlinear Hamilton-Jacobi-Bellman equation (cf. [21, 41] and references therein).

For applications, it is desirable that the problem under consideration has a unique solution. If a convergent iterative scheme is available, one would know a priori that the resulting solution corresponds to the function that actually describes the modeled phenomenon. However, this is not always the case, and many interesting nonlinear partial differential equations arising in mathematics, physics, and other sciences have multiple geometrically different solutions.

This dissertation studies problems with uniquely defined solutions as well as problems having multiple solutions. In Chapter 3, we will study a nonlinear Dirichlet problem that has several positive solutions and prove a lower bound on their number. Chapter 4, to the contrary, will be concerned with a nonlinear problem that has a unique solution. In that chapter, we will study a practical problem of developing an optimal trading strategy in the high-frequency setting, and the uniqueness of solutions will come in very handy. So, this dissertation covers two aspects of nonlinear partial differential equations and their applications: The theoretical problem of estimating the number of solutions and the practical aspect of applying uniquely solvable equations to modeling real life phenomena.

Estimating the number of distinct solutions of a nonlinear partial differential equation or a system of equations is one of the central and most difficult problems in the modern nonlinear analysis. Such a multiplicity problem is important both from a theoretical and practical perspectives. For example, a priori information about the number of solutions of a nonlinear problem under investigation could be extremely valuable when developing numerical algorithms. This is because on one hand, it is a priori not known which solution a (convergent) iterative scheme converge will to, and on the other hand, not every solution of the problem under consideration may have a physical, economic, or any other 'target' interpretation or meaning. In the theory of optimal control, multiplicity of solutions is also very important in the study of mathematical models described by nonlinear equations of mathematical physics [52]. In physical models, the existence of more than one solution to the underlying equation may imply that the initial hypotheses and assumptions are not sufficient to uniquely describe the observed phenomenon. Finally, the knowledge gained through a careful analysis of multiplicity of solutions can help with better understanding of the underlying process. For example, the existence of infinitely many distinct and unbounded (in the norm of an appropriate functional space) solutions of the Emden-Fowler equation suggests that we live in an expanding universe.

Chapter 3 of this dissertation will study a class of nonlinear equations that depend on a real parameter, λ. Such equations appear in different areas of mathematics and are used to model many processes arising in mechanics, physics, and other sciences (cf. [49, 53, 65]). For example, such problems occur in differential geometry when studying conformal deformation of Riemannian metric (cf. [31, 48, 63, 56]). After some transformations, this geometric problem can be reduced to studying solutions of the problem

$$\begin{cases} \Delta u + \lambda u - hu^p = 0 & \text{on } M, \\ u > 0 & \text{on } M \end{cases}$$

where $\lambda > 0$, $p > 1$ are constants and $h \geqslant 0$ is a C^1 function on a Riemannian manifold

M, which has been studied by T. Ouyang in [56].

It is often the case that there exists an λ^* such that the problem has a solution in a target function class if $\lambda < \lambda^*$ and ceases to have any when $\lambda > \lambda^*$. It is therefore important to have a method of finding such an λ^*. In [54], V. Lubyshev studied this problem for the Laplace operator and nonlinear Neumann boundary conditions using the extended functional method [46].

An interesting class of nonlinear problems are those involving concave-convex nonlinearities. Such class of problems was introduced by A. Ambrosetti, H. Brezis, and G. Cerami in [3] where they considered the problem

$$\begin{cases} -\Delta u = \lambda |u|^{\alpha-2}u + |u|^{\beta-2}u & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega, \end{cases}$$

with $1 < \alpha < 2 < \beta \leqslant 2^*$ and proved the existence of two positive solutions for a bounded domain Ω and a range of parameters $\lambda \in (0, \lambda^*)$. This result was generalized for $p \geqslant 2$ and/or other nonlinear operators, and for more general nonlinearities essentially preserving the concave-convex structure [45, 51, 50, 4]. In [55], V. Lubyshev studied a nonlinear Hamiltonian system involving polyharmonic operators and a concave-convex nonlinearity and proved a nonexistence result under certain assumptions on the coefficients and exponents of the underlying nonlinearity. The goal of Chapter 3 is to establish a higher order multiplicity result for problems with nonlinearities consisting of polynomial terms.

In Chapter 4, we will study a stochastic optimal control problem of finding an optimal trading strategy for a trader who continuously submits and cancels limit orders and can use market orders at his discretion. Nonlinear differential equations and inequalities that arise in the theory of stochastic control often appear in the form of the so-called Hamilton-Jacobi-Bellman equations and quasi-variational Hamilton-Jacobi-Bellman inequalities. For such problems, one normally works with the concept of viscosity solutions which was first introduced by M. Crandall and P.-L. Lions in 1980 (cf. [25]) and then extended by other mathematicians (cf. [36, 7, 6] and references therein).

One of the first works that applies optimal stochastic control to trading is [9], where M. Avellaneda and S. Stoikov studied a market making problem for a single dealer.

Under a Brownian motion assumption on the quote midpoint process, they provided a closed-form approximate solution in a stylized market model where the controls are continuous. Paper [37] derives an explicit formula for the optimal trading strategy in the Avellaneda and Stoikov model, and some generalizations of that model have been studied in [21, 20, 22].

In real markets, the price is not continuous and changes in multiples of the smallest price increment called a tick. From that perspective, models for the midpoint evolution that use point processes would be more realistic. Such models have been studied e.g. in [41, 33, 40]. Papers [41] and [40] study the optimal trading strategy for a single market maker who is allowed to use both limit and market orders. The clustering of market order flow was not modeled in these papers. Paper [33] models market order flow as a Cox process with intensity depending on the time elapsed since the last quote midpoint jump. This allowed to model the empirically observed clustering effect [11]. The market maker in [33] was only allowed to use limit orders. In Chapter 4, we generalize this model by allowing the trader to use market orders. This gives the trader more tools for inventory management and allows him to unwind inventory without waiting for a fill of one of his limit orders. From a mathematical standpoint, quasi-variational Hamilton-Jacobi-Bellman inequalities will be used instead of Hamilton-Jacobi-Bellman equations. We prove the existence and uniqueness of viscosity solutions and provide a numerical scheme for the problem at hand.

Summarizing, the contribution of this dissertation is as follows.

1. In Chapter 3, a higher-order multiplicity result for positive solutions of a class of nonlinear problems is obtained. This contrasts with many results in this area where the existence of two positive solutions is normally proved.

2. In Chapter 4, we generalize a high-frequency trading model introducing an opportunity to trade with market orders. This allows for a wider range of trading decisions and replaces a Hamilton-Jacobi-Bellman equation with a more complex quasi-variational Hamilton-Jacobi-Bellman inequality for which the existence and uniqueness of viscosity solutions are proved and a numerical scheme is constructed.

Chapter 2

Preliminaries

2.1 Sobolev spaces

Sobolev spaces are extremely important function spaces for the theory of partial differential equations. They are an extension of smooth functions and provide a more natural framework to study partial differential equations from a functional analysis perspective. It is often easier to show that a given partial differential equation has a solution in a Sobolev space rather than proving the existence of a smooth solution using the machinery and properties of smooth function spaces only. Some powerful embedding theorems exist that then help deduce that the obtained generalized solution is in fact smooth. This is akin to using complex numbers to solve a polynomial equation: It is often useful to go beyond the field of real numbers. For detailed expositions of Sobolev spaces, see [1, 32, 18, 61].

Definition 2.1.1 ([1])**.** Given an arbitrary domain $\Omega \subset \mathbf{R}^n$, an integer $m \geqslant 0$, and an $p \in [1, \infty]$, let

$$\|v\|_{m,p} := \begin{cases} \left(\sum_{0\leqslant|\alpha|\leqslant m} \|D^\alpha v\|^p_{L^p(\Omega)}\right)^{1/p} & \text{if } p \in [1,\infty), \\ \sup_{0\leqslant|\alpha|\leqslant m} \|D^\alpha v\|_{L^\infty(\Omega)} & \text{if } p = \infty. \end{cases}$$

Then the Sobolev space $W^{m,p}(\Omega)$ is the Banach space defined as

$$W^{m,p}(\Omega) := \{v \in L^p(\Omega) \mid D^\alpha v \in L^p(\Omega) \text{ for any } |\alpha| \leqslant m\}$$

and endowed with the norm $\|\cdot\|_{m,p}$. Here the derivatives are taken in the sense of the distribution theory. The zero-trace Sobolev space is the Banach space defined as

$$W_0^{m,p}(\Omega) := \text{the closure of } C_c^\infty(\Omega) \text{ in } W^{m,p}(\Omega)$$

with the same norm as above.

Theorem 2.1.1 (Sobolev embedding, [1]). *For any* $p \in [1, \infty)$,

$$W^{1,p}(\mathbf{R}^n) \hookrightarrow \begin{cases} L^{p^*}(\mathbf{R}^n) & \text{if } p < n, \\ C^{0,1-n/p}(\mathbf{R}^n) & \text{if } p > n \end{cases}$$

continuously, where $p^* := pn/(n-p)$ *is called the critical Sobolev exponent.*

Theorem 2.1.2 (Rellich-Kondrachov, [1]). *Let* $\Omega \subset \mathbf{R}^n$ *be a bounded domain with Lipschitz boundary, and let* $1 \leqslant p < n$. *Then*

$$W^{1,p}(\Omega) \hookrightarrow L^q(\Omega) \qquad \text{compactly for any } 1 \leqslant q < p^*.$$

2.2 Critical point theory

For many nonlinear partial differential equations, finding weak solutions is equivalent to finding critical points of a functional defined on a suitable Sobolev space. More abstractly, any such problem reduces to solving the equation

$$DJ(u) = 0, \qquad u \in X \tag{2.1}$$

where $(X, \|\cdot\|)$ is a Banach space and $J\colon X \to \mathbf{R}$ is a (Fréchet) differentiable functional. Throughout this section, by saying "differentiable" we will always mean "Fréchet differentiable". B_ρ and S_ρ will denote respectively the closed ball and sphere of radius ρ of the Banach space X.

Many methods exist in nonlinear analysis that help studying problem (2.1) (cf. [62, 23, 24, 8, 30, 66]). Among them are classical optimization theorems relying on the coercivity of the functional, infinite dimensional Morse theory, linking methods, bifurcation theorems, etc. Being powerful in one context, some of these methods stop being effective when the problem under consideration becomes more nonlinear. For example, many theorems in the bifurcation calculus, such as the two theorems of M. Crandall and P. Rabinowitz [26, 27], are the most effective when the differential operator is linear. Therefore, for the problem we will study in Chapter 3, we will have to rely on other machinery. Namely, we will extensively use the Fibering Method developed by S. I. Phozohaev [57] and the Mountain Pass Theorem developed by A. Ambrosetti and P. H. Rabinowitz [5]. The former is particularly useful when the problem has a polynomial-like nonlinearity.

2.2.1 Fibering method

In this subsection, we present the Fibering Method of S. Pohozaev [57, 59, 59, 60]. We will state its special form, the *spherical fibering*, which is most suitable for our problem.

Suppose that the norm $\|\cdot\|$ of the Banach space X is differentiable away from the origin and that

$$J = \frac{1}{p}\|\cdot\|^p + R$$

where $p > 1$ and $R\colon X \to \mathbf{R}$ is a differentiable functional.

Instead of studying critical points of the functional J directly, constrained critical points of the fibering functional

$$I(v) = J(t(v)v),$$

where $t = t(v)$ is a solution of

$$\frac{\partial}{\partial t}J(tv) = 0,$$

are studied on $S_1 := \{v \in X \mid \|v\| = 1\}$. Under certain assumptions, a constrained critical point v_* of I corresponds to a critical point $u_* = t(v_*)v_*$ of the original functional J.

These constrained critical points typically arise as extrema of I on S_1. But because S_1 is not weakly closed, the limit of a maximizing or minimizing sequence, if it exists, may no longer be in S_1. There is a regularity theorem that helps circumvent this difficulty. It reduces the problem to studying constrained critical points of an extension of I to the unit ball. For reflexive X, this new problem is often more tractable because of the weak compactness of B_1.

A rigorous formulation of the (spherical) fibering method combined with the regularity theorem is presented below. In what follows, $Df(a)$ will always denote the Fréchet derivative of the functional f at point a.

Theorem 2.2.1 ([57, 59, 58])**.** *Let $\mathscr{U}$ be an open subset of X with $\mathscr{U} \cap S_1 \neq \varnothing$. Suppose that the equation*

$$t^{p-1} + DR(tv)v = 0 \tag{2.2}$$

has a solution $t\colon B_1 \cap \mathscr{U} \to [0, \infty)$ that is differentiable on $(B_1 \cap \mathscr{U}) \setminus \{0\}$. Consider the functional $I\colon (B_1 \cap \mathscr{U}) \setminus \{0\} \to \mathbf{R}$ defined by

$$I(v) = \frac{1}{p} t(v)^p + R(t(v)v)$$

and let

$$\mathscr{M} := \{v \in B_1 \cap \mathscr{U} \mid t(v) \neq 0\}.$$

Then for any critical point $v_* \in \mathscr{M}$ of I on $B_1 \cap \mathscr{U}$,

(a) *$v_* \in S_1$;*

(b) *$u_* = t(v_*)v_*$ is a critical point of J, provided that $DH(v_*)v_* \neq 0$ where $H := \|\cdot\|$.*

2.2.2 Mountain pass theorem

It is a "classical" result that if the sub-level sets of a weak sequentially lower semicontinuous functional $J\colon X \to \mathbf{R}$ on a Banach space X are compact, then J attains its global minimum. Unfortunately, this compactness criteria is too strong even for basic functionals like

$$J(u) := \frac{1}{2} \int_\Omega |\nabla u|^2 - \int_\Omega fu, \qquad u \in H_0^1(\Omega) := W_0^{1,2}(\Omega),\ f \in L^2(\Omega)$$

which corresponds to the problem $-\Delta u = f$ on $H_0^1(\Omega)$ for a bounded domain $\Omega \subset \mathbf{R}^N$. Therefore a weaker compactness condition, known as the Palais-Smale compactness condition, has been introduced. In Chapter 3, we will need to prove the existence of a saddle point of the underlying variational functional. For that purpose, we will use the Mountain Pass Theorem by A. Ambrosetti and P. Rabinowitz [5].

Let us first define the Palais-Smale compactness condition.

Definition 2.2.1 ([62])**.** Given $c \in \mathbf{R}$, a sequence $(u_n) \subset X$ is called a Palais-Smale sequence at level c ($(\mathrm{PS})_c$-sequence, in short) for the differentiable functional $J\colon X \to \mathbf{R}$ if

$$J(u_n) \to c \qquad \text{and} \qquad DJ(u_n) \to 0 \text{ in } X^*.$$

Definition 2.2.2 ([62])**.** A differentiable functional $J\colon X \to \mathbf{R}$ is said to satisfy the Palais-Smale condition if given $c \in \mathbf{R}$, every $(\mathrm{PS})_c$-sequence for J contains a (strongly) convergent subsequence.

The theorem is question is presented below.

Theorem 2.2.2 ([5])**.** *Let $J \in C^1(X, \mathbf{R})$ satisfy the Palais-Smale condition and let*

$$\max(J(0), J(w)) < \inf_{S_\rho} J$$

for some $\|w\| > \rho > 0$. Then J has a critical point u_* such that

$$J(u_*) = \inf_{\gamma \in \Gamma} \max_{t \in [0,1]} J(\gamma(t))$$

where

$$\Gamma := \{\gamma \in C([0,1], X) \mid \gamma(0) = 0 \;\; \textit{and} \;\; \gamma(1) = w\}.$$

2.3 Ishii's lemma

In this section, we state Ishii's lemma, a technical result that is often used to prove the uniqueness of viscosity solutions. It will be be used in Chapter 4 to prove the uniqueness of viscosity solutions of a quasi-variational Hamilton-Jacobi-Bellman inequality. Viscosity solutions were introduced in the early 1980 by Pierre-Louis Lions and Michael G. Crandall as a generalization of classical solutions. They are particularly useful in studying partial differential equations arising in optimal control and differential games.

First, let us introduce some definitions. In what follows, $\mathscr{S}(N)$ will denote the set of symmetric $N \times N$ matrices.

Definition 2.3.1 ([25])**.** Let $u\colon \mathscr{O} \to \mathbf{R}$ where $\mathscr{O} \subset \mathbf{R}^N$, and let $\hat{x} \in \mathscr{O}$.

(i) The second order superjet $J^{2,+}_{\mathscr{O}} u(\hat{x})$ is defined as the set of all pairs $(p, X) \in \mathbf{R}^N \times \mathscr{S}(N)$ such that

$$u(x) \leqslant u(\hat{x}) + \langle p, x - \hat{x} \rangle + \frac{1}{2}\langle X(x - \hat{x}), x - \hat{x} \rangle + o(|x - \hat{x}|^2),$$

as $\mathscr{O} \ni x \to \hat{x}$. The second order subjet is defined by $J^{2,-}_{\mathscr{O}} u(\hat{x}) := -J^{2,+}_{\mathscr{O}}(-u)(\hat{x})$.

(ii) $\bar{J}^{2,\pm}_{\mathscr{O}}u(\hat{x})$ consists of all pairs $(p,X) \in \mathbf{R}^N \times \mathscr{S}(N)$ for which there exists a sequence $(x_n, p_n, X_n) \in \mathscr{O} \times \mathbf{R}^N \times \mathscr{S}(N)$ such that

$$(p_n, X_n) \in J^{2,\pm}_{\mathscr{O}}u(x_n) \qquad \text{and} \qquad (x_n, u(x_n), p_n, X_n) \to (\hat{x}, u(\hat{x}), p, X).$$

Definition 2.3.2 ([25])**.** Let $u\colon \mathscr{O}_T \to \mathbf{R}$ where $\mathscr{O}_T := (0,T) \times \mathscr{O}$ for some $T > 0$ and $\mathscr{O} \subset \mathbf{R}^N$, and let $(\hat{t}, \hat{x}) \in \mathscr{O}_T$.

(i) The second order parabolic superjet $\mathscr{P}^{2,+}_{\mathscr{O}}u(\hat{t},\hat{x})$ consists of all $(a,p,X) \in \mathbf{R} \times \mathbf{R}^N \times \mathscr{S}(N)$ such that

$$u(t,x) \leqslant u(\hat{t},\hat{x}) + a(t-\hat{t}) + \langle p, x-\hat{x}\rangle + \frac{1}{2}\langle X(x-\hat{x}), x-\hat{x}\rangle + o(|t-\hat{t}| + |x-\hat{x}|^2),$$

as $\mathscr{O}_T \ni (t,x) \to (\hat{t},\hat{x})$. The second order parabolic subjet is defined by $\mathscr{P}^{2,-}_{\mathscr{O}}u(\hat{t},\hat{x}) := -\mathscr{P}^{2,+}_{\mathscr{O}}(-u)(\hat{t},\hat{x})$.

The sets $\bar{\mathscr{P}}^{2,\pm}_{\mathscr{O}}u(\hat{t},\hat{x})$ are defined similarly to $\bar{J}^{2,\pm}_{\mathscr{O}}u(\hat{x})$.

For notational convenience, let

$$\mathrm{USC}(\mathscr{O}) := \{\text{upper semicontinuous functions } u\colon \mathscr{O} \to \mathbf{R}\},$$

$$\mathrm{LSC}(\mathscr{O}) := \{\text{lower semicontinuous functions } u\colon \mathscr{O} \to \mathbf{R}\}.$$

Proposition 2.3.1 ([25])**.** *Let* $\mathscr{O} \subset \mathbf{R}^N$, $\Phi \in \mathrm{USC}(\mathscr{O})$, $\Psi \in \mathrm{LSC}(\mathscr{O})$, $\Psi \geqslant 0$, *and*

$$M_\epsilon := \sup_{\mathscr{O}} F_\epsilon(x), \qquad \epsilon > 0$$

where $F_\epsilon(x) := \Phi(x) - \frac{1}{\epsilon}\Psi(x)$. *Let* $-\infty < \lim_{\epsilon\downarrow 0} M_\epsilon < \infty$ *and* $x_\epsilon \in \mathscr{O}$ *be chosen so that*

$$\lim_{\epsilon\downarrow 0}(M_\epsilon - F_\epsilon(x_\epsilon)) = 0.$$

Then the following hold:

(i) $\lim_{\epsilon\downarrow 0} \epsilon\Psi(x_\epsilon) = 0$,

(ii) $\Psi(\hat{x}) = 0$ *and* $\lim_{\epsilon\downarrow 0} M_\epsilon = \Phi(\hat{x}) = \sup_{\Psi^{-1}(0)} \Phi$ *whenever* $\hat{x} \in \mathscr{O}$ *is a limit point of* x_ϵ *as* $\epsilon \to 0$.

The following result is very useful for proving the uniqueness of viscosity solutions.

Theorem 2.3.1 (Ishii's Lemma, [25]). *Let $\mathscr{O}_i$ be a locally compact subset of $\mathbf{R}^{N_i}$ for $i = 1, \ldots, k$,*

$$\mathscr{O} := \mathscr{O}_1 \times \cdots \times \mathscr{O}_k,$$

$u_i \in \mathrm{USC}(\mathscr{O}_i)$, and φ_i be twice continuously differentiable in a neighborhood of $\mathscr{O}$. Set

$$w(x) = u_1(x_1) + \cdots + u_k(x_k) \qquad \textit{for } x = (x_1, \cdots, x_k) \in \mathscr{O},$$

and suppose that $\hat{x} = (\hat{x}_1, \cdots, \hat{x}_k) \in \mathscr{O}$ is a local maximum for $w - \varphi$ relative to $\mathscr{O}$. Then for each $\epsilon > 0$, there exists $X_i \in \mathscr{S}(N_i)$ such that

$$(D_{x_i}\varphi(\hat{x}), X_i) \in \bar{J}^{2,+}_{\mathscr{O}_i} u_i(\hat{x}_i) \qquad \textit{for } i = 1, \cdots, k,$$

and the block diagonal matrix with entries X_i satisfies

$$-\left(\frac{1}{\epsilon} + \|A\|\right) I \leqslant \begin{bmatrix} X_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & X_k \end{bmatrix} \leqslant A + \epsilon A^2$$

where $A := D^2\varphi(\hat{x}) \in \mathscr{S}(N)$, $N := N_1 + \cdots + N_k$, and

$$\|A\| := \sup\{|\lambda| \mid \lambda \textit{ is an eigenvalue of } A\} = \sup_{|\xi| \leqslant 1} (A\xi, \xi).$$

Chapter 3

Three positive solutions of a nonlinear Dirichlet problem

3.1 Introduction

In this chapter, we study the following nonlinear Dirichlet problem:

$$\begin{cases} -\Delta_p u = \lambda |u|^{\alpha-2}u + |u|^{\beta-2}u - \epsilon\, f(x)|u|^{p^*-2}u & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega. \end{cases} \tag{3.1}$$

where $\Omega \subset \mathbf{R}^N$ is a bounded smooth domain, λ is a real parameter, and $\Delta_p u = \nabla \cdot (|\nabla u|^{p-2}\nabla u)$ is the p-Laplacian operator. We will assume that $p < N$. Problem (3.1) is studied under the following hypothesis on the exponents:

(H0) $1 < \alpha < p < \beta < p^*$,

(H1) $f \in C(\bar{\Omega})$ is nonnegative and $f^{-\beta/(p^*-\beta)} \in L^1(\Omega)$

where $p^* := pN/(N-p)$ is the critical Sobolev exponent for $W^{1,p}(\mathbf{R}^N)$. Our goal is to show that problem (3.1) has at least three positive solutions for any $\lambda \in (0, \lambda^*)$ and $\epsilon \in (0, \epsilon^*)$ where λ^* and ϵ^* are some positive real numbers.

Multiplicity of positive solutions to the nonlinear Dirichlet problem

$$\begin{cases} -\Delta_p u = f_\lambda(\cdot, u) & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega \end{cases} \tag{3.2}$$

has been extensively studied by many authors. In [3], problem (3.2) was studied with $p = 2$, bounded Ω, and concave-convex nonlinearity $f_\lambda(x,u) = \lambda|u|^{\alpha-2}u + |u|^{\beta-2}u$, $1 < \alpha < p < \beta \leqslant p^*$, where the existence of two positive solutions for $\lambda \in (0, \lambda^*)$, $\lambda^* > 0$, has been proved. This result was generalized for $p \geqslant 2$ and/or other nonlinear operators, and for more general nonlinearities essentially preserving the concave-convex structure (cf. [45, 51, 50, 4]). It is known that, at least for the case where Ω is the unit ball, two positive solutions for the concave-convex problem considered in [3] is the maximum one can expect (cf. [64]).

Some recent papers that study higher order multiplicity of positive solutions to nonlinear problems are [44, 70, 12]. In [44], a one-dimensional problem (3.2) with $p = 2$ and $f_\lambda(x, u) = \lambda g(u)$ has been considered for $\Omega = (-1, 1)$. Under the assumption that g is concave on $(0, \gamma)$ and convex on (γ, ∞), $\gamma > 0$, as well as other assumptions such as g having a unique positive zero and $\lim_{u\to\infty} g(u)/u = \infty$, the existence of three positive solutions has been proved for $\lambda \in (\lambda_*, \lambda^*)$, $\lambda^* > \lambda_* > 0$. In [70], problem (3.2) was considered for $p > N$ and $f_\lambda(x, u) = \lambda a(x)u^{-\gamma} + \lambda g(x, u)$, $\gamma > 0$. Under some assumptions on the coefficients and the exponent γ, the existence of three positive solutions has been proved for λ belonging to an open subinterval of $(0, \infty)$. Finally, in [12], a one-dimensional problem (3.2) (the authors considered a more general form of the differential operator) with $f_\lambda(x, u) = \lambda a(x)g(u)$ and $\Omega = (0, 1)$ has been considered. Under various assumptions, including $g(u) \leqslant u^{p-1}$ in a positive neighborhood of 0, the existence of three positive solutions has been shown for λ belonging to an open subinterval of $(0, \infty)$.

The main result of this chapter is presented below.

Theorem 3.1.1. *Under hypotheses (H0)-(H1), there exist* $\lambda^* > 0$ *and* $\epsilon^* > 0$ *such that for any* $\lambda \in (0, \lambda^*)$ *and* $\epsilon \in (0, \epsilon^*)$*, problem* (3.1) *has three distinct positive weak solutions* $u_1, u_2, u_3 \in C^{1,\mu}(\bar{\Omega})$, $0 < \mu < 1$, *such that*

$$\max(J_{\lambda,\epsilon}(u_1), J_{\lambda,\epsilon}(u_3)) < 0 < J_{\lambda,\epsilon}(u_2)$$

where

$$J_{\lambda,\epsilon}(u) = \frac{1}{p}\int_\Omega |\nabla u|^p - \frac{\lambda}{\alpha}\int_\Omega |u|^\alpha - \frac{1}{\beta}\int_\Omega |u|^\beta + \frac{\epsilon}{p^*}\int_\Omega f|u|^{p^*}$$

is an energy functional for problem (3.1).

Many powerful methods exist in nonlinear analysis that help study the multiplicity of solutions to nonlinear differential equations. These methods include Morse theory, the mountain pass lemma, fixed point theorems, and the Pohozaev fibering method to name a few (cf. [23, 24, 62, 60] and references therein). The fibering method is especially useful when the nonlinearity contains polynomial components and will serve as the underlying method in this chapter. For its applications to multiplicity of solutions to nonlinear equations and systems, cf. [59, 17, 45] and references therein.

This chapter is organized as follows. Section 3.2 studies properties of the energy functional, $J_{\lambda,\epsilon}$, necessary to prove Theorem 3.1.1. Sections 3.3–3.5 study three optimization problems, each leading to a positive solution of problem (3.1). Finally, Section 3.6 is dedicated to proving our main result, Theorem 3.1.1.

3.2 Some properties of the energy functional

It will be convenient for us to represent $J_{\lambda,\epsilon}$ as

$$J_{\lambda,\epsilon}(u) = \frac{1}{p}\int_\Omega |\nabla u|^p - \frac{\lambda}{\alpha}A(u) - \frac{1}{\beta}B(u) + \frac{\epsilon}{p^*}C(u)$$

where

$$A(u) := \int_\Omega |u|^\alpha, \qquad B(u) := \int_\Omega |u|^\beta, \qquad C(u) := \int_\Omega f|u|^{p^*}.$$

Lemma 3.2.1. *The energy functional $J_{\lambda,\epsilon}$ is coercive*[1] *for any $\lambda > 0$ and $\epsilon > 0$.*

Proof. It suffices to show that the functional

$$R_{\lambda,\epsilon}(u) := -\frac{\lambda}{\alpha}A(u) - \frac{1}{\beta}B(u) + \frac{\epsilon}{p^*}C(u) \tag{3.3}$$

is bounded from below on $W_0^{1,p}(\Omega)$. By (H1), $L^{p^*}(\Omega, f(x)dx) \hookrightarrow L^\beta(\Omega)$ continuously. Therefore, by Hölder's inequality,

$$R_{\lambda,\epsilon}(u) \geqslant c_\epsilon \|u\|_{L^\beta}^{p^*} - \frac{1}{\beta}\|u\|_{L^\beta}^\beta - \lambda c\|u\|_{L^\beta}^\alpha$$

for some constants c_ϵ and c independent of u. We complete the proof by noticing that the function

$$s \mapsto c_\epsilon s^{p^*} - \frac{1}{\beta}s^\beta - \lambda c s^\alpha$$

is bounded from below on $[0,\infty)$. $\square$

[1] *Recall that a functional $F\colon X \to \mathbf{R}$ defined on a normed vector space $(X, \|\cdot\|)$ is coercive if and only if* $\lim_{\|x\|\to\infty} F(x) = \infty$.

Let

$$\hat{J}_{\lambda,\epsilon}(t,v) := \frac{1}{p}t^p + R_{\lambda,\epsilon}(tv) \qquad \text{and} \qquad \Phi_{\lambda,\epsilon}(t,v) := t^{-\alpha}\frac{\partial}{\partial t}\hat{J}_{\lambda,\epsilon}(t,v).$$

Throughout this chapter, S_ρ and B_ρ will denote respectively the sphere and the closed ball of radius ρ in $W_0^{1,p}(\Omega)$. Clearly, $\hat{J}_{\lambda,\epsilon}(t,v) = J_{\lambda,\epsilon}(tv)$ for $v \in S_1$. Consider the equation

$$\Phi_{\lambda,\epsilon}(t,v) = t^{p-\alpha} - t^{\beta-\alpha}B(v) + \epsilon\, t^{p^*-\alpha}C(v) - \lambda A(v) = 0, \qquad t > 0. \tag{3.4}$$

Because of our assumption, (H0), on the exponents, this equation has at most three solutions for each fixed $v \in W_0^{1,p}(\Omega)$. Set

$$\mathscr{U}_{\lambda,\epsilon} := \{v \in W_0^{1,p}(\Omega) \mid \text{ equation (3.4) has three solutions in } t\}.$$

Given $v \in W_0^{1,p}(\Omega)$, the i^{th} solution of (3.4) will be denoted $t_i(v) = t_{i,\lambda,\epsilon}(v)$, where the subscripts λ and ϵ will be dropped for notational simplicity.

Lemma 3.2.2. *There exist $\lambda^* > 0$ and $\epsilon^* > 0$ such that for any $\lambda \in (0,\lambda^*)$ and $\epsilon \in (0,\epsilon^*)$, the following hold.*

(a) $\mathscr{U}_{\lambda,\epsilon} \cap S_1 \neq \varnothing$.

(b) $\mathscr{U}_{\lambda,\epsilon}$ *is open in* $W_0^{1,p}(\Omega)$.

(c) *For any* $v \in \mathscr{U}_{\lambda,\epsilon} \cap B_1$,

$$\hat{J}_{\lambda,\epsilon}(t_1(v),v) < 0 < \hat{J}_{\lambda,\epsilon}(t_2(v),v).$$

There exists an $w \in \mathscr{U}_{\lambda,\epsilon} \cap S_1$ *such that*

$$J_{\lambda,\epsilon}(t_3(w)w) < 0.$$

Proof. (a): By the Rellich-Kondrachov Theorem 2.1.2, there is an $v^* \in S_1$ such that

$$B(v^*) = \max_{v \in S_1} B(v).$$

Choose $\lambda^* > 0$ and $\epsilon^* > 0$ so that

$$\max_{t>0}\left[t^{p-\alpha} - t^{\beta-\alpha}B(v^*)\right] > \lambda^* \max_{v\in S_1} A(v) \tag{3.5}$$

and

$$\min_{t>0}\left[t^{p-\alpha} - t^{\beta-\alpha}B(v^*) + \epsilon^*\, t^{p^*-\alpha}C(v^*)\right] < 0.$$

Then it is clear that $v^* \in \mathscr{U}_{\lambda,\epsilon}$ for all $\lambda \in (0,\lambda^*)$ and $\epsilon \in (0,\epsilon^*)$.

(b): Consider the map

$$T_{\lambda,\epsilon}(t,v) := \left(\Phi_{\lambda,\epsilon}(t,v), \frac{\partial}{\partial t}\Phi_{\lambda,\epsilon}(t,v)\right).$$

We will write $T_{\lambda,\epsilon}(t,v) > 0$ (resp., $T_{\lambda,\epsilon}(t,v) < 0$) to mean that the two components of $T_{\lambda,\epsilon}(t,v)$ are strictly positive (resp., strictly negative).

It is readily seen that $v \in \mathscr{U}_{\lambda,\epsilon}$ if and only if

$$T_{\lambda,\epsilon}(s_1,v) > 0 \qquad \text{and} \qquad T_{\lambda,\epsilon}(s_2,v) < 0$$

for some $0 < s_1 < s_2 < \infty$. Now fix any $v_0 \in \mathscr{U}_{\lambda,\epsilon}$ and choose $0 < s_1^0 < s_2^0 < \infty$ with

$$T_{\lambda,\epsilon}(s_1^0,v_0) > 0 \qquad \text{and} \qquad T_{\lambda,\epsilon}(s_2^0,v_0) < 0.$$

Since $T_{\lambda,\epsilon}(t,\cdot)$ is continuous on $W_0^{1,p}(\Omega)$ for any $t>0$, there is a neighborhood $\mathscr{N}_{v_0}$ of v_0 in $W_0^{1,p}(\Omega)$ such that

$$T_{\lambda,\epsilon}(s_1^0,v) > 0 \qquad \text{and} \qquad T_{\lambda,\epsilon}(s_2^0,v) < 0 \qquad \text{for all } v \in \mathscr{N}_{v_0}.$$

This implies that $\mathscr{N}_{v_0} \subset \mathscr{U}_{\lambda,\epsilon}$. Since v_0 was arbitrary, we conclude that $\mathscr{U}_{\lambda,\epsilon}$ is open.

(c): Reduce λ^* and ϵ^*, if necessary, so that

$$\max_{t>0}\left[\frac{t^{p-\alpha}}{p} - \frac{t^{\beta-\alpha}}{\beta}B(v^*)\right] > \frac{\lambda^*}{\alpha}\max_{v\in S_1} A(v) \tag{3.6}$$

and

$$\min_{t>0}\left[\frac{t^p}{p} - \frac{t^\beta}{\beta}B(v^*) + \epsilon^*\frac{t^{p^*}}{p^*}C(v^*)\right] < 0. \tag{3.7}$$

Fix any $v \in \mathscr{U}_{\lambda,\epsilon} \cap B_1$. Then the inequality

$$\hat{J}_{\lambda,\epsilon}(t_1(v),v) < 0$$

follows from the fact that $\alpha < \min(p,\beta,p^*)$. It is clear that

$$t \mapsto \hat{J}_{\lambda,\epsilon}(t,v) \qquad \text{is increasing on } [t_1(v),t_2(v)]. \tag{3.8}$$

By (3.5), for any $v \in B_1 \setminus \{0\}$, the equation

$$\tilde{\Phi}_\lambda(t,v) := t^{p-\alpha} - t^{\beta-\alpha}B(v) - \lambda A(v) = 0$$

has exactly two solutions, $0 < \tilde{t}_1(v) < \tilde{t}_2(v) < \infty$.

Let us show that

$$t_1(v) < \tilde{t}_1(v) < \tilde{t}_2(v) < t_2(v). \tag{3.9}$$

Since $\Phi_{\lambda,\epsilon}(t,v) > \tilde{\Phi}_{\lambda,\epsilon}(t,v)$ for all $t > 0$,

$$\Phi_{\lambda,\epsilon}(\cdot,v) > 0 \qquad \text{on } [\tilde{t}_1(v), \tilde{t}_2(v)].$$

Therefore, either $[\tilde{t}_1(v), \tilde{t}_2(v)] \subset (t_1(v), t_2(v))$ or $[\tilde{t}_1(v), \tilde{t}_2(v)] \subset (t_3(v), \infty)$. We want to verify that the second option is impossible. Indeed, since $\frac{\partial}{\partial t}\Phi_{\lambda,\epsilon}(t,v) > \frac{\partial}{\partial t}\tilde{\Phi}_\lambda(t,v)$ for $t > 0$ and $\frac{\partial}{\partial t}\tilde{\Phi}_\lambda(t,v) > 0$ on $(0, \tilde{t}_1(v)]$,

$$\Phi_{\lambda,\epsilon}(\cdot,v) \qquad \text{is increasing on } (0, \tilde{t}_1(v)].$$

This implies that $\tilde{t}_1(v) < t_2(v)$, finishing the proof of (3.9).

Entertaining (3.8), (3.9), we deduce that

$$\hat{J}_{\lambda,\epsilon}(t_2(v),v) > \hat{J}_{\lambda,\epsilon}(\tilde{t}_2(v),v) \geqslant H_\lambda(\tilde{t}_2(v),v)$$

where

$$H_\lambda(t,v) := \frac{t^p}{p} - \lambda\frac{t^\alpha}{\alpha}A(v) - \frac{t^\beta}{\beta}B(v).$$

But by (3.6),

$$H_\lambda(\tilde{t}_2(v),v) = \max_{t>0} H_\lambda(t,v) > 0,$$

yielding that $\hat{J}_{\lambda,\epsilon}(t_2(v),v) > 0$.

Finally, take $w = v^*$. The minimum point for the left hand side of (3.7) is bounded from below by a positive constant as $\epsilon^* \downarrow 0$ while the first zero of $t \mapsto J_{\lambda,\epsilon}(tw)$ is $o(1)$ as $\lambda \downarrow 0$. This and (3.7) mean that we can decrease λ^*, if necessary, so that the third critical value of $t \mapsto J_{\lambda,\epsilon}(tw)$ is negative or, equivalently, $J_{\lambda,\epsilon}(t_3(w)w) < 0$.

$\square$

3.3 First critical point

Let $\lambda^* > 0$ be as in part (b) of Lemma 3.2.2. In view of inequality (3.5), equation (3.4) has the minimal solution $t = t_1(v)$ for any $v \in B_1 \setminus \{0\}$. Moreover,

$$\left.\frac{\partial}{\partial t}\right|_{t=t_1(v)} \Phi_{\lambda,\epsilon}(t, v) > 0. \tag{3.10}$$

Therefore, by the Implicit Function Theorem, t_1 is continuously differentiable on $B_1 \setminus \{0\}$.

Define the functional $I_1 \colon B_1 \setminus \{0\} \to \mathbf{R}$ by

$$I_1(v) := \hat{J}_{\lambda,\epsilon}(t_1(v), v) < 0.$$

Theorem 3.3.1. *For any $\lambda \in (0, \lambda^*)$ and $\epsilon \in (0, \epsilon^*)$, there is a nonnegative $v_* \in S_1$ such that $u_* = t_1(v_*)v_*$ is a critical point of $J_{\lambda,\epsilon}$ with $J_{\lambda,\epsilon}(u_*) < 0$.*

Proof. Consider the problem

$$c_1 := \inf_{v \in B_1 \setminus \{0\}} I_1(v). \tag{3.11}$$

Let (v_n) be its minimizing sequence, which we may assume to be nonnegative. According to Lemma 3.2.1, $t_{1,n} := t_1(v_n)$ is a bounded sequence. Therefore, without loss of generality we can assume that

$$\begin{aligned} &v_n \rightharpoonup v_* \qquad \text{weakly in } W^{1,p}(\Omega), \\ &t_{1,n} \to t_* \end{aligned}$$

for some nonnegative $v_* \in B_1$ and $0 \leqslant t_* < \infty$.

Since $c_1 < 0$, we must have $0 < t_* < \infty$. In particular, equation (3.4) implies that $v_* \not\equiv 0$. Since C is weakly lower semicontinuous on $W_0^{1,p}(\Omega)$, the infimum in (3.11) is attained at v_*.

Entertaining Theorem 2.2.1, we conclude that $v_* \in S_1$ and that $u_* = t_1(v_*)v_*$ is a critical point of $J_{\lambda,\epsilon}$. Since $v_* \in S_1$, $J_{\lambda,\epsilon}(u_*) = I_1(v_*) < 0$.

□

3.4 Second critical point

Consider the continuously differentiable functional $\bar{J}_{\lambda,\epsilon}\colon W_0^{1,p}(\Omega) \to \mathbf{R}$ defined by

$$\bar{J}_{\lambda,\epsilon}(v) := \frac{1}{p}\int_\Omega |\nabla v|^p + R_{\lambda,\epsilon}(v^+)$$

where $v^+ := \max(v, 0)$.

In $L^r(\Omega)$, $r > 1$, and $W^{1,p}(\Omega)$, a weakly and a.e. convergent subsequence does not, in general, contain a strongly convergent subsequence. The two lemmas below, which will be used in proving part (a) of Lemma 3.4.3, demonstrate which additional conditions must be met in order to be able to extract such a subsequence.

Lemma 3.4.1 (Brezis-Lieb, [19])**.** *Let (X, Σ, μ) be a measure space and (f_n) be a bounded sequence in $L^r(X, \mu)$, $1 < r < \infty$, such that $f_n \to f$ a.e. Then*

$$\|f_n - f\|_{L^r}^r = \|f_n\|_{L^r}^r - \|f\|_{L^r}^r + o(1), \qquad \text{as } n \to \infty.$$

Lemma 3.4.2 ([29])**.** *Let $p > 1$ and let $\Omega \subset \mathbf{R}^N$ be an open bounded set. Let $T(s) = s \cdot \chi_{[-1,1]}(s) + \frac{s}{|s|} \cdot \chi_{[-1,1]^c}(s)$. If $(u_n) \subset W^{1,p}(\Omega)$ is such that $u_n \rightharpoonup u$ weakly in $W^{1,p}(\Omega)$ and*

$$\int_\Omega \left(|\nabla u_n|^{p-2}\nabla u_n - |\nabla u|^{p-2}\nabla u\right) \cdot \nabla T(u_n - u) = o(1), \qquad \text{as } n \to \infty,$$

then

(a) *$\nabla u_{n_k} \to \nabla u$ a.e. for some subsequence (n_k).*

(b) *$\|\nabla(u_n - u)\|_{L^p}^p = \|\nabla u_n\|_{L^p}^p - \|\nabla u\|_{L^p}^p + o(1)$, as $n \to \infty$.*

Lemma 3.4.3. *The following statements are true.*

(a) *$\bar{J}_{\lambda,\epsilon}$ satisfies the Palais-Smale condition for any $\lambda > 0$ and $\epsilon > 0$.*

(b) *There exists an $\rho > 0$ such that for any $\lambda \in (0, \lambda^*)$ and $\epsilon \in (0, \epsilon^*)$,*

$$\max\left(\inf_{B_\rho} \bar{J}_{\lambda,\epsilon}, \inf_{W_0^{1,p}\setminus B_\rho} \bar{J}_{\lambda,\epsilon}\right) < 0 < \inf_{S_\rho} \bar{J}_{\lambda,\epsilon}.$$

Proof. (a): Fix any $c \in \mathbf{R}$ and a $(\mathrm{PS})_c$-sequence (u_n) for $\bar{J}_{\lambda,\epsilon}$:

$$\bar{J}_{\lambda,\epsilon}(u_n) \to c, \tag{3.12}$$

$$D\bar{J}_{\lambda,\epsilon}(u_n) \to 0 \qquad \text{in } W^{-1,p'}(\Omega). \tag{3.13}$$

Since the functional $\bar{J}_{\lambda,\epsilon}$ is coercive, the sequence (u_n) is bounded. Therefore, we can assume that

$$\begin{aligned} &u_n \rightharpoonup u \quad \text{weakly in } W_0^{1,p}(\Omega),\\ &u_n \to u \quad \text{in } L^r(\Omega), \qquad 1 \leqslant r < p^*,\\ &u_n \to u \quad \text{a.e. in } \Omega. \end{aligned}$$

In view of (3.13) and Lemma 3.4.2, we deduce that, up to a subsequence,

$$\nabla u_n \to \nabla u \qquad \text{a.e. in } \Omega$$

and $DJ_{\lambda,\epsilon}(u)u = 0$. Making use of the Brezis-Lieb Lemma 3.4.1, we can also assume that

$$\begin{gathered} \|u_n - u\|^p_{W_0^{1,p}} = \|u_n\|^p_{W_0^{1,p}} - \|u\|^p_{W_0^{1,p}} + o(1),\\ C(u_n - u) = C(u_n) - C(u) + o(1). \end{gathered}$$

Hence,

$$\|u_n - u\|^p_{W_0^{1,p}} + \epsilon\, C(u_n - u) = DJ_{\lambda,\epsilon}(u_n)u_n - DJ_{\lambda,\epsilon}(u)u + o(1) = o(1),$$

yielding that $u_n \to u$ strongly in $W_0^{1,p}(\Omega)$.

(b): Denote by $\rho > 0$ the value of t where the left hand side of (3.6) attains its maximum. It is clear that

$$\inf_{B_\rho} \bar{J}_{\lambda,\epsilon} < 0 < \inf_{S_\rho} \bar{J}_{\lambda,\epsilon}.$$

Since the value of t where the left hand side of (3.7) attains its minimum is $> \rho$, we also deduce that

$$\inf_{W_0^{1,p}(\Omega)\setminus B_\rho} \bar{J}_{\lambda,\epsilon} < 0.$$

□

Theorem 3.4.1. *For any $\lambda \in (0, \lambda^*)$ and $\epsilon \in (0, \epsilon^*)$, the functional $J_{\lambda,\epsilon}$ has a critical point $u_* \geqslant 0$ with $J_{\lambda,\epsilon}(u_*) > 0$.*

Proof. According to statement (b) of Lemma 3.4.3, there is an $\rho > 0$ and an $w \in W_0^{1,p}(\Omega) \setminus B_\rho$ such that

$$\max(\bar{J}_{\lambda,\epsilon}(0), \bar{J}_{\lambda,\epsilon}(w)) = 0 < \inf_{S_\rho} \bar{J}_{\lambda,\epsilon}.$$

Since, by statement (a) of the same Lemma, $\bar{J}_{\lambda,\epsilon}$ satisfies the Palais-Smale condition, we conclude from the Mountain Pass Theorem 2.2.2 that $\bar{J}_{\lambda,\epsilon}$ has a critical point u_* such that

$$\bar{J}_{\lambda,\epsilon}(u_*) = \inf_{\gamma\in\Gamma} \max_{t\in[0,1]} \bar{J}_{\lambda,\epsilon}(\gamma(t)) > 0$$

where

$$\Gamma := \{\gamma \in C([0,1], W_0^{1,p}(\Omega)) \mid \gamma(0) = 0 \text{ and } \gamma(1) = w\}.$$

To finish the proof, it suffices to show that $u_* \geqslant 0$ because then u_* will be a critical point of the original functional $J_{\lambda,\epsilon}$. Since

$$0 = D\bar{J}_{\lambda,\epsilon}(u_*)u_*^- = \int_\Omega |\nabla u_*|^{p-2}\nabla u_* \cdot \nabla u_*^- = \int_\Omega |\nabla u_*^-|^p$$

and $u_* \in W_0^{1,p}(\Omega)$, we must have $u_*^- = 0$ or, equivalently, that $u_* \geqslant 0$.

□

3.5 Third critical point

Define the functional $I_3 \colon \mathscr{U}_{\lambda,\epsilon} \to \mathbf{R}$ by

$$I_3(v) := \hat{J}_{\lambda,\epsilon}(t_3(v), v) = \min_{t>t_2(v)} \hat{J}_{\lambda,\epsilon}(t, v).$$

Since

$$\left.\frac{\partial}{\partial t}\right|_{t=t_3(v)} \Phi_{\lambda,\epsilon}(t, v) > 0, \qquad v \in \mathscr{U}_{\lambda,\epsilon},$$

the Implicit Function Theorem implies that t_3 is continuously differentiable in $W_0^{1,p}(\Omega)$.

Theorem 3.5.1. *For any $\lambda \in (0, \lambda^*)$ and $\epsilon \in (0, \epsilon^*)$, there is a nonnegative $v_* \in S_1 \cap \mathscr{U}_{\lambda,\epsilon}$ such that $u_* = t_3(v_*)v_*$ is a critical point of $J_{\lambda,\epsilon}$ with $J_{\lambda,\epsilon}(u_*) < 0$.*

Proof. By statement (c) of Lemma 3.2.2,

$$c_3 := \inf_{v\in\mathscr{U}_{\lambda,\epsilon}\cap B_1} I_3(v) < 0. \tag{3.14}$$

Let (v_n) be a minimizing sequence for problem (3.14) which we may assume to be nonnegative. Let $t_{3,n} := t_3(v_n)$. Then we can assume that

$$v_n \rightharpoonup v_* \qquad \text{weakly in } W_0^{1,p}(\Omega), \tag{3.15}$$

$$t_{3,n} \to t_3 \tag{3.16}$$

for some nonnegative $v_* \in B_1$ and $0 \leqslant t_3 \leqslant \infty$.

Since $J_{\lambda,\epsilon}$ is coercive and $I_3(v_n) \geqslant J_{\lambda,\epsilon}(t_{3,n}v_n)$, we must have $t_3 < \infty$. Entertaining (3.14), we also obtain that $0 < t_3$. In other words, $0 < t_3 < \infty$. Taking into account equation (3.4), we infer that $v_* \not\equiv 0$.

Let us show that $v_* \in \mathscr{U}_{\lambda,\epsilon}$. Let $t_{2,n} := t_2(v_n)$. Without loss of generality, we may assume that

$$t_{2,n} \to t_2$$

for some $0 \leqslant t_2 \leqslant t_3$. Since $v_* \not\equiv 0$, we deduce from (3.4) that $t_2 \neq 0$ and hence

$$0 < t_2 \leqslant t_3.$$

By the weak lower semicontinuity of C on $W_0^{1,p}(\Omega)$, we must have

$$\hat{J}_{\lambda,\epsilon}(t_3, v_*) \leqslant c_3(\lambda) < 0. \tag{3.17}$$

By statement (c) of Lemma 3.2.2,

$$0 < \hat{J}_{\lambda,\epsilon}(t_{2,n}, v_n), \qquad n \geqslant 1.$$

Without loss of generality, we can assume that the sequence $(C(v_n))$ is convergent.

If $C(v_*) = \lim_{n\to\infty} C(v_n)$, then

$$0 \leqslant \hat{J}_{\lambda,\epsilon}(t_2, v_*). \tag{3.18}$$

So, we deduce from (3.17) that

$$0 < t_2 < t_3.$$

Since α is the smallest exponent, we deduce from (3.18) that there exists an $t_1 \in (0, t_2)$ such that

$$\hat{J}_{\lambda,\epsilon}(t_1, v_*) = \min_{0<s<t_2} \hat{J}_{\lambda,\epsilon}(s, v_*) < 0. \tag{3.19}$$

Taking into account that $\lim_{t\to\infty} \hat{J}_{\lambda,\epsilon}(t, v_*) = \infty$ and entertaining (3.17)-(3.19), we conclude that equation (3.4) for $v = v_*$ has three solutions, yielding that $v_* \in \mathscr{U}_{\lambda,\epsilon}$.

Now suppose that $C(v_*) < \lim_{n\to\infty} C(v_n)$. Then

$$\Phi_{\lambda,\epsilon}(t_2, v_*) < \lim_{n\to\infty} \Phi_{\lambda,\epsilon}(t_{2,n}, v_n) = 0,$$
$$\frac{\partial}{\partial t}\Phi_{\lambda,\epsilon}(t_2, v_*) < \lim_{n\to\infty} \frac{\partial}{\partial t}\Phi_{\lambda,\epsilon}(t_{2,n}, v_n) \leqslant 0.$$

By (3.5), there is a point $s > 0$ such that $\Phi_{\lambda,\epsilon}(s, v_*) > 0$ and $\frac{\partial}{\partial t}\Phi_{\lambda,\epsilon}(\cdot, v_*) > 0$ on $(0, s)$. Consequently, $0 < s < t_2$ and $\Phi_{\lambda,\epsilon}(s, v_*) > 0 > \Phi_{\lambda,\epsilon}(t_2, v_*)$, implying that $v_* \in \mathscr{U}_{\lambda,\epsilon}$.

So, we deduce that the infimum in (3.14) is attained at v_*. Applying Theorem 2.2.1, we conclude that $v_* \in S_1$ and that $u_* = t_3(v_*)v_*$ is a critical point of $J_{\lambda,\epsilon}$. Since $v_* \in S_1$, $J_{\lambda,\epsilon}(u_*) = I_3(v_*) < 0$.

□

3.6 Proof of the main result

In this section, we prove the main result of this chapter, Theorem 3.1.1. We will also need the following regularity result.

Theorem 3.6.1 ([47]). *Let $\Omega \subset \mathbf{R}^N$ be a bounded domain with C^2 boundary and let $u \in W_0^{1,p}(\Omega)$ be a weak solution of the problem*

$$\begin{cases} -\Delta_p u = f(\cdot, u) & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega. \end{cases}$$

where $f\colon \Omega \times \mathbf{R}$ is a Carathéodory function that satisfies the inequality

$$|f(x,u)| \leqslant c(1 + |u|^{m-1}) \qquad \text{for a.e. } x \in \Omega \text{ and all } u \in \mathbf{R},$$

where $m = p^*$ if $p < N$ and $m > 1$ if $p \geqslant N$. Then $u \in C^{1,\mu}(\bar{\Omega})$ for some $\mu \in (0,1)$.

We now move on to proving the main result.

Proof of Theorem 3.1.1. Let λ^*, ϵ^* be as in Lemma 3.2.2 and fix any $\lambda \in (0, \lambda^*)$ and $\epsilon \in (0, \epsilon^*)$.

Let u_1, u_2, u_3 be functions u_* in Theorem 3.3.1, Theorem 3.4.1, and Theorem 3.5.1, respectively. Since they are critical points of the energy functional, they are weak solutions of problem (3.1). Since $J_{\lambda,\epsilon}(u_i) < 0 < J_{\lambda,\epsilon}(u_2)$ for $i = 1, 3$, $u_2 \neq u_1$ and $u_2 \neq u_3$. Let us show that $u_1 \neq u_3$. If this was not the case, we would have $v_1 = v_3$ and $t_1(v_1) = t_3(v_3)$. But then that would imply $t_1(v_3) = t_1(v_1) = t_3(v_3)$, contradicting the definition of $\mathscr{U}_{\lambda,\epsilon}$.

Applying Theorem 3.6.1, we infer that $u_1, u_2, u_3 \in C^{1,\mu}(\bar{\Omega})$ for some $\mu \in (0, 1)$, which completes our proof.

$\square$

Chapter 4

Optimal trading with limit and market orders

4.1 Introduction

In this chapter, we consider the problem of maximizing trader's utility function through a systematic trading with limit and market orders. It is assumed that all the trading happens in a single trading venue so that only one limit order book (LOB) is responsible for price formation.

The trading happens over a finite time interval $[0, T]$ where the trader makes decisions to either submit a limit order to the LOB or trade with a market order. A LOB can be thought of as a price ladder with limit orders attached to each price level. Prices with buy interest are called bids and prices with sell interest are called asks with the highest bid (resp., ask) called the best bid (resp., best ask). The quote midpoint, which is the average of the best bid and the best ask, moves according to changes in the supply (asks) and demand (bids) dynamics. In most financial markets, limit orders are filled with respect to price-time priority. In such a market, if the trader submits a buy limit order of size L at the bid price p with the cumulative interest of Q units on it, his limit order will be added to the end of the limit order queue at price p and the new quantity at that level becomes $Q + L$. The trader will receive an execution of quantity $0 < L' \leqslant L$ if and only if (a) Q units that are in front of the queue are depleted (as a result of cancellations and/or matching with incoming sell market orders) and (b) the L' units are matched with incoming sell limit orders. One could also submit the so called hidden limit orders but we do not consider them in this chapter. Submitting and canceling limit orders are free of charge. The benefit of using a limit order is that such an order is executed at a price no worse that the one specified when submitting the order. However, a limit order will be filled only if the scenario described above is realized, so that there is no guarantee of execution, and even if filled, the executed quantity can be smaller than the submitted quantity. A market order is, to the contrary, guaranteed execution but the price of execution will always be at least 1 tick worse than the best bid for a buy market order and the best ask for a sell market order. Despite its disadvantages, market orders can be useful if the trader has an urgency to liquidate his inventory (signed number of security units in portfolio).

High-frequency trading (HFT) and market microstructure have been subjects of active research recently. Early work on optimal posting of limit orders by a dealer is [42]. More recently, optimal trading with limit orders for a single dealer has been studied by Avellaneda and Stoikov in [9] by means of stochastic control theory. Under a Brownian motion assumption on the quote midpoint process, they provided a closed-form approximate solution in a stylized market model where the controls are continuous. Paper [37] derives an explicit formula for the optimal trading strategy in the Avellaneda and Stoikov model. Some generalizations of that model can be found in works [21, 20, 22]. Paper [22] models market order arrivals as a self-exciting point process, which is in line with empirical observations [11]. For studies of optimal execution, we refer to [2, 35, 39, 15, 14, 38, 43].

In real markets, the quote midpoint is a pure jump process. Problems where the quote midpoint was modeled as a point process have been studied in [41, 33, 40]. Papers [41] and [40] study the optimal trading strategy for a single market maker who is allowed to use both limit and market orders. The clustering of market order flow was not modeled in these papers. Paper [33] models market order flow as a Cox process with intensity depending on the time elapsed since the last quote midpoint jump. This allowed to model the aforementioned clustering. The market maker in [33] was only allowed to use limit orders. In this chapter, we generalize [33] and study the optimal trading strategy when both limit and market orders are allowed. One can argue that high-frequency trading firms that utilize aggressive trading earn substantially higher profits than purely passive high-frequency trading firms [13]. The availability of market orders gives the trader more tools for inventory management. The resulting impulse control problem and the associated quasi-variational Hamilton-Jacobi-Bellman inequality are studied in Sections 4.4 and 4.5.

This chapter is organized as follows. Section 4.2 discusses underlying assumptions on market variables to be used in our trading model. Section 4.3 formalizes the trading problem at hand and states the main theorem of the chapter. Section 4.4 gives a numerical scheme for solving the trading problem. Finally, Section 4.5 proves the main theorem, and Section 4.6 lists estimation procedures for different market variables found in this chapter.

4.2 Models for market observables

In this section, we specify models for observable market variables that will be used in making optimal trading decision in Section 4.3. We restrict ourselves with time horizon $[0, T]$ where time 0 is market open and T is a terminal time within the day, say, the time of market close. Statistical estimation procedures for models introduced in this section are delegated to Section 4.6.

The model for the quote midpoint. We suppose that the bid-ask spread is constant and is equal to 1 tick. Then the quote midpoint process, $(P_t)_{t \geqslant 0}$, is a jump process with jumps of ± 1 tick. More precisely, let (T_n, J_n) be the marked point process consisting of midpoint jump times, T_n, and corresponding jump directions, $J_n \in \{-1, 1\}$. Then

$$P_t = P_0 + \sum_{n=1}^{N_t} J_n$$

where P_0 is the midpoint value at market open and N_t is the number of midpoint jumps up to time t.

As in [33], we model the interdependence of jump directions by

$$J_n = J_{n-1} \zeta_n$$

where (ζ_n) is an i.i.d. sequence of $\{-1, 1\}$-valued Bernoulli random variables independent of (J_n) with

$$\mathsf{P}[\zeta_n = 1] = \frac{1+\alpha}{2}, \qquad -1 \leqslant \alpha < 1.$$

It is straightforward that

$$\alpha = \mathsf{E}[J_n J_{n-1}].$$

The Markov chain (J_n) is clearly irreducible with the stationary distribution $\pi = (1/2, 1/2)$. We have

$$\alpha = \mathrm{Corr}_\pi\,[J_n, J_{n-1}],$$

leading to the following interpretation of α with respect to distribution π: If $\alpha = 0$, then midpoint jumps are independent, while $\alpha < 0$ (resp., $\alpha > 0$) corresponds to mean-reversion (resp., momentum) of midpoint jumps.

Next, we assume that for the jump inter-arrival times $\Delta_n := T_n - T_{n-1}$, the conditional distributions

$$F_\pm(s) = \mathsf{P}[\Delta_n \leqslant s \mid J_n J_{n-1} = \pm 1]$$

are independent of n. Then N is a renewal process with inter-arrival time distribution

$$F = \frac{1+\alpha}{2} F_+ + \frac{1-\alpha}{2} F_-.$$

Figures 4.1a and 4.1b show distribution functions $F_\pm$ estimated for CME's ESU3 future contract on July 15, 2013. One can observe that F_- decays faster than F_+. This can be explained by the fact that in order for the midpoint to, say, jump up twice, two contiguous ask levels must be executed. The gamma distribution serves as a good proxy for $F_\pm$ as demonstrated by Figures 4.2a and 4.2b. We denote by $f_\pm$ and f respectively the conditional and unconditional densities of the jump inter-arrival time.

Now consider the following two processes:

$$I_t := J_{N_t} \qquad \text{and} \qquad S_t := t - \sup\{T_n \mid T_n \leqslant t\},$$

i.e., the last midpoint jump direction and the time elapsed since the last midpoint jump. Then S_t, (S_t, I_t), and (P_t, S_t, I_t) are all Markov processes with the later having infinitesimal generator

$$(\mathscr{L}\varphi)(p,s,i) = \frac{\partial \varphi}{\partial s} + j^+(s)\left(\varphi(p+i,0,i) - \varphi(p,s,i)\right) + j^-(s)\left(\varphi(p-i,0,-i) - \varphi(p,s,i)\right)$$

where

$$j^\pm(s) := \lim_{h \downarrow 0} \frac{\mathsf{P}[N_{t+h} - N_t = 1,\, I_{t+h} I_t = \pm 1 \mid S_t = s, I_t = i]}{h} = \frac{1 \pm \alpha}{2} \frac{f_\pm(s)}{1 - F(s)}$$

are midpoint jump intensities in the same and the opposite direction, respectively. Clearly,

$$j := j^+ + j^-$$

is the intensity of the renewal process N. We will assume that $j^\pm \in C_b(\mathbf{R}_+)$.

Observe that the midpoint P_t can become negative, which does not happen with equities, for example. But since we are only interested in a short time horizon, the model above should serve as a good enough approximation.

The model for market order flow. We model the flow of market orders as a marked point process (θ_n, Z_n) where θ_n and $Z_n \in \{-1, 1\}$ are respectively timestamps of market order arrivals and their 'directions' ($Z_n = \pm 1$ means that the n^{th} market order was executed at the best ask/best bid). Let $M = (M_t)_{t \geqslant 0}$ be the counting process for market orders arrived up to time t. We assume that M is a Cox process with intensity $t \mapsto \lambda(S_t)$ depending on the time, S_t, elapsed since the last quote midpoint jump. Because of trading activity clustering [16, 34], we want $\lambda = \lambda(s)$ to be decreasing. The market order flow is highly correlated with midpoint jumps, and so $\lambda = \lambda(s)$ and $j = j(s)$ should have similar shapes. When modeling $F_\pm$ through the gamma distribution, j decays exponentially, as seen in (4.19). Therefore, we will assume that

$$\lambda(s) = \lambda_0 + Ae^{-ks}, \qquad A \geqslant 0, \tag{4.1}$$

which is parsimonious and captures the exponential decay. Figure 4.3 shows the plot of λ estimated for CME's ESU3 future contract on July 15, 2013.

To evaluate market order flow intensities on the strong and weak sides of the LOB[1], let us model the interdependence of midpoint jump and trade directions through

$$Z_n = I_{\theta_n -} \kappa_n$$

where (κ_n) is an i.i.d. sequence of $\{-1, 1\}$-valued Bernoulli random variables that are independent of all other processes. Let

$$\mathsf{P}[\kappa_n = 1] = \frac{1+\rho}{2}, \qquad |\rho| \leqslant 1,$$

[1] "Strong side" (resp., "weak side") of the LOB will mean in the direction I_t (resp., $-I_t$). E.g., if $I_t = 1$, then a trade occurred on the strong side of the LOB if it was executed at the ask.

then we define the market order arrival intensities on the strong and the weak sides of the LOB by

$$\lambda^{\pm}(s) = \frac{1 \pm \rho}{2}\,\lambda(s).$$

It is readily seen that

$$\rho = \mathsf{E}[Z_n I_{\theta_n -}] = \mathrm{Corr}_{\pi}[Z_n, I_{\theta_n -}],$$

similarly to the case with α. Hence, with respect to the stationary distribution π, the sign of ρ can tell us whether market orders are more likely to be initiated in the direction of the most recent midpoint jump or in the direction opposite to the most recent midpoint jump. The former corresponds to $\rho > 0$ while the later corresponds to $\rho < 0$. If $\rho = 0$, then trades happen independently of the midpoint jump directions.

Limit order fill rates. We suppose that the trader submits limit orders of constant size of L units and gets filled on the strong (resp., weak) side of the LOB with a size distributed according to the probability law $\vartheta^{\pm} = \vartheta^{\pm}_{L}$.

4.3 The trading problem

In the trading model we are about to describe, the trader continuously submits and cancels his buy and sell limit orders at the BBO (best bid and offer) as well as submits market orders at discrete times. More precisely, his limit order submission is modeled through the control

$$\ell^{\pm}_{t} \in \{0, 1\}$$

where $\ell^{\pm}_{t} = 1$ if and only if he has a limit order on the strong (resp., weak) side of the LOB at time t, while his market order submission activity is modeled through the sequence

$$(\tau_n, \xi_n)$$

where τ_n is the n^{th} time when the trader submitted a market order and ξ_n is the corresponding signed trading quantity ($\xi_n > 0$ means buying $|\xi_n|$ units while $\xi_n < 0$ means selling $|\xi_n|$ units).

At time t, the trader will have accumulated inventory q_t and cash X_t satisfying the equations:

$$dq_t = -\ell^+_{t-} I_{t-} \left(z^+ \, dM^+_t + L \, dN^+_t\right) + \ell^-_{t-} I_{t-} \left(z^- \, dM^-_t + L \, dN^-_t\right),$$
$$dX_t = \ell^+_{t-} \left(I_{t-} P_{t-} + \frac{1}{2}\right) \left(z^+ \, dM^+_t + L \, dN^+_t\right) + \ell^-_{t-} \left(-I_{t-} P_{t-} + \frac{1}{2}\right) \left(z^- \, dM^-_t + L \, dN^-_t\right)$$

for $\tau_n \leqslant t < \tau_{n+1}$ and, right after submitting a market order,

$$\Delta q_{\tau_n} = \xi_n, \qquad \Delta X_{\tau_n} = -\xi_n P_{\tau_n -} - \left(\frac{1}{2} + \epsilon\right) |\xi_n| - \epsilon_0$$

where $\epsilon, \epsilon_0 > 0$ are respectively a per-share and a fixed trading fees. We have also assumed, for the sake of simplicity, that there are no rebates. Here $z^\pm$ are independent $\mathbf{Z}_+$-valued random variables with probability distributions $\vartheta^\pm$ introduced in the previous section.

Since the trader both provides and takes liquidity, his trading strategy is characterized by the impulse control

$$\beta = (\beta^{\text{make}}, \beta^{\text{take}}), \qquad \beta^{\text{make}} = (\ell^+_t, \ell^-_t)_{t \geqslant 0}, \qquad \beta^{\text{take}} = (\tau_n, \xi_n)_{n \in \mathbf{N}}.$$

In order to mitigate inventory risk, the trader sets a maximum inventory, $Q \in \mathbf{N}$, he is willing to hold. This introduces the constraint $q_t \in \mathscr{Q}$ where $\mathscr{Q} := \{-Q, \ldots, Q\}$.

Overall, the goal is to solve the following mean-variance optimization problem:

$$\text{maximize } \mathsf{E}\left[X_T - \gamma \int_0^T q^2_{r-} \, d[P]_r\right] \text{ over all strategies } \beta \in \mathscr{A} \text{ with } q_T = 0$$

where $\gamma \geqslant 0$ and $\mathscr{A}$ is the set of all admissible controls $\beta = (\beta^{\text{make}}, \beta^{\text{take}})$ such that

$$q_t \in \mathscr{Q}, \qquad |\xi_n| \leqslant |q_{\tau_n -}|, \qquad n \geqslant 0.$$

This last constraint on market order size means that the trader will not submit market orders larger than the current inventory. Parameter γ measures the degree to which the trader is unwilling to approach the threshold inventory. If $\gamma = 0$, then he does not care about the inventory as long as it does not exceed the threshold Q, while $\gamma > 0$ forces him to trade less on the inventory side and more on the opposite side.

It will be convenient to remove the constraint $q_T = 0$ by introducing the liquidation function

$$U(p, x, q) := x + qp - \left(\frac{1}{2} + \epsilon\right) |q| - \epsilon_0.$$

This transforms our optimization problem to

$$\text{maximize } \mathsf{E}\left[U(P_T, X_T, q_T) - \gamma \int_t^T q_r^2\, j(r)\, dr\right] \text{ over all strategies } \beta \in \mathscr{A},$$

where we have also used that $[P]_r = \sum_{k=1}^{N_r} J_k^2 = N_r$. The corresponding value function is therefore

$$u(t, p, x, s, i, q) = \max_{\beta \in \mathscr{A}} \mathsf{E}_{t,p,x,q,s,i}\left[U(P_T, X_T, q_T) - \gamma \int_t^T q_r^2\, j(r)\, dr\right], \qquad \gamma \geqslant 0 \tag{4.2}$$

for $(t, p, x, s, i, q) \in [0, T] \times G$ where $G := \mathbf{R}^2 \times \mathbf{R}_+ \times \{-1, 1\} \times \mathscr{Q}$ is the state space.

Below is the main result of this section.

Theorem 4.3.1. *The value function u is of the form*

$$u(t, p, x, s, i, q) = U(p, x, q) + v(t, s, iq) \tag{4.3}$$

where $v \in C_b([0, T] \times \mathbf{R}_+ \times \mathscr{Q})$, $v \geqslant 0$, is the unique bounded viscosity solution of the problem

$$\begin{cases} \min\big(-v_t - v_s + \Lambda(s)\, v - \left(\mathscr{L}^+ + \mathscr{L}^-\right) v \\ \quad -\left(j^+(s) - j^-(s)\right) q + \gamma j(s)\, q^2, v - \mathscr{B}v\big) = 0 \quad \text{on } [0, T) \times \mathbf{R}_+ \times \mathscr{Q}, \\ \qquad\qquad\qquad\qquad\qquad\qquad v(T, \cdot) = 0 \quad \text{on } \mathbf{R}_+ \times \mathscr{Q}. \end{cases} \tag{4.4}$$

Here $\Lambda := \lambda + j$ is the combined intensity of the midpoint jumps and trade arrivals,

$$\begin{aligned}(\mathscr{L}^\pm v)(t, s, q) := \sup_{\substack{\ell \in \{0,1\} \\ q \mp \ell L \in \mathscr{Q}}} \Bigg[&\lambda^\pm(s) \left(\int_0^\infty v(t, s, q \mp \ell z)\, \vartheta^\pm(dz)\right. \\ &+ \left.\int_0^\infty \left[\frac{1}{2}\ell z - \left(\frac{1}{2} + \epsilon\right)(|q \mp \ell z| - |q|)\right] \vartheta^\pm(dz)\right) \\ &+ j^\pm(s) \left(v(t, 0, \pm q - \ell L) - \frac{L}{2}\ell - \left(\frac{1}{2} + \epsilon\right)(|q \mp \ell L| - |q|)\right)\Bigg],\end{aligned}$$

and

$$(\mathscr{B}v)(t, s, q) := \sup_{\substack{|\eta| \leqslant |q| \\ q + \eta \in \mathscr{Q}}} \left[v(t, s, q + \eta) - \left(\frac{1}{2} + \epsilon\right)(|q + \eta| + |\eta| - |q|) - \epsilon_0\right].$$

The proof of this theorem can be found in Section 4.5.

Equation (4.3) shows that the value function consists of two components: the P&L value that the trader would lock in if he decided to unwind the inventory with a market order and a nonnegative correction term that is determined by the remaining trading time, the time elapsed since the last midpoint jump, and the "signed inventory" iq. In particular, it is important to know both the absolute inventory value and whether the inventory is on the strong or the weak side of the LOB. The trader would opt to trade with a market order when $\mathscr{B}v = v$ and would rest a limit order at the strong (resp., weak) side of the LOB if the supremum in the expression for $\mathscr{L}^+$ (resp., $\mathscr{L}^-$) is attained at $\ell = 1$.

4.4 Numerical scheme

In this section, we provide a numerical scheme for solving problem (4.4) and prove its convergence.

Discretize the time domain $[0, T]$ as $\mathbf{T}_N := \{nh\}_{n=0}^N$ where $h := T/N$. Next, discretize and localize the domain $\mathbf{R}_+$ for the inter-arrival time: Fix an $K > 0$ and set $\mathbf{G}_{M,K} := \{ih'\}_{i=0}^M$ where $h' := K/M$. We will be choosing h and h' so that $h' = \omega(h) = o(1)$ as $h \downarrow 0$ and

$$h/h' \leqslant 1/2. \tag{4.5}$$

Now approximate

$$\begin{aligned}
\frac{\partial v}{\partial t}(t, s, q) &\approx \frac{v(t+h, s, q) - v(t, s, q)}{h}, \\
\frac{\partial v}{\partial s}(t, s, q) &\approx \frac{v(t+h, s+h', q) - v(t+h, s, q)}{h'}
\end{aligned}$$

and let

$$(\Psi^{h,h'}\varphi)(t, s, q) := \max\left(\frac{h}{h'}\varphi(t+h, s+h', q) + (\Gamma^{h,h'}\varphi)(t+h, s, q), (\mathscr{B}\varphi)(t+h, s, q)\right)$$

where

$$\begin{aligned}
(\Gamma^{h,h'}\varphi)(t, s, q) := &\left(1 - \frac{h}{h'} - h\,\Lambda(s)\right)\varphi(t, s, q) \\
&+ h\left[\left(\mathscr{L}^+\varphi + \mathscr{L}^-\varphi\right)(t, s, q) + \left(j^+(s) - j^-(s)\right)q - \gamma j(s)\,q^2\right].
\end{aligned}$$

We then approximate the unique viscosity solution of (4.4) by the solution $v^{h,h',S}\colon \mathbf{T}_N \times \mathbf{G}_{M,K} \times \mathscr{Q} \to \mathbf{R}$ to the following numerical scheme:

$$\begin{cases} v^{h,h',K}(t_N,\cdot,\cdot) = 0, \\ v^{h,h',K}(t_n,s_i,q) = (\Psi^{h,h'}v^{h,h',K})(t_n,s_i,q), \quad n=0,\dots,N-1,\ i=0,\dots,M-1, \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad q \in \mathscr{Q}, \\ v^{h,h',K}(\cdot,s_M,\cdot) = v^{h,h',K}(\cdot,s_{M-1},\cdot) \end{cases} \tag{4.6}$$

Because $h' = \omega(h)$, we will often drop the superscript h' for notational convenience.

Theorem 4.4.1. *$v^{h,K} \to v$ locally uniformly, as $(h,K) \to (0,\infty)$.*

Proof. **Monotonicity.** The scheme is monotonic for

$$h < \frac{1}{2\|\Lambda\|_\infty}.$$

Indeed, for such h,

$$1 - \frac{h}{h'} - h\,\Lambda(s) \geqslant \frac{1}{2} - h\,\|\Lambda\|_\infty > 0,$$

implying that $\varphi \mapsto \Gamma^h\varphi$ is increasing. Since this is also true for $\varphi \mapsto \mathscr{B}\varphi$, we deduce that $\varphi \mapsto \Psi^h\varphi$ is increasing.

Stability. The existence and uniqueness is a direct consequence of the explicit backwards form of scheme (4.6). We now claim that there exists a constant $C > 0$ independent of h, t_n, s_i, q, K, M such that

$$0 \leqslant v^{h,K} \leqslant C \qquad \text{on } \mathbf{T}_N \times \mathbf{G}_{M,K} \times \mathscr{Q}. \tag{4.7}$$

To prove the left inequality, notice that

$$v^{h,K}(t_n,s_i,q) \geqslant (\mathscr{B}v^{h,K})(t_{n+1},s_i,q) \geqslant v^{h,K}(t_{n+1},s_i,0).$$

Since $v^{h,K}(t_N,\cdot,\cdot) = 0$, we conclude by induction that $v^{h,K} \geqslant 0$. To prove the right inequality, consider the function

$$\psi(t) := (T-t)\left[\left(\frac{1}{4\gamma} + \epsilon L\right)\|j\|_\infty + (1+\epsilon)\,\|\lambda\|_\infty(\bar\vartheta^+ \vee \bar\vartheta^-)\right]$$

where $\bar{\vartheta}^{\pm} := \int_0^{\infty} z\,\vartheta^{\pm}(dz)$. Then direct computations show that

$$\Psi^h \psi \leqslant \psi.$$

Since $v^{h,K}(t_N, \cdot, \cdot) = \psi(t_N) = 0$ and scheme (4.6) is monotonic, we conclude, by induction on t_n, that $v^{h,K} \leqslant \psi$.

Consistency. Let $\varphi \in C_b^1([0,T] \times \mathbf{R}_+ \times \mathscr{Q})$. Then it is readily seen that as $h \downarrow 0$,

$$(\mathscr{B}\varphi)(t+h, s, q) \to (\mathscr{B}\varphi)(t, s, q)$$

and

$$\begin{aligned}
-\varphi(t,s,q) + \frac{1}{h'}\varphi(t+h, s+h', q) + \frac{1}{h}\left(1 - \frac{h}{h'} - h\,\Lambda(s)\right)\varphi(t+h,s,q) \\
+ \left(\mathscr{L}^{+}\varphi + \mathscr{L}^{-}\varphi\right)(t+h,s,q) \\
\to \left(\varphi_t + \varphi_s - \Lambda(s)\,\varphi + \mathscr{L}^{+}\varphi + \mathscr{L}^{-}\varphi\right)(t,s,q).
\end{aligned}$$

Convergence. Consider the functions on $[0,T] \times \mathbf{R}_+ \times \mathscr{Q}$ defined by

$$v_*(\zeta) := \liminf_{\substack{(h,K)\to(0,\infty) \\ \zeta' \to \zeta}} v^{h,K}(\zeta'), \qquad v^*(\zeta) := \limsup_{\substack{(h,K)\to(0,\infty) \\ \zeta' \to \zeta}} v^{h,K}(\zeta').$$

By (4.7), these functions are bounded and nonnegative. To prove the convergence, it suffices to show that these functions are respectively viscosity super- and subsolutions of problem (4.4). Indeed, suppose this is true. Then by the comparison principle (Lemma 4.5.3), $v^* \leqslant v_*$. But it follows from the definition of v_* and v^* that $v_* \leqslant v^*$, implying that $v := v_* \equiv v^*$ is the unique continuous bounded viscosity solution of problem (4.4). The locally uniform convergence follows from the definitions of v_* and v^* (cf. Remark 6.4 in [25]).

We will only show that v_* is a viscosity supersolution of problem (4.4); the proof that v^* is a viscosity subsolution is similar. To this end, fix a point $\bar{\zeta} := (\bar{t}, \bar{s}, \bar{q}) \in [0,T) \times \mathbf{R}_+ \times \mathscr{Q}$ and let $\varphi \in C_b^1([0,T] \times \mathbf{R}_+ \times \mathscr{Q})$ be a test function such that $\bar{\zeta}$ is a strict global minimum point for $v_* - \varphi$. Without loss of generality, we can assume that $(v_* - \varphi)(\bar{\zeta}) = 0$. Then there are sequences $B_R(\bar{\zeta}) \ni \bar{\zeta}_n \to \bar{\zeta}$ and $(h_n, K_n) \to (0, \infty)$ such that

$$v^{h_n,K_n}(\bar{\zeta}_n) \to v_*(\bar{\zeta}) \qquad \text{and} \qquad \bar{\zeta}_n \text{ is a global minimum point for } v^{h_n,K_n} - \varphi.$$

Now let $\theta_n := (v^{h_n,K_n} - \varphi)(\bar{\zeta}_n)$, then $v^{h_n,K_n} \geqslant \varphi + \theta_n$. Hence, by the monotonicity of Ψ^h,

$$\varphi(\bar{\zeta}_n) + \theta_n = v^{h_n,K_n}(\bar{\zeta}_n) \geqslant \Psi^{h_n}(\varphi + \theta_n)(\bar{\zeta}_n) = (\Psi^{h_n}\varphi)(\bar{\zeta}_n) + \theta_n$$

or, equivalently,

$$\varphi(\bar{\zeta}_n) - (\Psi^{h_n}\varphi)(\bar{\zeta}_n) \geqslant 0.$$

Consequently,

$$\begin{aligned}\min\Big(&\frac{\varphi(\bar{\zeta}_n) - h_n'^{-1}\varphi(\bar{\zeta}_n + h_n\,\boldsymbol{e}_1 + h_n'\,\boldsymbol{e}_2) - (\Gamma^h\varphi)(\bar{\zeta}_n + h_n\,\boldsymbol{e}_1)}{h_n},\\ &\varphi(\bar{\zeta}_n) - (\mathscr{B}\varphi)(\bar{\zeta}_n + h_n\,\boldsymbol{e}_1)\Big) \geqslant 0\end{aligned}$$

where $\boldsymbol{e}_i \in \mathbf{R}^3$ is the vector having 1 at the i^{th} position and 0 elsewhere. Proceeding to the limit as $n \to \infty$ and using the consistency and equivalence of different definitions of a viscosity solution (cf. [7]), we conclude that v_* is a viscosity supersolution of (4.4). The proof is complete.

□

4.5 Proof of Theorem 4.3.1

We first establish some bounds on the value function.

Lemma 4.5.1. *The value function, u, satisfies the following a priori bounds:*

$$U(p, x, q) \leqslant u(t, p, x, q, s, i) \leqslant x + qp + C \tag{4.8}$$

where C is a constant independent of t, x, p, s, i, q.

Proof. The left inequality in (4.8) easily follows when considering the strategy that immediately unwinds all the inventory through a market order and takes no action afterwards.

To prove the right inequality in (4.8), let $V_t := X_t + q_t P_t$. Then

$$\begin{aligned} dV_t &= \frac{1}{2}\left(\ell_{t-}^{+}\,dH_t^{+} + \ell_{t-}^{-}\,dH_t^{-}\right) + q_t\,dP_t + d[q, P]_t, \qquad \tau_{n-1} < t < \tau_n,\\ \Delta V_{\tau_n} &= -\left(\frac{1}{2} + \epsilon\right)|\xi_n| - \epsilon_0 \end{aligned}$$

where $dH_t^\pm := z^\pm\, dM_t^\pm + L\, dN_t^\pm$. Since $U(P_t, X_t, q_t) \leqslant V_t$, $\Delta V_{\tau_n} \leqslant 0$, and $[q, P]_t \leqslant 0$, we must have

$$\begin{aligned} u(t,p,x,s,i,q) &\leqslant x + qp + \mathsf{E}_{t,p,x,q,s,i}\Big[\frac{L}{2}\,(N_T - N_t + M_T - M_t) + \int_t^T q_{r-}\, dP_r \\ &\quad - \gamma \int_t^T q_r^2 j(r)\, dr\Big] \\ &\leqslant x + qp + \left(\frac{L}{2} + Q\right) \mathsf{E}_{t,s}[N_T - N_t] + \frac{L}{2}\, \mathsf{E}_{t,s}[M_T - M_t]. \end{aligned}$$

Since N is a renewal process,

$$\mathsf{E}_{t,s}[N_T - N_t] \leqslant \mathsf{E}[N_T] + 1 < \infty,$$

and since M is a Cox process with intensity $\lambda(S_\cdot)$,

$$\mathsf{E}_{t,s}[M_T - M_t] = \mathsf{E}_{t,s}\left[\int_t^T \lambda(S_u)\, du\right] \leqslant \|\lambda\|_{L^\infty(\mathbf{R}_+)}\,(T-t) \leqslant \|\lambda\|_{L^\infty(\mathbf{R}_+)}\, T.$$

The proof is complete.

□

Now consider the problem

$$\begin{cases} \min\left(-u_t - u_s - \left(\mathscr{I}^+ + \mathscr{I}^-\right) u + \gamma j(s)\, q^2, u - \mathscr{K} u\right) = 0 & \text{on } [0,T) \times G, \\ u(T,p,x,s,i,q) = U(p,x,q) & \text{on } G \end{cases} \tag{4.9}$$

where

$$\begin{aligned} \mathscr{I}^\pm u := \sup_{\substack{\ell \in \{0,1\} \\ q \mp i\ell L \in \mathscr{Q}}} \Big[\lambda^\pm(s) \int_0^\infty \left[u(t,p,x+\ell z\,(\pm ip + 1/2)\,, s, i, q \mp i\ell z) - u\right] \vartheta^\pm(dz) \\ + j^\pm(s)\,(u(t, p \pm i, x + \ell L\,(\pm ip + 1/2)\,, 0, \pm i, q \mp i\ell L) - u)\Big] \end{aligned}$$

and

$$\mathscr{K} u := \sup_{\substack{|\eta| \leqslant |q|, \\ q + \eta \in \mathscr{Q}}} u(t, p, x - \eta p - |\eta|(1/2 + \epsilon) - \epsilon_0, s, i, q + \eta).$$

Our goal is to prove that the value function is the unique viscosity solution (defined below) of problem (4.9) that satisfies growth condition (4.8).

Lemma 4.5.2. *There exist positive constants A, B, C such that the function*

$$w(t,p,x,s,q) = e^{A(T-t)}\left[(U^+)^2 + s^2 + p^2 + B\right] + U$$

satisfies

$$\min\left[-w_t - w_s - (\mathscr{I}^+ + \mathscr{I}^-)\,w + \gamma j(s)q^2, w - \mathscr{K}w\right] \geqslant C \qquad \text{on } [0,T)\times G.$$

Proof. Let

$$H_0\,u := -u_t - u_s - (\mathscr{I}^+ + \mathscr{I}^-)\,u \qquad \text{and} \qquad Hu := H_0u + \gamma j(s)q^2.$$

Then

$$Hw \geqslant H_0w \geqslant H_0U + H_0w_0$$

where $w_0(t,x) := e^{A(T-t)}\left[(U^+)^2 + s^2 + p^2 + B\right]$. Since

$$\begin{aligned} U(p, x+\ell z(\pm ip + 1/2), q \mp i\ell z) - U(p,x,q) &= \frac{1}{2}\ell z - \left(\frac{1}{2}+\epsilon\right)(|q \mp i\ell z| - |q|) \\ &\leqslant (1+\epsilon)L, \end{aligned} \tag{4.10}$$

$$\begin{aligned} &U(p \pm i, x+\ell L(\pm ip + 1/2), q \mp i\ell L) - U(p,x,q) \\ &\qquad = \pm iq - \frac{L}{2}\ell - \left(\frac{1}{2}+\epsilon\right)(|q \mp i\ell L| - |q|) \\ &\qquad \leqslant \epsilon L + Q, \end{aligned} \tag{4.11}$$

we must have $H_0U \geqslant C_0$ for some constant C_0 independent of the state variables and t so that

$$\begin{aligned} Hw \geqslant C_0 + H_0w_0 &= C_0 - (\partial_t + \partial_s)\,w_0 - (\mathscr{I}^+ + \mathscr{I}^-)\,w_0 \\ &\geqslant C_0 + (A-1)\,w_0 - (\mathscr{I}^+ + \mathscr{I}^-)\,w_0, \end{aligned}$$

provided that $B \geqslant 1$.

Now notice that for any $a, b \in \mathbf{R}$,

$$(b^+)^2 - (a^+)^2 \leqslant (b-a)^+\left((a^+)^2 + b - a + 1\right). \tag{4.12}$$

This inequality is obvious for $b \leqslant a$. For $b > a$, it follows from the fact that $(a+m)^+ \leqslant a^+ + m$ for any $m \geqslant 0$ and hence

$$\begin{aligned}(b^+)^2 - (a^+)^2 \leqslant (b-a)(b^+ + a^+) &\leqslant (b-a)(2a^+ + b - a) \\ &\leqslant (b-a)\left((a^+)^2 + b - a + 1\right).\end{aligned}$$

Therefore, it follows from (4.10), (4.11), and (4.12) that $(\mathscr{I}^+ + \mathscr{I}^-)\, w_0 \leqslant C_1$ for large $B > 0$ and some constant C_1 independent of A. Hence,

$$Hw \geqslant C_0 + (A - 1 - C_1)\, w_0 \geqslant C_0 + (A - 1 - C_1)\, B > 0$$

for A large enough.

Finally, since

$$\begin{aligned}U(p, x - \eta p - |\eta|(1/2+\epsilon) - \epsilon_0, q + \eta) &= U(p,x,q) - \left(\frac{1}{2} + \epsilon\right)(|q+\eta| + |\eta| - |q|) - \epsilon_0 \\ &\leqslant U(p,x,q),\end{aligned}$$

we must have $\mathscr{K}(U^+)^2 - (U^+)^2 \leqslant 0$. So, we conclude that

$$w - \mathscr{K} w \geqslant e^{A(T-t)}\left((U^+)^2 - \mathscr{K}(U^+)^2\right) + (U - \mathscr{K} U) \geqslant U - \mathscr{K} U = \epsilon_0.$$

□

Let

$$\mathscr{I}[\cdot, u] := (\mathscr{I}^+ + \mathscr{I}^-)\, u$$

and define the function $F \colon G \times \mathbf{R}^3 \times \mathbf{R} \to \mathbf{R}$ by

$$F(x, p, \mathscr{I}) = -p_3 + \gamma j(x_3)\, x_5^2 - \mathscr{I}$$

where for notation parsimony we have denoted by $x = (x_1, x_2, x_3, x_4, x_5) = (p, x, s, i, q)$ a point in the state space, and p will henceforth denote a vector in $\mathbf{R}^3$. Then problem (4.9) can be written in the form

$$\begin{cases} \min\left(-u_t + F(x, D_{\tilde{x}} u, \mathscr{I}[t,x,u], u - \mathscr{K} u\right) = 0 & \text{on } [0,T) \times G, \\ u(T, \cdot) = U & \text{on } G \end{cases} \tag{4.13}$$

where $\tilde{x} := (x_1, x_2, x_3)$ is the "continuous part" of x.

Definition 4.5.1. Given $T > 0$, $p \geqslant 0$, we set

$$\mathscr{P}_p([0,T]\times G) := \left\{ f\colon [0,T]\times G \to \mathbf{R} \,\middle|\, f \text{ is measurable}\ ,\ \sup_{(t,x)\in[0,T]\times G} \frac{|f(t,x)|}{1+|x|^p} < \infty \right\}$$

and denote by

$$\mathscr{P}([0,T]\times G) := \bigcup_{p\geqslant 0} \mathscr{P}_p([0,T]\times G)$$

the space of measurable functions of polynomial growth on $[0,T]\times G$.

Following notations in [25], for any function $u\colon [0,T]\times G \to \mathbf{R}$, we will use u_* (resp., u^*) to denote the lower (resp., upper) semicontinuous envelope of u.

Definition 4.5.2. An u.s.c. function $u \in \mathscr{P}([0,T]\times G)$ is a viscosity subsolution of (4.13) if $u(T,\cdot) \leqslant U$ on G and for any $(a,p,X) \in \mathscr{P}^{2,+}u(\bar{t},\bar{x})$,

$$\min\left(-a + F(\bar{x}, p, \mathscr{I}[\bar{t},\bar{x},u], u(\bar{t},\bar{x}) - \mathscr{K}u(\bar{t},\bar{x})\right) \leqslant 0.$$

Similarly, a l.s.c. function $u \in \mathscr{P}([0,T]\times G)$ is a viscosity supersolution of (4.13) if $u(T,\cdot) \geqslant U$ on G and for any $(a,p,X) \in \mathscr{P}^{2,-}u(\bar{t},\bar{x})$,

$$\min\left(-a + F(\bar{x}, p, \mathscr{I}[\bar{t},\bar{x},u], u(\bar{t},\bar{x}) - \mathscr{K}u(\bar{t},\bar{x})\right) \geqslant 0.$$

Finally, a function $u \in \mathscr{P}([0,T]\times G)$ is a viscosity solution of (4.13) if u^* and u_* are respectively viscosity sub- and supersolution of (4.13).

One can also use an equivalent definition of viscosity sub- and supersolutions (cf. [7]).

Definition 4.5.3. An u.s.c. function $u \in \mathscr{P}([0,T]\times G)$ is a viscosity subsolution of (4.13) if $u(T,\cdot) \leqslant U$ on G and for any $(\bar{t},\bar{x}) \in [0,T)\times G$ and a function $\varphi \in C^{1,2}([0,T]\times G)^2$ for which $(\bar{t},\bar{x})$ is a zero global maximum point of $u-\varphi$,

$$\min\left(-\varphi_t(\bar{t},\bar{x}) + F(\bar{x}, D_{\tilde{x}}\varphi(\bar{t},\bar{x}), \mathscr{I}[\bar{t},\bar{x},u], u(\bar{t},\bar{x}) - \mathscr{K}u(\bar{t},\bar{x})\right) \leqslant 0.$$

Similarly, a l.s.c. function $u \in \mathscr{P}([0,T]\times G)$ is a viscosity supersolution of (4.13) if $u(T,\cdot) \geqslant U$ on G and for any $(\bar{t},\bar{x}) \in [0,T)\times G$ and a function $\varphi \in C^{1,2}([0,T]\times G)$ for

[2] Here, $C^{1,2}([0,T]\times G)$ denotes the space of continuous functions $[0,T]\times G \to \mathbf{R}$ that are of class $C^{1,2}$ in the "continuous variable" $(t,\tilde{x})$.

which $(\bar t,\bar x)$ is a zero global minimum point of $u-\varphi$,

$$\min\left(-\varphi_t(\bar t,\bar x)+F(\bar x,D_{\tilde x}\varphi(\bar t,\bar x),\mathscr{I}[\bar t,\bar x,u],u(\bar t,\bar x)-\mathscr{K}u(\bar t,\bar x)\right)\geqslant 0.$$

Finally, a function $u\in\mathscr{P}([0,T]\times G)$ is a viscosity solution of (4.13) if u^* and u_* are respectively viscosity sub- and supersolution of (4.13).

Lemma 4.5.3 (Uniqueness)**.** *Problem* (4.9) *has at most one viscosity solution satisfying a priori bounds* (4.8)*.*

Proof. **1.** Let

$$w(t,x)=U(x_1,x_2,x_5)+e^{A(T-t)}\left((U(x_1,x_2,x_5)^+)^2+x_1^2+x_3^2+B\right)$$

be as in Lemma 4.5.2. We claim that for any $\delta\in(0,1)$, the functions

$$u^\delta:=(1+\delta)\,u-\delta\,w\qquad\text{and}\qquad v^\delta:=(1-\delta)\,v+\delta\,w$$

are respectively viscosity sub- and supersolutions of the problems

$$\min(-u_t+F(x,D_{\tilde x}u,\mathscr{I}[t,x,u]),u-\mathscr{K}u)+\delta C=0\qquad\text{on }[0,T)\times G\tag{4.14}$$

and

$$\min(-u_t+F(x,D_{\tilde x}u,\mathscr{I}[t,x,u]),u-\mathscr{K}u)-\delta C=0\qquad\text{on }[0,T)\times G.\tag{4.15}$$

First, let us prove that u^δ is a viscosity subsolution of (4.14). To this end, fix any φ^δ such that $(\bar t,\bar x)$ is a zero global maximum point for $u^\delta-\varphi^\delta$. Set $\varphi:=(1+\delta)^{-1}(\varphi^\delta+\delta\,w)$ so that $u-\varphi$ attains its zero global maximum at $(\bar t,\bar x)$. We now work on the case by case basis. Suppose that

$$-\varphi_t(\bar t,\bar x)+F(\bar x,D_{\tilde x}\varphi(\bar t,\bar x),\mathscr{I}[\bar t,\bar x,u])\leqslant 0$$

and let $R[\bar t,\bar x,\varphi,u]:=-\varphi_t(\bar t,\bar x)+F(\bar x,D_{\tilde x}\varphi(\bar t,\bar x),\mathscr{I}[\bar t,\bar x,u])$. Then

$$\begin{aligned}R[\bar t,\bar x,\varphi^\delta,u^\delta]&\leqslant R[\bar t,\bar x,\varphi^\delta,u^\delta]+\delta\,R[\bar t,\bar x,w,w]-\delta C\\&\leqslant R[\bar t,\bar x,\varphi^\delta+\delta w,u^\delta+\delta w]+\delta\,\gamma j(x_3)x_5^2-\delta C\\&=(1+\delta)\,R[\bar t,\bar x,\varphi,u]-\delta C\\&\leqslant-\delta C.\end{aligned}$$

Now suppose that $(u - \mathscr{K} u)(\bar{t}, \bar{x}) \leqslant 0$. Then

$$\begin{aligned}(u^\delta - \mathscr{K} u^\delta)(\bar{t}, \bar{x}) &= (1+\delta)\, u(\bar{t}, \bar{x}) - \delta\, w(\bar{t}, \bar{x}) - \mathscr{K} u^\delta(\bar{t}, \bar{x}) \\ &\leqslant (1+\delta)\, u(\bar{t}, \bar{x}) - \delta\, w(\bar{t}, \bar{x}) - ((1+\delta)\mathscr{K} u(\bar{t}, \bar{x}) - \delta \mathscr{K} w(\bar{t}, \bar{x})) \\ &= (1+\delta)\,(u - \mathscr{K} u)(\bar{t}, \bar{x}) - \delta\,(w - \mathscr{K} w)(\bar{t}, \bar{x}) \\ &\leqslant -\delta C.\end{aligned}$$

This concludes the proof that u^δ is a viscosity subsolution of (4.14). Now let us show that v^δ is a viscosity supersolution of (4.15). To this end, fix any φ^δ such that $(\bar{t}, \bar{x})$ is a zero global minimum point for $v^\delta - \varphi^\delta$. Set $\varphi := (1-\delta)^{-1}(\varphi^\delta - \delta\, w)$ so that $v - \varphi$ attains its zero global minimum at $(\bar{t}, \bar{x})$. We have

$$\begin{aligned}R[\bar{t}, \bar{x}, \varphi^\delta, v^\delta] &\geqslant R[\bar{t}, \bar{x}, \varphi^\delta, v^\delta] - \delta\, R[\bar{t}, \bar{x}, w, w] + \delta C \\ &\geqslant R[\bar{t}, \bar{x}, \varphi^\delta - \delta\, w, v^\delta - \delta\, w] + \delta C \\ &= (1+\delta)\, R[\bar{t}, \bar{x}, \varphi, v] + \delta C \\ &\geqslant \delta C.\end{aligned}$$

By the convexity of $\mathscr{K}$,

$$\begin{aligned}(v^\delta - \mathscr{K} v^\delta)(\bar{t}, \bar{x}) &\geqslant v^\delta(\bar{t}, \bar{x}) - (1-\delta)\mathscr{K} v(\bar{t}, \bar{x}) - \delta \mathscr{K} w(\bar{t}, \bar{x}) \\ &\geqslant (v^\delta - (1-\delta)\, v)(\bar{t}, \bar{x}) - \delta \mathscr{K} w(\bar{t}, \bar{x}) \\ &= \delta\,(w - \mathscr{K} w)(\bar{t}, \bar{x}) \\ &\geqslant \delta C.\end{aligned}$$

Finally, since $w(T, \cdot) \geqslant U$, $u(T, \cdot) \leqslant U$, and $v(T, \cdot) \geqslant U$, we must have

$$u^\delta(T, \cdot) \leqslant U \leqslant v^\delta(T, \cdot).$$

2. Let us show that

$$\lim_{|x| \to \infty,\, x \in S} u^\delta(t, x) = -\infty \qquad \text{and} \qquad \lim_{|x| \to \infty,\, x \in S} v^\delta(t, x) = \infty. \tag{4.16}$$

We will only prove the statement for u^δ because the proof for v^δ is similar. Since $\mathscr{Q}$ is bounded, it follows from (4.8) that $u \leqslant U + \text{const}$ and hence

$$u^\delta \leqslant U - \delta\left((U^+)^2 + x_1^2 + x_3^2\right) + (1+\delta)\, C.$$

Now let $|x_n| \to \infty$. Since x_5 is discrete and bounded, we can assume without loss of generality that $x_5^{(n)} = x_5^*$ for large n. If $U(x_n) \to \pm\infty$, then it is obvious that $u^\delta(x_n) \to -\infty$. Suppose that the sequence $U(x_n)$, or equivalently the sequence $x_2^{(n)} + x_1^{(n)} x_5^*$, is bounded. If $(x_1^{(n)})^2 + (x_3^{(n)})^2 \to \infty$, then the statement is immediate. Finally, if $|x_2^{(n)}| \to \infty$, then we must have $|x_1^{(n)}| \to \infty$ so we are in the situation of the previous sentence, and hence $u^\delta(x_n) \to -\infty$ again.

Since $u^\delta \to u$ and $v^\delta \to v$ as $\delta \downarrow 0$, it suffices to show that $u^\delta \leqslant v^\delta$ for small δ. So, we can assume that $u = u^\delta$, $v = v^\delta$ and that

$$\lim_{|x|\to\infty,\, x\in S} u(t,x) = -\infty \qquad \text{and} \qquad \lim_{|x|\to\infty,\, x\in S} v(t,x) = \infty \tag{4.17}$$

uniformly in $t \in [0,T]$.

Now suppose, contrary to our claim, that

$$M := \max_{(t,x)\in[0,T)\times G} (u(t,x) - v(t,x)) > 0$$

and let $(\bar{t}, \bar{x}) \in [0,T) \times G$ be the corresponding maximum point. Fix an arbitrary $\epsilon > 0$ and consider the optimization problem

$$M_\epsilon := \sup_{(t,x),\,(s,y)\in[0,T)\times G} \left[u(t,x) - v(s,y) - \frac{1}{2\epsilon}\left((t-s)^2 + |x-y|^2\right) \right].$$

Thanks to (4.17), M_ϵ must be attained at some points (t_ϵ, x_ϵ) and (s_ϵ, y_ϵ). By Proposition 2.3.1, we can assume that as $\epsilon \downarrow 0$,

$$(t_\epsilon, x_\epsilon) \to (\bar{t}, \bar{x}) \in [0,T) \times G, \qquad \frac{1}{2\epsilon}\left((t_\epsilon - s_\epsilon)^2 + |x_\epsilon - y_\epsilon|^2\right) \to 0, \qquad M_\epsilon \to M.$$

Moreover, because the last two components of x_ϵ and y_ϵ are discrete, we must also have $x_\epsilon = (\bar{x}_1, z_\epsilon, \bar{x}_4, \bar{x}_5)$ and $y_\epsilon = (\bar{x}_1, z'_\epsilon, \bar{x}_4, \bar{x}_5)$ for some $z_\epsilon, z'_\epsilon \in \mathbf{R} \times \mathbf{R}_+$. So, by the Ishii Lemma (Theorem 2.3.1), there exist $(a_\epsilon, \frac{1}{\epsilon}(x_\epsilon - y_\epsilon), X_\epsilon) \in \overline{\mathscr{P}}^{2,+} u(t_\epsilon, x_\epsilon)$ and $(b_\epsilon, \frac{1}{\epsilon}(x_\epsilon - y_\epsilon), Y_\epsilon) \in \overline{\mathscr{P}}^{2,-} v(t_\epsilon, x_\epsilon)$ such that

$$a_\epsilon - b_\epsilon = 0.$$

3. We must have

$$\begin{cases} \min\left[-a_\epsilon + F(x_\epsilon, \epsilon^{-1}(\tilde{x}_\epsilon - \tilde{y}_\epsilon), \mathscr{I}[t_\epsilon, x_\epsilon, u]), u(t_\epsilon, x_\epsilon) - \mathscr{K} u(t_\epsilon, x_\epsilon)\right] \leqslant 0, \\ \min\left[-b_\epsilon + F(y_\epsilon, \epsilon^{-1}(\tilde{x}_\epsilon - \tilde{y}_\epsilon), \mathscr{I}[s_\epsilon, y_\epsilon, v]), v(s_\epsilon, y_\epsilon) - \mathscr{K} v(s_\epsilon, y_\epsilon)\right] \geqslant \delta\, C \end{cases}$$

We can assume that $u(t_\epsilon, x_\epsilon) - \mathscr{K} u(t_\epsilon, x_\epsilon) > 0$ for small $\epsilon > 0$. Indeed, suppose, to the contrary, that $u(t_\epsilon, x_\epsilon) - \mathscr{K} u(t_\epsilon, x_\epsilon) \leqslant 0$ for some infinite set of ϵ's converging to 0. Then since $v(s_\epsilon, y_\epsilon) - \mathscr{K} v(s_\epsilon, y_\epsilon) \geqslant \delta C$,

$$\begin{aligned} M &= \limsup_{\epsilon \downarrow 0} \left(u(t_\epsilon, x_\epsilon) - v(s_\epsilon, y_\epsilon) \right) \\ &\leqslant \limsup_{\epsilon \downarrow 0} \mathscr{K} u(t_\epsilon, x_\epsilon) - \liminf_{\epsilon \downarrow 0} \mathscr{K} v(s_\epsilon, y_\epsilon) - \delta C \\ &\leqslant \mathscr{K} u(\bar{t}, \bar{x}) - \mathscr{K} v(\bar{t}, \bar{x}) - \delta C \end{aligned}$$

where we have used the upper (resp., lower) semicontinuity of $\mathscr{K} u$ (resp., $\mathscr{K} v$). But since $\mathscr{K} u(\bar{t}, \bar{x}) = u(\bar{t}, \Gamma(\bar{x}, \bar{\zeta}))$ for some $\bar{\zeta} \in Z(\bar{t}, \bar{x})$,

$$\begin{aligned} M \leqslant u(\bar{t}, \Gamma(\bar{x}, \bar{\zeta})) - \mathscr{K} v(\bar{t}, \bar{x}) - \delta C &\leqslant u(\bar{t}, \Gamma(\bar{x}, \bar{\zeta})) - v(\bar{t}, \Gamma(\bar{x}, \bar{\zeta})) - \delta C \\ &\leqslant M - \delta C, \end{aligned}$$

which is a contradiction. So, we can assume that

$$\begin{cases} -a_\epsilon + F(x_\epsilon, \epsilon^{-1}(\tilde{x}_\epsilon - \tilde{y}_\epsilon), \mathscr{I}[t_\epsilon, x_\epsilon, u]) \leqslant 0, \\ -b_\epsilon + F(y_\epsilon, \epsilon^{-1}(\tilde{x}_\epsilon - \tilde{y}_\epsilon), \mathscr{I}[s_\epsilon, y_\epsilon, v]) \geqslant \delta C. \end{cases}$$

4. Since $a_\epsilon - b_\epsilon = 0$, we deduce that

$$\begin{aligned} \delta C &\leqslant F(y_\epsilon, \epsilon^{-1}(\tilde{x}_\epsilon - \tilde{y}_\epsilon), \mathscr{I}[s_\epsilon, y_\epsilon, v]) - F(x_\epsilon, \epsilon^{-1}(\tilde{x}_\epsilon - \tilde{y}_\epsilon), \mathscr{I}[t_\epsilon, x_\epsilon, u]) \\ &\leqslant o_\epsilon(1) + \mathscr{I}[t_\epsilon, x_\epsilon, u] - \mathscr{I}[s_\epsilon, y_\epsilon, v]. \end{aligned}$$

To complete the proof, it suffices to show that

$$\mathscr{I}[t_\epsilon, x_\epsilon, u] - \mathscr{I}[s_\epsilon, y_\epsilon, v] \leqslant o_\epsilon(1) \tag{4.18}$$

because this would yield $C \leqslant 0$, which is a contradiction.

Using the definition of the points (t_ϵ, x_ϵ) and (s_ϵ, y_ϵ) and the identity $|x + y|^2 = |x|^2 + 2\langle x, y \rangle + |y|^2$, we have

$$u(t_\epsilon, x_\epsilon + a) - u(t_\epsilon, x_\epsilon) \leqslant v(s_\epsilon, y_\epsilon + b) - v(s_\epsilon, y_\epsilon) + \frac{1}{\epsilon} \langle x_\epsilon - y_\epsilon, a - b \rangle + \frac{1}{2\epsilon} |a - b|^2$$

for any a, b such that $x_\epsilon + a, y_\epsilon + b \in S$. Hence, for any state-depending intensity $\theta\colon G \to \mathbf{R}_+$ that is continuous at $\bar{x}$, finite measure μ on $\mathbf{R}_+$, and function $g\colon G \times \{0, 1\} \times \mathbf{R}_+ \to G$

such that

$$\int_0^\infty |g(x,\ell,z)-g(y,\ell,z)|^p\,\mu(dz)\leqslant A|x-y|^p$$

and

$$\max_{\ell\in\{0,1\}}\int_0^\infty |g(x_0,\ell,z)|^p\,\mu(dz)<\infty$$

for all $p=1,2$, some $x_0\in S$, and a constant A which is independent of x,y,ℓ, we must have

$$\begin{aligned}
\theta(x_\epsilon)&\int_0^\infty \Big(u(t_\epsilon,x_\epsilon+g(x_\epsilon,\ell,z))-u(t_\epsilon,x_\epsilon)\Big)\,\mu(dz)\\
&-\theta(x_\epsilon)\int_0^\infty \Big(v(s_\epsilon,y_\epsilon+g(y_\epsilon,\ell,z))-v(s_\epsilon,y_\epsilon)\Big)\,\mu(dz)\\
&\leqslant \theta(x_\epsilon)\Big[\frac{1}{\epsilon}\int \langle x_\epsilon-y_\epsilon, g(x_\epsilon,\ell,z)-g(y_\epsilon,\ell,z)\rangle\,\mu(dz)\\
&+\frac{1}{2\epsilon}\int |g(x_\epsilon,\ell,z)-g(y_\epsilon,\ell,z)|^2\,\mu(dz)\Big]\\
&\leqslant \theta(x_\epsilon)\,\frac{3A}{2\epsilon}\,|x_\epsilon-y_\epsilon|^2=o_\epsilon(1).
\end{aligned}$$

This implies that

$$\begin{aligned}
\theta(x_\epsilon)&\int_0^\infty \Big(u(t_\epsilon,x_\epsilon+g(x_\epsilon,\ell,z))-u(t_\epsilon,x_\epsilon)\Big)\,\mu(dz)\\
&\leqslant \theta(y_\epsilon)\int_0^\infty \Big(v(s_\epsilon,y_\epsilon+g(y_\epsilon,\ell,z))-v(s_\epsilon,y_\epsilon)\Big)\,\mu(dz)\\
&+|\theta(x_\epsilon)-\theta(y_\epsilon)|\cdot\sup_{\ell\in\{0,1\}}\int_0^\infty |v(s_\epsilon,y_\epsilon+g(y_\epsilon,\ell,z))-v(s_\epsilon,y_\epsilon)|\,\mu(dz)+o_\epsilon(1)\\
&=\theta(y_\epsilon)\int_0^\infty \Big(v(s_\epsilon,y_\epsilon+g(y_\epsilon,\ell,z))-v(s_\epsilon,y_\epsilon)\Big)\,\mu(dz)+o_\epsilon(1)
\end{aligned}$$

where $o_\epsilon(1)$ is independent of ℓ. Using this inequality for $\mu=\vartheta^\pm$ and $\mu=\delta_L$ (the Dirac mass at L) and the corresponding intensity functions, we derive (4.18). The proof is complete.

□

Proof of Theorem 4.3.1. That u is a continuous viscosity solution of problem (4.9) can be proved by repeating the arguments in the proof of Theorem 5.3 in [67]. By Lemma 4.5.3, it must be the only viscosity solution satisfying growth condition (4.8). The existence of viscosity solution of problem (4.4) follows similarly.

Now let

$$\mathscr{M}^{\pm}u := \int \left[u(t,p,x+\ell z\,(\pm ip+1/2)\,,s,i,q\mp i\ell z)-u\right]\vartheta^{\pm}(dz),$$

$$\mathscr{N}^{\pm}u := u(t,p\pm i,x+\ell L\,(\pm ip+1/2)\,,0,\pm i,q\mp i\ell L)-u.$$

Then using the ansatz

$$u(t,p,x,s,i,q)=U(p,x,q)+w(t,s,i,q),$$

we obtain

$$\begin{aligned}
\mathscr{M}^{\pm}u &= \int_0^{\infty} w(t,s,i,q\mp i\ell z)\,\vartheta^{\pm}(dz)\\
&\quad+\int_0^{\infty}\left[\frac{1}{2}\ell z-\left(\frac{1}{2}+\epsilon\right)(|q\mp i\ell z|-|q|)\right]\vartheta^{\pm}(dz)-w(t,s,i,q),\\
\mathscr{N}^{\pm}u &= w(t,0,\pm i,q\mp i\ell L)\pm iq-\frac{L}{2}\ell-\left(\frac{1}{2}+\epsilon\right)(|q\mp i\ell L|-|q|)-w(t,s,i,q),\\
\mathscr{K}u-u &= \sup_{\substack{|\eta|\leqslant|q|\\ q+\eta\in\mathscr{Q}}}\left[w(t,s,i,q+\eta)-\left(\frac{1}{2}+\epsilon\right)(|q+\eta|+|\eta|-|q|)-\epsilon_0\right]-w(t,s,i,q).
\end{aligned}$$

Using the ansatz

$$w(t,s,i,q)=v(t,s,iq)$$

and taking into account that $i^2=1$, we obtain

$$\begin{aligned}
\mathscr{M}^{\pm}u &= \int_0^{\infty} v(t,s,\hat{q}\mp\ell z)\,\vartheta^{\pm}(dz)\\
&\quad+\int_0^{\infty}\left[\frac{1}{2}\ell z-\left(\frac{1}{2}+\epsilon\right)(|\hat{q}\mp\ell z|-|\hat{q}|)\right]\vartheta^{\pm}(dz)-v(t,s,\hat{q}),\\
\mathscr{N}^{\pm}u &= v(t,0,\pm\hat{q}-\ell L)-\frac{L}{2}\ell-\left(\frac{1}{2}+\epsilon\right)(|\hat{q}\mp\ell L|-|\hat{q}|)\pm\hat{q}-v(t,s,\hat{q}),\\
\mathscr{K}u-u &= \sup_{\substack{|\eta|\leqslant|\hat{q}|\\ \hat{q}+\eta\in\mathscr{Q}}}\left[v(t,s,\hat{q}+\eta)-\left(\frac{1}{2}+\epsilon\right)(|\hat{q}+\eta|+|\eta|-|\hat{q}|)-\epsilon_0\right]-v(t,s,\hat{q})
\end{aligned}$$

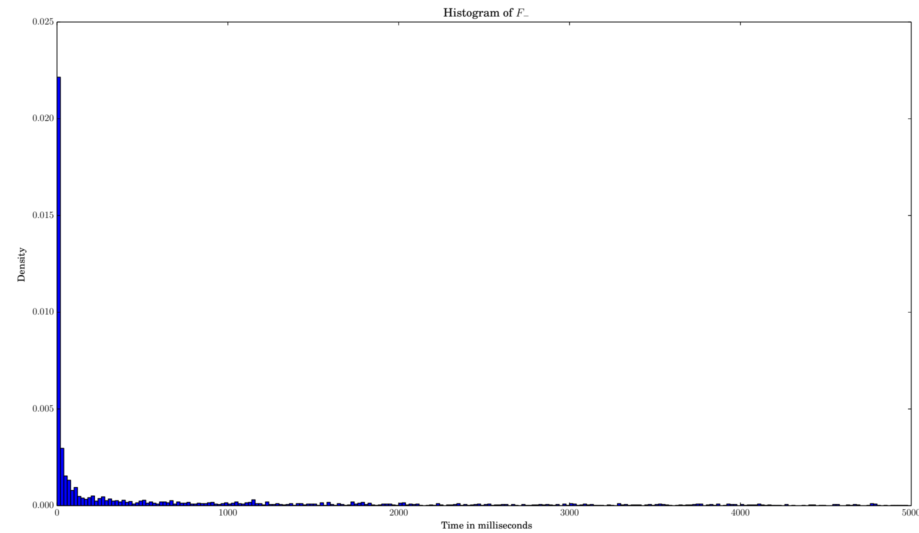
where $\hat{q}:=iq$.

The proof is complete.

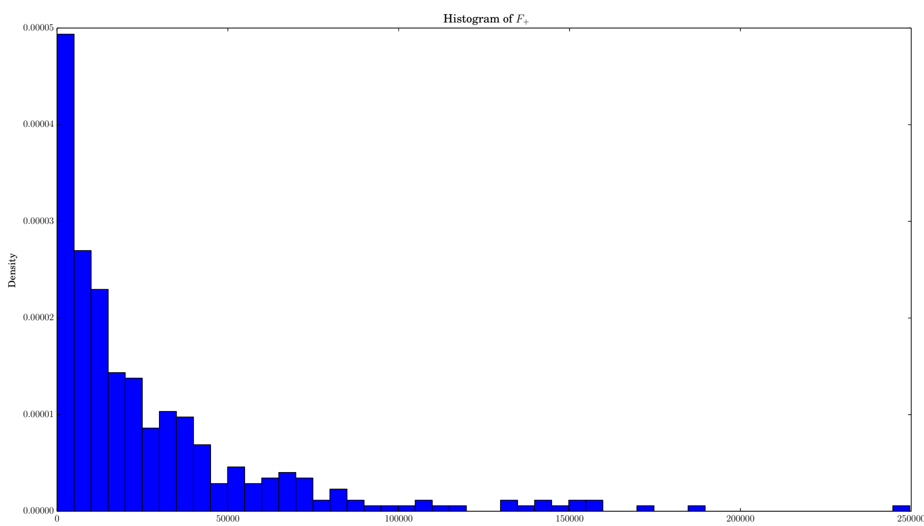
□

4.6 Parameter estimation

This section is devoted to statistical estimation of parameters introduced in the previous sections.

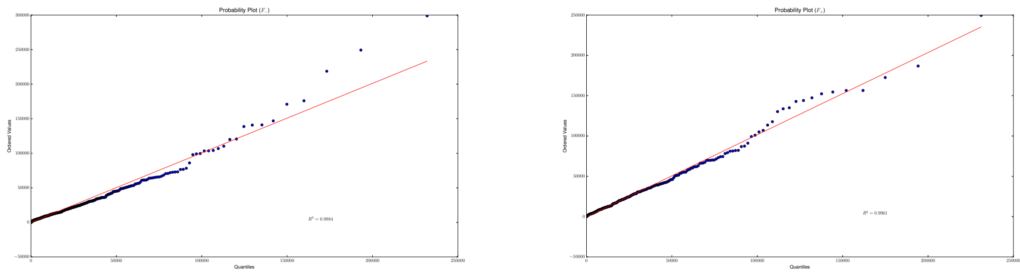


(a) Histogram of F_-



(b) Histogram of F_+

Figure 4.1: Histograms of $F_\pm$ for the CME ESU3 contract. Estimated on July 15, 2013 data.



(a) Probability plot for F_-

(b) Probability plot for F_+

Figure 4.2: Probability plots for $F_\pm$ for the CME ESU3 contract. Estimated on July 15, 2013 data.

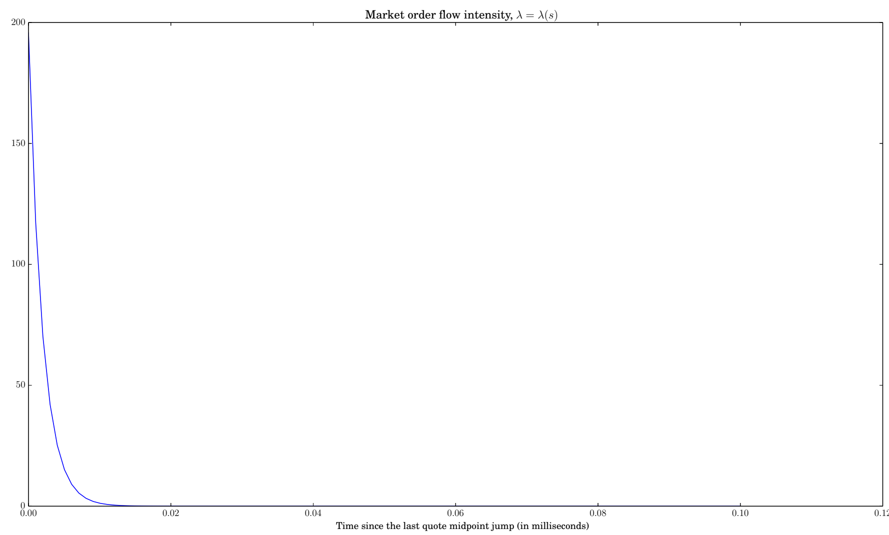


Figure 4.3: Market order flow intensity, $\lambda = \lambda(s)$ for the CME ESU3 contract. Estimated on July 15, 2013 data.

Midpoint jump intensity. The distributions $F_\pm$ can be very well approximated by Gamma distributions, which can be seen from probability plots 4.2a and 4.2b. The corresponding coefficients can be estimated as follows. Let the conditional random variables $\Delta_n \mid I_n I_{n-1} = \pm 1$ have Gamma distributions $\Gamma(\beta_\pm, \theta_\pm)$ with shapes $\beta_\pm$ and scales $\theta_\pm$. Since the mean and the variance of $\Gamma(\beta, \theta)$ are equal to $\beta\theta$ and $\beta\theta^2$, we can estimate

$$\hat{\theta}_\pm = \frac{\frac{1}{|\mathscr{Z}_\pm|}\sum_{n\in\mathscr{Z}_\pm}(\Delta_n - \bar{\Delta}_\pm)^2}{\bar{\Delta}_\pm}, \qquad \hat{\beta}_\pm = \frac{\bar{\Delta}_\pm}{\hat{\theta}_\pm}$$

where $\mathscr{Z}_\pm$ denotes the subset of indices n for which $I_n I_{n-1} = \pm 1$ and

$$\bar{\Delta}_\pm := \frac{1}{|\mathscr{Z}_\pm|} \sum_{n \in \mathscr{Z}_\pm} \Delta_n.$$

Under the Gamma distribution assumption,

$$j^\pm(s) \equiv \frac{f_\pm(s)}{1 - F_\pm(s)} = \frac{1}{\theta_\pm} \frac{(s/\theta_\pm)^{\beta-1} e^{-s/\theta_\pm}}{\Gamma(\theta_\pm) - \gamma_0(\beta_\pm, s/\theta_\pm)}, \tag{4.19}$$

where $\gamma_0(\cdot,\cdot)$ is the lower incomplete Gamma function. This intensity function is decreasing if and only if $\beta < 1$[3]. For such values of β, $j^\pm$ are unbounded at the origin. So, in order to make j bounded, one needs to modify $j^\pm$ in a right neighborhood $[0, \epsilon)$ of 0. The most simple way to accomplish that is to let $f_\pm$ be a linear function on $[0, \epsilon)$.

Alternately, one could also perform a nonparametric estimation of $f_\pm$ and $j^\pm$. To do that, fix a kernel $K\colon \mathbf{R} \to \mathbf{R}_+$, i.e. a nonnegative even function with $\int K(s)\,ds = 1$. Then the densities $f_\pm$ of the conditional inter-jump times can be estimated by the standard kernel density estimation method:

$$\hat{f}_\pm(s) = \frac{1}{|\mathscr{Z}_\pm|} \sum_{n \in \mathscr{Z}_\pm} K_h(s - \Delta_n)$$

where $K_h(s) := h^{-1} K(s/h)$ is the smoothing scaled kernel with bandwidth $h > 0$. We will assume that K is the Gaussian density with mean 0 and variance h^2.

The conditional jump intensities

$$\begin{aligned} j^\pm(s) &= \lim_{\delta \downarrow 0} \frac{1}{\delta} \frac{\mathsf{P}[s \leqslant \Delta_k < s + \delta, J_k J_{k-1} = \pm 1 \mid J_{k-1} = i]}{\mathsf{P}[\Delta_k \geqslant s \mid J_{k-1} = i]} \\ &= \lim_{\delta \downarrow 0} \frac{1}{\delta} \frac{\mathsf{P}[s \leqslant \Delta_k < s + \delta, J_k J_{k-1} = \pm 1]}{\mathsf{P}[\Delta_k \geqslant s]}. \end{aligned}$$

Therefore, we can use the estimate

$$\hat{j}^\pm(s) = \frac{1}{|\{n \mid \Delta_n \geqslant s\}|} \sum_{k \in \mathscr{Z}_\pm} K_h(s - \Delta_k).$$

---

[3]Indeed, let $g(x) := x^{\beta-1} e^{-x} / (\Gamma(\beta) - \Gamma_x(\beta))$. Then $g'(x) < 0$ is equivalent to $h(x) := (\beta - 1 - x)(\Gamma(\beta) - \Gamma_x(\beta)) + x^\beta e^{-x} < 0$. Proceeding to the limit as $x \downarrow 0$, we deduce that $\beta < 1$ is a necessary condition. Now let $\beta < 1$. Then

$$h(x) \leqslant (\beta - 1) \int_x^\infty u^{\beta-1} e^{-u}\,du - x^\beta \int_x^\infty e^{-u}\,du + x^\beta e^{-x} = (\beta - 1) \int_x^\infty u^{\beta-1} e^{-u}\,du < 0,$$

implying that $g' < 0$ and hence $\beta < 1$ is a sufficient condition.

Finally, since $\alpha = \mathsf{E}[J_n J_{n-1}]$, the correlation parameter can be estimated as

$$\hat{\alpha} = \frac{1}{N} \sum_{n=1}^{N} J_n J_{n-1}$$

where N is the sample size.

Trade intensity. We start with an estimation procedure for the market order flow intensity, $\lambda = \lambda(S_t)$. By Proposition 7.2.III in [28], the log-likelihood function for M over $[0, T]$ has the form

$$L_T = \int_0^T \log \lambda(S_t)\, dM_t - \int_0^T \lambda(S_t)\, dt.$$

Since

$$\begin{aligned} \int_0^T \lambda(S_t)\, dt &= \sum_{k=1}^{N_T} \int_{T_{k-1}}^{T_k} \lambda(t - T_{k-1})\, dt + \int_{T_{N_T}}^{T} \lambda(t - T_{N_T})\, dt \\ &= \sum_{k=1}^{N_T} \int_0^{\Delta_k} \lambda(t)\, dt + \int_0^{T - T_{N_T}} \lambda(t)\, dt, \end{aligned}$$

we must have

$$L_T = \sum_{k=1}^{M_T} \log \lambda(S_{\theta_k -}) - \sum_{k=1}^{N_T} \int_0^{\Delta_k} \lambda(t)\, dt - \int_0^{T - T_{N_T}} \lambda(t)\, dt.$$

Take $T = T_{N_T}$. Then in view of (4.1),

$$L_T = \sum_{j=1}^{M_T} \log\left(\lambda_0 + A e^{-k S_{\theta_j -}}\right) - \lambda_0 T - \frac{A}{k}\left(N_T - \sum_{j=1}^{N_T} e^{-k\Delta_j}\right).$$

The parameter (λ_0, k, A) can now be estimated by any standard iterative method. Figure 4.3 shows a plot of the market order flow intensity function estimated with the July 15, 2013 data for the CME ESU3 future contract.

Finally, since $\rho = \mathsf{E}[Z_n I_{\theta_n -}]$, we estimate it as

$$\hat{\rho} = \frac{1}{N} \sum_{n=1}^{N} Z_n I_{\theta_n -}.$$

List of main notations

Chapter 3

Δ_p	p-Laplacian operator, $\Delta_p u := \operatorname{div}(\lvert\nabla u\rvert^{p-2}\nabla u)$
$W_0^{1,p}(\Omega)$	space of functions $u \in W^{1,p}(\Omega)$ with zero trace on $\partial\Omega$ (Definition 2.1.1)
$C^{k,\mu}(\bar{\Omega})$	Hölder space of functions $u \in C^2(\bar{\Omega})$ such that the derivatives $\{D^\gamma u\}_{\lvert\gamma\rvert \leqslant k}$ are Lipschitz with order $\mu \in (0,1)$
p^*	critical Sobolev exponent defined by $\frac{1}{p^*} = \frac{1}{p} - \frac{1}{N}$ where N is the dimension of the ambient space
B_ρ	closed ball of radius ρ in $W_0^{1,p}(\Omega)$ centered at 0
S_ρ	sphere of radius ρ in $W_0^{1,p}(\Omega)$ centered at 0

Chapter 4

$\mathscr{P}([0,T] \times G)$	space of functions of polynomial growth (Definition 4.5.1)
P_t	quote midpoint process
X_t	cash process
q_t	inventory process
N_t	counting process of midpoint jumps
M_t	counting process of market order arrivals
S_t	time elapsed since the last midpoint jump
I_t	last quote midpoint jump
$j(s)$	intensity of midpoint jumps
$\lambda(s)$	intensity of market order arrivals

References

[1] R. Adams, *Sobolev spaces.* New York-London: Academic Press, 1975.

[2] R. Almgren and N. Chriss, "Optimal execution of portfolio transactions," *J. Risk*, vol. 3, pp. 5–39, 2000.

[3] A. Ambrosetti, H. Brezis, and G. Cerami, "Combined effects of concave and convex nonlinearities in some elliptic problems," *J. Funct. Anal.*, vol. 122, pp. 519–543, 1994.

[4] A. Ambrosetti, A. Garcia, and I. Peral, "Multiplicity results for some nonlinear elliptic equations," *J. Funct. Anal.*, vol. 137, pp. 219–242, 1996.

[5] A. Ambrosetti and P. H. Rabinowitz, "Dual variational methods in critical point theory and applications," *J. Functional Analysis*, vol. 14, pp. 349–381, 1973.

[6] M. Arisawa, "A new definition of viscosity solutions for a class of second-order degenerate elliptic integro-differential equations," *Ann. Inst. H. Poincaré*, vol. 23, pp. 695–711, 2006.

[7] ——, "A remark on the definitions of viscosity solutions for the integro-differential equations with Lévy operators," *J. Math. Pures Appl.*, vol. 89, pp. 567–574, 2008.

[8] J.-P. Aubin and I. Ekeland, *Applied nonlinear analysis.* Dover Publications, 2006.

[9] M. Avellaneda and S. Stoikov, "High-frequency trading in a limit order book," *Quant. Fin.*, vol. 8, pp. 217–224, 2008.

[10] A. Babin and A. Figotin, "Nonlinear Maxwell equations in inhomogeneous media," *Commun. Math. Phys.*, vol. 241, pp. 519–581, 2003.

[11] E. Bacrya and J.-F. Muzy, "Hawkes model for price and trades high-frequency dynamics," *Quant. Fin.*, vol. 14, pp. 1147–1166, 2014.

[12] D. Bai and Y. Chen, "Three positive solutions for a generalized Laplacian boundary value problem with a parameter," *Appl. Math. Comput.*, vol. 219, pp. 4782–4788, 2013.

[13] M. Baron, J. Brogaard, and A. Kirilenko, "Risk and return in high frequency trading," 2014, available at SSRN: http://ssrn.com/abstract=2433118.

[14] E. Bayraktar and M. Ludkovski, "Optimal trade execution in illiquid markets," *Math. Finance*, vol. 21, pp. 681–701, 2011.

[15] ——, "Liquidation in limit order books with controlled intensity," *Math. Finance*, vol. 24, pp. 627–650, 2014.

[16] J. P. Bouchaud, J. D. Farmer, and F. Lillo, "How markets slowly digest changes in supply and demand," in *Handbook of Financial Markets: Dynamics and Evolution.* North Holland, 2009, pp. 57–160.

[17] Y. Bozhkov and E. Mitidieri, "Existence of multiple solutions for quasilinear systems via fibering method," *J. Differential Equations*, vol. 190, pp. 239–267, 2003.

[18] H. Brezis, *Functional analysis, Sobolev spaces and partial differential equations.* Springer, 2011.

[19] H. Brezis and E. Lieb, "A relation between pointwise convergence of functions and convergence of functionals," *Proc. Amer. Math. Soc.*, vol. 88, pp. 486–490, 1983.

[20] Á. Cartea, R. Donnelly, and S. Jaimungal, "Algorithmic trading with model uncertainty," 2014, available at SSRN: http://ssrn.com/abstract=2310645.

[21] Á. Cartea and S. Jaimungal, "Risk metrics and fine-tuning of high frequency trading strategies," *Math. Finance*, 2013, doi: 10.1111/mafi.12023.

[22] Á. Cartea, S. Jaimungal, and J. Ricci, "Buy low, sell high: a high frequency trading perspective," *SIAM J. Financial Math.*, vol. 5, pp. 415–444, 2014.

[23] K. C. Chang, *Infinite Dimensional Morse Theory and Multiple Solution Problems*, ser. Progress in Nonlinear Differential Equations and Their Applications. Birkhauser, 1993.

[24] ——, *Methods in Nonlinear Analysis*, ser. Springer Monographs in Mathematics. Springer, 2005.

[25] M. Crandall, H. Ishii, and P.-L. Lions, "User's guide to viscosity solutions of second order partial differential equations," *Bull. Amer. Math. Soc.*, vol. 27, pp. 1–67, 1992.

[26] M. Crandall and P. Rabinowitz, "Bifurcation from simple eigenvalues," *J. Func. Anal.*, vol. 8, pp. 321–340, 1971.

[27] ——, "Bifurcation, perturbation of simple eigenvalues and linearized stability," *Arch. Rational Mech. Anal.*, vol. 52, pp. 161–180, 1973.

[28] D. J. Daley and D. V. Jones, *An introduction to the theory of point processes*, 2nd ed. Springer Verlag, 2002, vol. 1.

[29] S. de Valeriola, "On some quasilinear critical problems," *Adv. Nonlinear Stud.*, vol. 9, pp. 825–836, 2009.

[30] I. Ekeland and R. Temam, *Convex analysis and variational problems.* Society for Industrial and Applied Mathematics, 1987.

[31] J. F. Escobar, "Conformal deformation of a Riemannian metric to a scalar flat metric with constant mean curvature on the boundary," *Ann. of Math.*, vol. 136, pp. 1–50, 1992.

[32] L. Evans, *Partial Differential Equations.* American Mathematical Society, 1998.

[33] P. Fodra and H. Pham, “High frequency trading in a Markov renewal model,” 2013, available at SSRN: http://ssrn.com/abstract=2333752.

[34] J. D. Fonseca and R. Zaatour, “Hawkes process: fast calibration, application to trade clustering, and diffusive limit,” *Journal of Futures Markets*, vol. 34, pp. 548–579, 2014.

[35] A. Fruth, T. Schöneborn, and M. Urusov, “Optimal trade execution and price manipulation in order books with time-varying liquidity,” *Math. Finance*, 2013, doi: 10.1111/mafi.12022.

[36] C. I. G. Barles, “Second-order elliptic integro-differential equations: viscosity solutions’ theory revisited,” *Ann. I. H. Poincaré*, vol. 25, pp. 567–585, 2008.

[37] O. Guéant, C.-A. Lehalle, and J. Fernandez-Tapia, “Dealing with the inventory risk: a solution to the market making problem,” *Math. Finan. Econ.*, vol. 7, pp. 477–507, 2013.

[38] O. Guéant and G. Royer, “VWAP execution and guaranteed VWAP,” *SIAM J. Financial Math.*, vol. 5, pp. 445–471, 2014.

[39] O. Gueéant, C.-A. Lehalle, and J. Fernandez-Tapia, “Optimal portfolio liquidation with limit orders,” *SIAM J. Finan. Math.*, vol. 3, pp. 740–764, 2012.

[40] F. Guilbaud and H. Pham, “Optimal high-frequency trading in a pro rata microstructure with predictive information,” *Math. Finance*, 2013, doi: 10.1111/mafi.12042.

[41] ——, “Optimal high-frequency trading with limit and market orders,” *Quant. Fin.*, vol. 13, pp. 79–94, 2013.

[42] T. Ho and H. R. Stoll, “Optimal pricing under transactions and return uncertainty,” *J. Fin. Econ.*, vol. 9, pp. 47–73, 1981.

[43] R. Huitema, “Optimal portfolio execution using market and limit orders,” 2014, available at SSRN: http://ssrn.com/abstract=1977553.

[44] K.-C. Hung, “Exact multiplicity of positive solutions of a semipositone problem with concave-convex nonlinearity,” *J. Differential Equations*, vol. 255, pp. 3811–3831, 2013.

[45] Y. Il’yasov, “On nonlocal existence results for elliptic equations with convex-concave nonlinearities,” *Nonlinear Anal.*, vol. 61, pp. 211–236, 2005.

[46] ——, “Bifurcation calculus by the extended functional method,” *Func. Anal. and Its Appl.*, vol. 41, pp. 18–30, 2007.

[47] D. A. Kandilakis and A. N. Lyberopoulos, “Multiplicity of positive solutions for some quasilinear Dirichlet problems on bounded domains in $\mathbf{R}^n$,” *Commentationes Mathematicae Universitatis Carolinae*, vol. 44, pp. 645–658, 2003.

[48] J. L. Kazdan and F. W. Warner, “Scalar curvature and conformal deformation of Riemannian structure,” *J. Differential Geometry*, vol. 10, pp. 113–134, 1975.

[49] J. B. Keller and S. Antman, Eds., *Bifurcation theory and nonlinear eigenvalue problems.* Courant Institute of Mathematical Sciences, 1967.

[50] S. T. Kyritsi and N. S. Papageorgiou, "Pairs of positive solutions for p-Laplacian equations with combined nonlinearities," *Commun. Pure and Appl. Anal.*, vol. 8, pp. 1031–1051, 2009.

[51] S. Li, S. Wu, and H.-S. Zhou, "Solutions to semilinear elliptic problems with combined nonlinearities," *J. Diff. Equat.*, vol. 185, pp. 200–224, 2002.

[52] J.-L. Lions, *Control of distributed singular systems.* Gauthier-Villars, 1985.

[53] P. L. Lions, "On the existence of positive solutions of semilinear elliptic equations," *SIAM*, vol. 24, pp. 441–467, 1982.

[54] V. Lubyshev, "Precise range of the existence of positive solutions of a nonlinear, indefinite in sign Neumann problem," *Commun. Pure Appl. Anal.*, vol. 8, pp. 999–1018, 2009.

[55] ——, "Multiple solutions of an even-order nonlinear problem with convex-concave nonlinearity," *Nonlinear Anal.*, vol. 74, pp. 1345–1354, 2011.

[56] T. Ouyang, "On the positive solutions of semilinear equations $\Delta u + \lambda u + hu^p = 0$ on compact manifolds. Part II," *Indiana Univ. Math. J.*, vol. 40, pp. 1083–1141, 1991.

[57] S. I. Pohozaev, "On one approach to nonlinear equations," *Dokl. Akad. Nauk SSSR*, vol. 247, pp. 1327–1331, 1979, (in Russian) (vol. 20, pp. 912-916, 1979 (in English)).

[58] ——, "On a constructive method in calculus of variations," *Dokl. Akad. Nauk SSSR*, vol. 298, pp. 1330–1333, 1988, (in Russian) (vol. 37, pp. 274-277, 1988 (in English).

[59] ——, "On fibering method for the solutions of nonlinear boundary value problems," *Trudy Mat. Inst. Steklov.*, vol. 192, pp. 146–163, 1990, (in Russian).

[60] ——, *Nonlinear variational problems via the fibering method*, ser. Handbook of Differential Equations. Stationary Partial Differential Equations. Amsterdam: Elsevier, 2008, vol. 5, pp. 49–209.

[61] E. Stein, *Singular integrals and differentiability properties of functions.* Princeton University Press, 1970.

[62] M. Struwe, *Variational Methods: Applications to Nonlinear Partial Differential Equations and Hamiltonian Systems.* Springer, 2008.

[63] K. Taira, "The Yamabe problem and nonlinear boundary value problems," *J. Differential Equations*, vol. 122, pp. 316–372, 1995.

[64] M. Tang, "Exact multiplicity for semilinear elliptic Dirichlet problems involving concave and convex nonlinearities," *Proc. Roy. Soc. Edinburgh Sect. A*, vol. 133, pp. 705–717, 2003.

[65] K. Umezu, "Behavior and stability of positive solutions of nonlinear elliptic boundary value problems arising in population dynamics," *Nonlinear Anal.*, vol. 49, pp. 817–840, 2002.

[66] M. M. Vainberg, *Variational methods for the study of nonlinear operators.* Holden-Day, 1964.

[67] V. L. Vath, M. Mnif, and H. Pham, "A model of optimal portfolio selection under liquidity risk and price impact," *Finance Stoch.*, vol. 11, pp. 51–90, 2007.

[68] H. Yamabe, "On a deformation of riemannian structures on compact manifolds," *Osaka Math J.*, vol. 12, pp. 21–37, 1960.

[69] S.-T. Yau, "On the Ricci curvature of a compact Kähler manifold and the complex monge-ampre equation. I," *Comm. Pure Appl. Math.*, vol. 31, pp. 339–411, 1978.

[70] L. Zhao, Y. He, and P. Zhao, "The existence of three positive solutions of a singular p-Laplacian problem," *Nonlinear Anal.*, vol. 74, pp. 5745–5753, 2011.