May 2019

A Stochastic Control Model for Electricity Producers

Charles William Beer
University of Wisconsin-Milwaukee

A Stochastic Control Model for Electricity Producers

by

Charles Beer

A Dissertation Submitted in
Partial Fulfillment of the
Requirements for the Degree of

Doctor of Philosophy
in Mathematics

at

The University of Wisconsin-Milwaukee
May 2019

ABSTRACT

A Stochastic Control Model for Electricity Producers

by

Charles Beer

The University of Wisconsin-Milwaukee, 2019
Under the Supervision of Professor Richard Stockbridge

Modern electricity pricing models include a strong reversion to a long run mean and a number of non-local operators to encapsulate the discontinuous price behavior observed in such markets. However, incorporating non-local processes into a stochastic control problem presents significant analytical challenges. The motivation for this work is to solve the problem of optimal control of the burn rate for a coal-powered electricity plant. We first construct a pricng model that is a good general representative of the class of models currently used for electricity pricing as well as a model for the supply of fuel to the plant. Under this model, we state the control problem of maximizing the expected discounted revenue until the first time at which the plant runs out of fuel. Deriving the HJB equation for this control problem results in a partial integro-differential equation, which does not fit the classical theory of viscosity solutions. Building off of work by Barles and Imbert on viscosity solutions for non-local processes, we extend their theory to apply to non-local processes which also include a mean-reversion component. We first show that the value function for the control problem is a solution to this HJB equation. In our main result, we prove a comparison principle for viscosity solutions which uses a slightly more regular structure of the non-local operators to relax some of the assumptions of Barles and Imbert. Using this comparison principle, we are able to show that the value function is in fact the unique solution to the HJB equation. Thus, we have the desired result that solving the HJB equation is equivalent to solving the control problem, giving us a direct method for finding the optimal control policy for the electricity producer.

Table of Contents

List of Figures

ACKNOWLEDGEMENTS

First of all I would like to thank my adviser, Professor Richard Stockbridge, whose guidance and encouragement throughout my graduate career at UWM have been invaluable. Without Professor Stockbridge's involvement, this research would not exist, and I would not have been able to complete my graduate education. Further, Professor Chao Zhu's instruction, and his direct input on this paper, have been of great value.

Professor Anthony Ross, founding director of the Supply Chain Management Institute in the Lubar School of Business, deserves special recognition for his contributions to this work. Professor Ross brought the problem of optimal control for a coal plant to our attention, starting the process that grew into this thesis. He also supported this research fully for one year and partially for another, allowing the time for significant advancements in the early stages.

I would also like to acknowledge the contributions of some the many researchers whose work inspired this paper. Huyên Pham's clear and precise presentation in his book on stochastic control inspired and directed my work showing that the value function is a viscosity solution to the HJB equation. Guy Barles' and Cyril Imbert's novel work on viscosity solutions for PIDEs, without which this disseration would not have been possible, formed the bedrock for my main comparison principle result. Not least of all, Oliver Meister's MS thesis fitting parameters and verifying that our pricing model was applicable to real market data gave an excellent tangible connection to the real world and motivation to our work.

Finally, I would like to thank some of the many people whose support outside of academia gave me the ability to focus on my research and got me here today. The many friends, family, colleagues and instructors whose impact shows in every page would take pages to list, but I would especially like to acknowledge three people. My wife, Ashleigh Koenigs, whose support and love was entirely undiminished by her complete lack of interest in this subject area, and most of all, my parents, Dorothy and Robert Beer. I could not begin to list the ways in which they have worked to give me every one of the opportunities that have led to my PhD, and I will be forever grateful to them for that among innumerable other things.

Chapter 1

Introduction and Problem Formulation

1 Background

Beginning with some northern European markets in the early 1990's, nations around the world have liberalized the markets for electricity, and later those for other related products like energy futures and options [4]. In the United States and Canada, this process began in 1996 with the creation of Independent System Operators (ISOs) with the responsibilities of operating electricity grids and administering wholesale electricity markets for large, multi-state regions of the United States and Canada. These ISOs allow large energy producers to act as sellers in a partially regulated commodities market.

The ISO actually operates two separate markets, a day-ahead market and a real-time market. The day-ahead market allows producers to plan their production in a relatively deterministic manner since they are guaranteed a certain price. Producers are also obligated to provide a certain amount of electricity, based on their production capacity, to the day-ahead market in order to participate in the ISO market. However, since this market is largely deterministically controlled by the ISO, it is of little interest in this thesis. Of much greater interest is the real-time market operated alongside the day-ahead. In particular, we consider the nodes of the Midcontinent ISO (MISO) real-time market which provides service to the Midwestern US and Manitoba, Canada [11]. We pay particular attention to those nodes near Milwaukee, WI.

Modeling spot prices for electricity in these markets presents a challenge. While most commodities markets have been modeled with great accuracy using standard financial mathematical models driven by Brownian motion or exponential Brownian motion, these are continuous-path processes and thus fail to capture the large, instantaneous spikes seen in electricity prices.

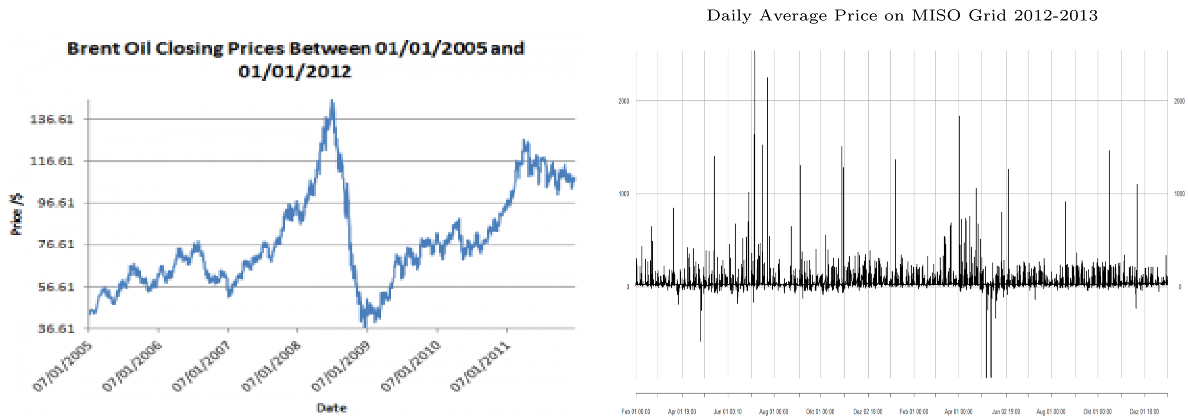

Figure 1: Crude oil spot price 2008-2016 [8] (left) and electricity spot price on one node of MISO grid [10] (right)

More sophisticated spot price models based on general Ornstein-Uhlenbeck processes were developed in the late 2000s (e.g. [3]) and compiled into a textbook on the subject by Fred Espen Benth in 2008 [4]. While these models contributed greatly to the pricing of energies derivatives like electricty futures and options, there has been little to no application of them to stochastic control problems, for example optimal action of an energy supplier in such a market. This thesis seeks to begin the examination of such applications through the particular case of a coal-powered electricty plant acting in an open electrictiy market.

2 Description of the Problem

The problem considered in this paper is to optimally control the rate at which a coal plant burns fuel in order to maximize its revenue from selling to a single node on the MISO grid. Because the costs of shutting down and restarting such a plant are extremely high, we wish to consider this as a first-exit problem with the terminal time being the first time the plant's coal supply reaches a specified minimum level.

3 The Model

3.1 The Coal Supply Process

Many coal plants have contracts with coal suppliers that specify an amount of coal to be delivered to the plant via freight train each day. However, several environmental factors including blockages of railways and failures of unloading equipment at the delivery point make it necessary to model this supply as a stochastic process rather than a deterministic one. Also, some minor delays may result in two trains (e.g. a train delayed from the previous day and the train intended to arrive on the current day) arriving in a single day. Further, since many of these delays stretch over multiple days, the arrival of coal each day cannot be considered completely independent of earlier days.

However, a Markovian supply process is needed for most analyses. So, the arrival of coal is modeled as a two coordinate Markov chain. Each element of the Markov chain has one coordinate indicating the number of coal trains arriving that day and a second coordinate that tracks the number of coal trains that arrived the previous day. In this way, some dependence on earlier information can be included while still creating a Markov process for the coal arrivals. Based on empirical data, we consider 9 such states s_1=(0,0), s_2=(0,1), s_3=(0,2), s_4=(1,0), s_5=(1,1), s_6=(1,2), s_7=(2,0), s_8=(2,1), and s_9=(2,2) (i.e. being in state s_6 means that one train arrived on the present day and two trains arrived on the previous day). Note that many transitions, e.g (0,0)→(1,1), are impossible since the second coordinate of the present state must match the first coordinate of the previous state. So, the transition probability matrix can be greatly simplified by making all such transition probabilities zero. We can then write the transition probability matrix, P, of the Markov chain as

$$P := \begin{pmatrix} p_{1,1} & 0 & 0 & p_{1,4} & 0 & 0 & p_{1,7} & 0 & 0 \\ p_{2,1} & 0 & 0 & p_{2,4} & 0 & 0 & p_{2,7} & 0 & 0 \\ p_{3,1} & 0 & 0 & p_{3,4} & 0 & 0 & p_{3,7} & 0 & 0 \\ 0 & p_{4,2} & 0 & 0 & p_{4,5} & 0 & 0 & p_{4,8} & 0 \\ 0 & p_{5,2} & 0 & 0 & p_{5,5} & 0 & 0 & p_{5,8} & 0 \\ 0 & p_{6,2} & 0 & 0 & p_{6,5} & 0 & 0 & p_{6,8} & 0 \\ 0 & 0 & p_{7,3} & 0 & 0 & p_{7,6} & 0 & 0 & p_{7,9} \\ 0 & 0 & p_{8,3} & 0 & 0 & p_{8,6} & 0 & 0 & p_{8,9} \\ 0 & 0 & p_{9,3} & 0 & 0 & p_{9,6} & 0 & 0 & p_{9,9} \end{pmatrix}$$

where these transition probabilities $p_{i,j}$ can be obtained empirically from records of coal shipment arrivals.

Since the problem we are considering is stated in a continuous-time framework, we instead think of this coal arrival process as a continuous-time Markov chain with state transition probabilities $p_{i,j}$ as above and intensity α such that there is a mean rate of one transition per day; that is a continuous-time Markov chain with transition rate matrix $Q := \alpha P$. Under the assumption that this transition matrix is irreducible (which, from empirical data, may require elimination of the last row and column since it is possible that the state $(2,2)$ is never reached) and noting that it is by construction positive recurrent, this continuous-time Markov chain will have unique stationary distribution $(\tilde{p}_1, \tilde{p}_2, \tilde{p}_3, \tilde{p}_4, \tilde{p}_5, \tilde{p}_6, \tilde{p}_7, \tilde{p}_8, \tilde{p}_9)$. We in fact only need the number of trains arriving currently at any given time t, that is the value of the first coordinate of the Markov chain's state. So, we define probabilities $p_0 := \tilde{p}_1 + \tilde{p}_2 + \tilde{p}_3$, $p_1 := \tilde{p}_4 + \tilde{p}_5 + \tilde{p}_6$, and $p_2 := \tilde{p}_7 + \tilde{p}_8 + \tilde{p}_9$, these being the probabilities that the current number of arrivals "today" is 0, 1, or 2, respectively. Further, the arrival of each coal shipment adds a constant amount $\tilde{\zeta}$ to the total available coal supply. Since coal can be unloaded and added to the supply faster than the maximum burn rate, we consider this addition to the supply to occur instantaneously upon arrival of a train. However, the coal supply is limited by the storage capacity of the plant to be below z_{max}. So, we define for $z \in [z_{min}, z_{max}]$, $\zeta := \tilde{\zeta} \wedge (z_{max} - z)$ and $2\zeta := 2\tilde{\zeta} \wedge (z_{max} - z)$. Then, for any bounded function f, the generator, Q, of this continuous-time Markov chain is given by

$$\begin{aligned} Qf(z) &= \lim_{h \to 0} \frac{\mathbb{E}\left[f(Z(h)) - f(Z(0)) \mid Z(0) = z\right]}{h} \\ &= \alpha \left(p_0[f(z) - f(z+0)] + p_1[f(z+\zeta) - f(z)] + p_2[f(z+2\zeta) - f(z)]\right) \\ &= \alpha \left(p_1[f(z+\zeta) - f(z)] + p_2[f(z+2\zeta) - f(z)]\right). \end{aligned} \tag{1.1}$$

Additionally, our control for this problem is the rate at which coal is being burned, $u(t)$. This rate is limited by the physical capacity of the plant such that $u(t) \in [u_{min}, u_{max}]$ for all $t \geq 0$. We further assume that u is a non-anticipating control. Since this control represents the rate at which coal is being used, it imposes a drift of $-u(t)$ on the process $Z(t)$. The coal supply is limited by the physical storage capacity of the plant such that $Z(t) \in [z_{min}, z_{max}] \subset \mathbb{R}^+$ for all $t \geq 0$. The i^{th} transition of the continuous-time Markov chain results in an addition of ξ_i to the coal supply where the ξ_i are i.i.d. random variables where

$$\mathbb{P}[\xi_i = x] = \begin{cases} p_0 \;, & \text{for } x = 0 \\ p_1 \;, & \text{for } x = \zeta \\ p_2 \;, & \text{for } x = 2\zeta \end{cases} \tag{1.2}$$

for all i. Letting ψ_i be the arrival time of the i^{th} shipment, we can express $Z(t)$ explicitly in the form

$$Z(t) = \min\left\{ Z(0) - \int_0^t u(t)\,dt + \sum_{i=1}^{\infty} \xi_i I_{\{\psi_i \le t\}} \ ,\ z_{max} \right\} \tag{1.3}$$

An inherent assumption of this problem is that the plant must be shut down the moment the coal supply reaches the level z_{min}. Therefore, the Z process terminates at the stopping time $\tau := \min\{t \ge 0 \,|\, Z(t) = z_{min}\}$. Note that since the singular behavior of this process is only in coal arrivals, any downward change in the coal supply will occur continuously due to the continuous drift rate $u(t)$. So, this minimum will exist, and therefore the terminal time τ is well-defined.

3.2 The Spot Price Process

The form of the spot price model is the same as that used by Gonzalez, Moriarty, and Palczewski [7], which is a specific form of the general model developed by Benth [4]. This multifactor model includes three components. The first is a Gaussian Ornstein-Uhlenbeck process which is the solution to the SDE

$$dY_0(t) = \frac{1}{\lambda}(\mu - Y_0(t))dt + \sigma dW(t) \quad , \quad Y_0(0) = y_0 \tag{1.4}$$

with $W(t)$ being a one-dimensional standard Brownian motion, $Y_0(0) = y_0$ being the spot price at the initial time, and μ being the long-term mean price in the market. This results in the explicit form

$$Y_0(t) = \mu + (y_0 - \mu)e^{-\frac{1}{\lambda}t} + \int_0^t e^{-\frac{1}{\lambda}(t-s)}\sigma dW(s). \tag{1.5}$$

This process is a mean-reverting Brownian motion which reverts exponentially towards the mean price μ at exponential rate $\frac{1}{\lambda}$.

The second two components are jump processes which are each driven by an independent compound Poisson process and revert to 0 at the same rate $\frac{1}{\lambda}$. (Note: The assumption that all three components have the same constant reversion rate is made in order to obtain a tractable HJB equation later in the problem, but was not made in the paper by Gonzalez et al [7].) One process models the relatively frequent large upward spikes in the spot price while the other models the much less frequent and smaller downward spikes. We define the driving compound Poisson processes to be

$$L_i(t) = \sum_{j=1}^{\infty} \xi_i^{(j)} I_{\{\tau_i^{(j)} \leq t\}} \tag{1.6}$$

for $i = 1, 2$ where the $\tau_i^{(j)}$s are the arrival times of independent Poisson processes with rate $\eta_i > 0$ (one for each jump process) and the $\xi_i^{(j)}$s are exponentially distributed jump sizes with parameter $\beta_i > 0$. The compensated compound Poisson processes,

$$\tilde{L}_i(t) = L_i(t) - \mathbb{E}[\xi_i]\eta_i t, \tag{1.7}$$

are therefore martingales, and we denote by $d\tilde{L}_i$ the compensated Poisson measure associated with each process. We then define $Y_1(t)$ and $Y_2(t)$ to be the unique strong solutions to

$$dY_i(t) = -\frac{1}{\lambda} Y_i(t)dt + dL_i(t) \quad , \quad Y_i(0-) = 0 \tag{1.8}$$

for $i = 1, 2$.

The sum of these random components is then multiplied by a deterministic exponential function representing the seasonal shifts in the mean price of electricity denoted by $e^{f(t)}$. Thus, we get the form of the spot price process, $S(t)$ to be

$$S(t) = e^{f(t)} \left[Y_0(t) + Y_1(t) - Y_2(t) \right]. \tag{1.9}$$

For simplicity of analysis, we take $f(t) \equiv 0$ and thus $e^{f(t)} \equiv 1$ for the majority of this paper. That is, we examine the process

$$X(t) = Y_0(t) + Y_1(t) + Y_2(t). \tag{1.10}$$

So, X satisfies the SDE

$$dX(t) = \frac{1}{\lambda}(\mu - X(t))dt + \sigma dW(t) + dL_1(t) + dL_2(t). \tag{1.11}$$

From an application standpoint, this assumption simply requires using the de-seasonalized spot price, which Gonzalez et al [7] provide a simple and effective method for producing from raw data, rather than the true spot price. This additive structure reproduces in an analytically tractable way the main characteristics of the energy spot price: its large, discontinuous jumps both upwards and downwards and its strong reversion towards a mean price. These characteristics can be seen in the plot below, which shows a sample path of each component process, Y_0, Y_1, and Y_2, as well as their sum, X.

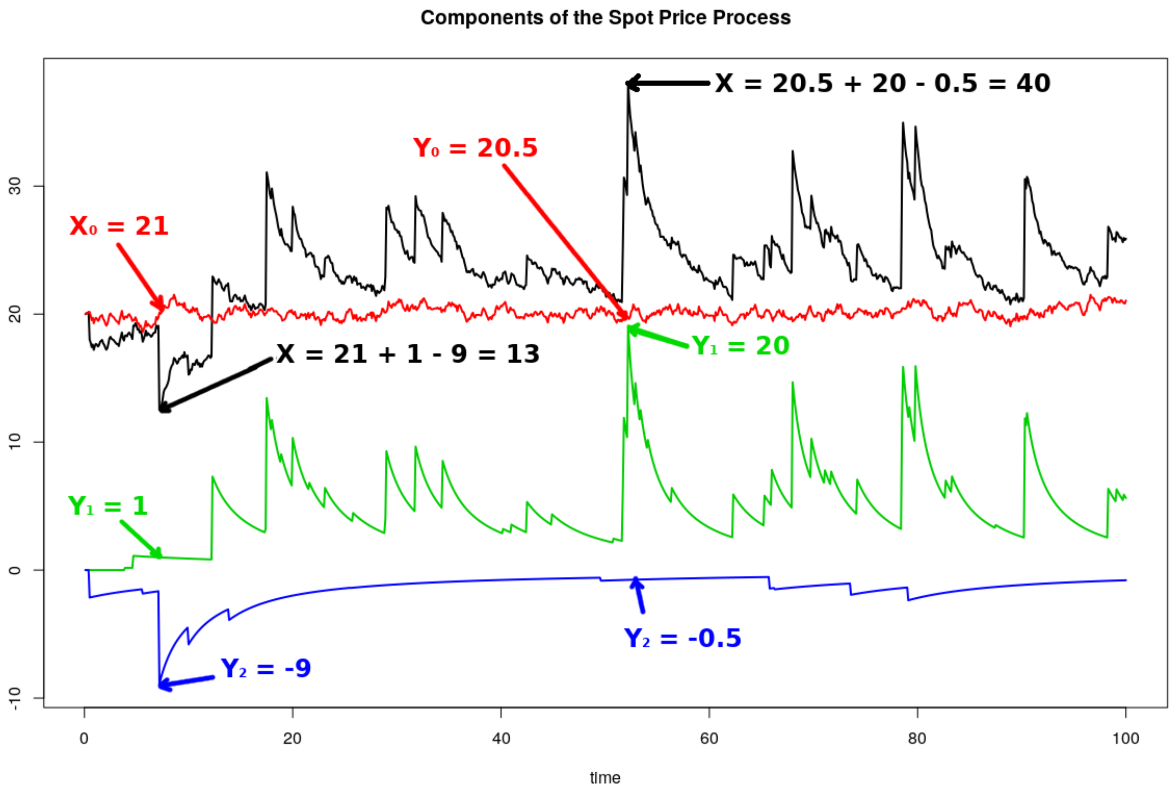

Figure 2: Sample path of spot price process model and its components

Referring to the paper by Gerber and Shiu [6] for the form of the compound Poisson terms, we can then write the generator of the spot price process for any bounded function $f : \mathbb{R} \to \mathbb{R}$ as

$$\tilde{A}f(x) = \frac{1}{\lambda}(\mu - x)f'(x) + \frac{\sigma^2}{2}f''(x) + \eta_1 \int_0^\infty (f(x+y) - f(x))\, \beta_1 e^{-\beta_1 y} dy + \eta_2 \int_{-\infty}^0 (f(x+y) - f(x))\, \beta_2 e^{\beta_2 y} dy. \tag{1.12}$$

3.3 The Paired Spot Price/Coal Process

We assume throughout that the spot price process and the coal process are independent of one another. This allows us to define the generator of the paired process (X, Z) for any bounded function $f : \mathbb{R} \times [z_{min}, z_{max}] \to \mathbb{R}$ to be

$$Af(x,z) := \frac{1}{\lambda}(\mu - x)f_x(x,z) + \frac{\sigma^2}{2}f_{xx}(x,z) + \eta_1 \int_0^\infty (f(x+y,z) - f(x,z))\, \beta_1 e^{-\beta_1 y}\, dy$$

$$+ \eta_2 \int_{-\infty}^{0} \left(f(x+y,z) - f(x,z)\right) \beta_2 e^{\beta_2 y}\, dy - u f_z(x,z)$$
$$+ \alpha \big[p_1 \left(f(x, z+\zeta) - f(x,z)\right) + p_2 \left(f(x, z+2\zeta) - f(x,z)\right) \big]. \tag{1.13}$$

3.4 Electricity Production/Revenue

The plant is paid the spot price at any given time for each unit of electricity production. We consider a continuous revenue function $R(x,z,u)$, which encapsulates the relationship between the burn rate, u, and the production rate of electricity as well as the structure of the payment received by the producer for that electricity. We can also consider a continuous cost function $C(x,z,u)$. The payoff function can then be written as

$$G(x,z,u) := R(x,z,u) - C(x,z,u). \tag{1.14}$$

4 Control Problem

The problem considered in this thesis is to maximize the revenue received by a producer selling electricity to a single node on the MISO grid in the real-time market. That is, a producer selling at a spot price determined by the market. The only control the producer has is the rate at which coal is being burned at the plant, which determines the rate at which electricity is produced as any electricity produced must be sold immediately. Since the coal supply is received based on a long-term contract with a supplier at a fixed price, the cost of burning coal is considered to be fixed regardless of the burn rate. So, we can consider the problem as simply maximizing revenue without including these fixed cost terms. However, if the coal supply ever reaches a certain minimum level, z_{min} the plant must be shut down in order to avoid damage. This shutdown and restart of the plant is extremely costly, so we wish to avoid it. Thus, we define the first hitting time of the minimum coal level to be $\tau := \min\{t \geq 0 \,|\, Z_t = z_{min}\}$, and we consider the problem of maximizing the expected revenue for the plant until the first time the coal supply reaches this minimum level. That is, for discount rate $\delta > 0$, we wish to maximize

$$\mathbb{E}\left[\int_0^{\tau} e^{-\delta t} G(X(t), Z(t), u(t))\, dt\right] \tag{1.15}$$

where $u(t) : \mathbb{R}^+ \to [u_{min}, u_{max}]$ is the non-anticipating control representing burn rate for the plant at time t and τ is the first time at which the process Z reaches z_{min} (at which time the paired process terminates).

5 Notation

For the sake of notational clarity and brevity, the following notational conventions will be used throughout the remainder of the paper.

$$\mathbf{1}_A(\cdot) \quad \text{denotes the indicator of the set A.} \tag{1.16}$$

$$U := [u_{min}, u_{max}] \quad \text{denotes the range of possible burn rates for the plant.} \tag{1.17}$$

$$\mathcal{U}_{(x,z)} \quad \text{denotes the space of admissible control functions for the paired process with initial position} (x, z). \tag{1.18}$$

$$\mathcal{D} := \mathbb{R} \times [z_{min}, z_{max}] \quad \text{denotes the range of the paired process } (X, Z). \tag{1.19}$$

We will also occasionally make use of the probabilists' notational convention for stochastic processes, X, of taking the notations $X(t)$ and X_t to be equivalent as some expressions are more clear with one notation or the other.

Chapter 2

Derivation of the HJB Equation

1 Overview of Continuous Time Stochastic Control

We will use a version of the dynamic programming principle presented in Pham [13], which differs slightly from the standard version in its conclusion of the equivalence of using supremum or infimum over the set of stopping times.

Theorem 2.1. *(Dynamic Programming Principle) Let $X(t)$ be a controlled Markov process, $\mathcal{A}(t,x)$ be the family of admissible controls for the initial point (t,x), and $T_{t,T}$ be the family of all stopping times between t and T. Then then following hold:*

(a) (Finite Time Horizon) Let $(t,x) \in [0,T] \times \mathbb{R}^n$. Then we have

$$v(t,x) = \sup_{\alpha \in \mathcal{A}(t,x)} \sup_{\theta \in \mathcal{T}_{t,T}} \mathbb{E}\left[\int_t^\theta f\left(s, X_s^{t,x}, \alpha_s\right)\, ds + v\left(\theta, X_\theta^{t,x}\right)\right] \tag{2.1}$$

$$= \sup_{\alpha \in \mathcal{A}(t,x)} \inf_{\theta \in \mathcal{T}_{t,T}} \mathbb{E}\left[\int_t^\theta f\left(s, X_s^{t,x}, \alpha_s\right)\, ds + v\left(\theta, X_\theta^{t,x}\right)\right]. \tag{2.2}$$

(b) (Infinite Time Horizon) Let $x \in \mathbb{R}^n$. Then we have

$$v(x) = \sup_{\alpha \in \mathcal{A}(x)} \sup_{\theta \in \mathcal{T}} \mathbb{E}\left[\int_0^\theta e^{-\beta s} f\left(X_s^x, \alpha_s\right)\, ds + v\left(X_\theta^x\right)\right] \tag{2.3}$$

$$= \sup_{\alpha \in \mathcal{A}(x)} \inf_{\theta \in \mathcal{T}} \mathbb{E}\left[\int_0^\theta e^{-\beta s} f\left(X_s^x, \alpha_s\right)\, ds + v\left(X_\theta^x\right)\right]. \tag{2.4}$$

Proof. For the sake of completeness, we quote the proof of this version of the dynamic programming principle in the finite time horizon case with terminal time T as seen in Theorem 3.3.1 in Pham [13].

Given an admissible control $\alpha \in \mathcal{A}(t,x)$, we have pathwise uniqueness of the flow of the SDE for X, the Markovian structure

$$X_s^{t,x} = X_s^{\theta, X_\theta^{t,x}} \quad , \quad s \geq 0$$

for any stopping time $\theta \in [t,T]$. By the law of iterated conditional expectation, we then get

$$J(t,x,\alpha) = \mathbb{E}\left[\int_t^\theta f(s, X_s^{t,x}, \alpha_s)\, ds + J(\theta, X_\theta^{t,x}, \alpha)\right],$$

and since $J(\cdot,\cdot,\alpha) \leq v$ and θ is arbitrary in $\mathcal{T}_{t,T}$

$$\begin{aligned} J(t,x,\alpha) &\leq \inf_{\theta \in \mathcal{T}_{t,T}} \mathbb{E}\left[\int_t^\theta f(s, X_s^{t,x}, \alpha_s)\, ds + v(\theta, X_\theta^{t,x})\right] \\ &\leq \sup_{\alpha \in \mathcal{A}(t,x)} \inf_{\theta \in \mathcal{T}_{t,T}} \mathbb{E}\left[\int_t^\theta f(s, X_s^{t,x}, \alpha_s)\, ds + v(\theta, X_\theta^{t,x})\right]. \end{aligned}$$

By taking the supremum over α in the left-hand side term, we obtain the inequality:

$$v(t,x) \leq \sup_{\alpha \in \mathcal{A}(t,x)} \inf_{\theta \in \mathcal{T}_{t,T}} \mathbb{E}\left[\int_t^\theta f(s, X_s^{t,x}, \alpha_s)\, ds + v(\theta, X_\theta^{t,x})\right]. \tag{2.5}$$

Fix some arbitrary control $\alpha \in \mathcal{A}(t,x)$ and $\theta \in \mathcal{T}_{t,T}$. By definition of the value functions, for any $\epsilon > 0$ and $\omega \in \Omega$, there exists $\alpha^{\epsilon,\omega} \in \mathcal{A}(\theta(\omega), X_{\theta(\omega)}^{t,x}(\omega))$, which is an ϵ-optimal control for $v(\theta(\omega), X_{\theta(\omega)}^{t,x}(\omega))$, i.e.

$$v(\theta(\omega), X_{\theta(\omega)}^{t,x}(\omega)) - \epsilon \leq J(\theta(\omega), X_{\theta(\omega)}^{t,x}(\theta(\omega)), \alpha^{\epsilon,\omega}).$$

Let us now define the process

$$\hat{\alpha}_s(\omega) = \begin{cases} \alpha_s(\omega) & , \ s \in [0, \theta(\omega)] \\ \alpha_s^{\epsilon,\omega}(\omega) & , \ s \in [\theta(\omega), T] \end{cases}.$$

It can be shown by the measurable selection theorem (see, e.g. Chapter 7 in [5]) that the process $\hat{\alpha}$ is progressively measurable, and so lies in $\mathcal{A}(t,x)$. By using again the law of iterated conditional expectation, we obtain

$$\begin{aligned} v(t,x) \geq J(t,x,\hat{\alpha}) &= \mathbb{E}\left[\int_t^\theta f(s, X_s^{t,x}, \alpha_s)\, ds + J(\theta, X_\theta^{t,x}, \alpha^\epsilon)\right] \\ &\geq \mathbb{E}\left[\int_t^\theta f(s, X_s^{t,x}, \alpha_s)\, ds + J(\theta, X_\theta^{t,x}, \alpha^\epsilon)\right] - \epsilon. \end{aligned}$$

From the arbitrariness of $\alpha \in \mathcal{A}(t,x)$, $\theta \in \mathcal{T}_{t,T}$ and $\epsilon > 0$, and we obtain the inequality

$$\sup_{\alpha \in \mathcal{A}(t,x)} \sup_{\theta \in \mathcal{T}_{t,T}} \mathbb{E}\left[\int_t^\theta f\left(s, X_s^{t,x}, \alpha_s\right) ds + v\left(\theta, X_\theta^{t,x}\right)\right] \leq v(t,x). \tag{2.6}$$

By combining the two relations (2.5) and (2.6), we get the required result. □

Pham remarks that the following extension to existence of ϵ-optimal controls also holds.

Remark 2.2. (Existence of ϵ-Optimal Control)

(a) In the finite time horizon case, for all $\alpha \in \mathcal{A}(t,x)$ and $\theta \in \mathcal{T}_{t,T}$:

$$v(t,x) \geq \mathbb{E}\left[\int_t^\theta f(s, X_t^{t,x}, \alpha_s) ds + v(\theta, X_\theta^{t,x})\right].$$

In the infinite time horizon case, for all $\alpha \in \mathcal{A}(x)$ and $\theta \in \mathcal{T}$:

$$v(x) \geq \mathbb{E}\left[\int_t^\theta e^{-\beta s} f(s, X_t^x, \alpha_s) ds + v(\theta, X_\theta^x)\right].$$

(b) In the finite time horizon case, for all $\epsilon > 0$, there exists $\alpha \in \mathcal{A}(t,x)$ such that for all $\theta \in \mathcal{T}_{t,T}$:

$$v(t,x) - \epsilon \leq \mathbb{E}\left[\int_t^\theta f(s, X_t^{t,x}, \alpha_s) ds + v(\theta, X_\theta^{t,x})\right].$$

In the infinite time horizon case, for all $\epsilon > 0$, there exists $\alpha \in \mathcal{A}(x)$ such that for all $\theta \in \mathcal{T}$:

$$v(x) - \epsilon \leq \mathbb{E}\left[\int_t^\theta e^{-\beta s} f(s, X_t^x, \alpha_s) ds + v(\theta, X_\theta^x)\right].$$

2 Derivation of the HJB Equation

Recall that the generator of the paired spot price and coal process (X, Z) is given by

$$\begin{aligned} Af(x,z,u) := \frac{1}{\lambda}(\mu - x) f_x(x,z) + \frac{\sigma^2}{2} f_{xx}(x,z) + \eta_1 \int_0^\infty \left(f(x+y,z) - f(x,z)\right) \beta_1 e^{-\beta_1 y}\, dy \\ + \eta_2 \int_{-\infty}^0 \left(f(x-y,z) - f(x,z)\right) \beta_2 e^{-\beta_2 y}\, dy - u f_z(x,z) \\ + \alpha\big[p_1 \left(f(x, z+\zeta) - f(x,z)\right) + p_2 \left(f(x, z+2\zeta) - f(x,z)\right)\big] \end{aligned}$$

which, using standard dynamic programming techniques (see, e.g., Chapter 3 of Pham [13]), leads to the HJB equation

$$\delta f(x,z) - \sup_{u\in U}\{Af(x,z) + G(x,z,u)\} = 0 \tag{2.7}$$

for all $(x,z) \in \mathbb{R}\times[z_{min}, z_{max}]$ where δ is the constant discount rate. The boundary condition with respect to Z imposed by the stopping time τ indicating the first hitting time of the minimum coal supply is $V(x, z_{min}) = 0$ for all $x \in \mathbb{R}$. So, noting that the control, u, appears only in the f_z term of the generator and in the payoff function, we have the explicit form

$$\begin{aligned}
0 = \delta f(x,z) - \frac{1}{\lambda}(\mu - x)f_x(x,z) - \frac{\sigma^2}{2}f_{xx}(x,z) - \sup_{u\in U}\{-uf_z(x,z) + G(x,z,u)\}& \\
- \eta_1 \int_0^\infty \left(f(x+y,z) - f(x,z)\right)\beta_1 e^{-\beta_1 y}\,dy& \\
- \eta_2 \int_{-\infty}^0 \left(f(x-y,z) - f(x,z)\right)\beta_2 e^{\beta_2 y}\,dy& \\
- \alpha\big[p_1\left(f(x,z+\zeta) - f(x,z)\right) + p_2\left(f(x,z+2\zeta) - f(x,z)\right)\big].&
\end{aligned} \tag{2.8}$$

Chapter 3

Value Function is a Viscosity Solution to the HJB Equation

We limit our consideration to a market in which the producer may sell energy only to a single node on the grid at the spot price for that node.

Recall that the generator of the paired spot price and coal process (X, Z) is given by

$$\begin{aligned} Af(x,z,u) := \frac{1}{\lambda}(\mu - x)f_x(x,z) + \frac{\sigma^2}{2} f_{xx}(x,z) + \eta_1 \int_0^\infty \left(f(x+y,z) - f(x,z)\right)\beta_1 e^{-\beta_1 y}\, dy \\ + \eta_2 \int_{-\infty}^0 \left(f(x+y,z) - f(x,z)\right)\beta_2 e^{-\beta_2 y}\, dy - u f_z(x,z) \\ + \alpha\big[p_1\left(f(x,z+\zeta) - f(x,z)\right) + p_2\left(f(x,z+2\zeta) - f(x,z)\right)\big] \end{aligned}$$

which leads to the HJB equation

$$\delta V(x,z) - \sup_{u\in[u_{min},u_{max}]} \{AV(x,z) + G(x,z,u)\} = 0 \tag{3.1}$$

for all $(x,z) \in \mathbb{R} \times [z_{min}, z_{max}]$ where δ is the constant discounting rate. The boundary condition with respect to Z imposed by the stopping time τ indicating the first hitting time of the minimum coal supply is $V(x, z_{min}) = 0$ for all $x \in \mathbb{R}$.

The nonlocal behavior due to the jumps in the spot price and the instantaneous nature of the coal arrivals requires an extension of the classical theory of viscosity solutions. This chapter follows closely the structure of the classical proofs seen in Chapter 4 of Pham's text [13], but some careful adjustments must be made to ensure that the nonlocal terms remain locally bounded.

1 Definition of Viscosity Solution

Let $\mathcal{O}$ be an open domain in $\mathbb{R}^d$.

Definition 3.1. Let $w : \mathcal{O} \to \mathbb{R}$ be locally bounded. Then we define

(a) the upper-semicontinuous envelope w^* by $w^*(\mathbf{x}) = \limsup_{\mathbf{x}' \to \mathbf{x}} w(\mathbf{x}')$

(b) the lower-semicontinuous envelope w_* by $w_*(\mathbf{x}) = \liminf_{\mathbf{x}' \to \mathbf{x}} w(\mathbf{x}')$

We will consider a second-order PIDE (partial integral-differential equation) of the form

$$F(\mathbf{x}, w(\mathbf{x}), Dw(\mathbf{x}), D^2 w(\mathbf{x}), \mathcal{I}[\mathbf{x}, w]) = 0 \tag{3.2}$$

where $\mathcal{I}[\mathbf{x}, w]$ is an operator which contains all non-local terms of the PIDE.

Definition 3.2. Let $w : \mathcal{O} \to \mathbb{R}$ be locally bounded.

(a) w is a (possibly discontinuous) viscosity subsolution of (3.2) on $\mathcal{O}$ if

$$F(\bar{\mathbf{x}}, w^*(\mathbf{x}), D\phi(\bar{\mathbf{x}}), D^2\phi(\bar{\mathbf{x}}), \mathcal{I}[\bar{\mathbf{x}}, \phi]) \le 0$$

for all $\bar{\mathbf{x}} \in \mathcal{O}$ and for all $\phi \in C^2(\mathcal{O})$ such that $\bar{\mathbf{x}}$ is a maximum point of $w^* - \phi$.

(b) w is a (possibly discontinuous) viscosity supersolution of (3.2) on $\mathcal{O}$ if

$$F(\bar{\mathbf{x}}, w_*(\mathbf{x}), D\phi(\bar{\mathbf{x}}), D^2\phi(\bar{\mathbf{x}}), \mathcal{I}[\bar{\mathbf{x}}, \phi]) \ge 0$$

for all $\bar{\mathbf{x}} \in \mathcal{O}$ and for all $\phi \in C^2(\mathcal{O})$ such that $\bar{\mathbf{x}}$ is a minimum point of $w_* - \phi$.

(c) w is a (possibly discontinuous) viscosity solution of (3.2) on $\mathcal{O}$ if it is both a viscosity subsolution and a viscosity supersolution of (3.2).

2 Verification That the Value Function is a Viscosity Solution of the HJB Equation

Proposition 3.1. *Suppose the value function*

$$V(x, z) = \sup_{u \in \mathcal{U}} \sup_{\theta \in \mathcal{T}} \mathbb{E}\left[\int_0^{\theta \wedge \tau} e^{-\delta r} G(X(r), Z(r), u(r))\, dr\right]$$

is locally bounded. Then, V is a viscosity supersolution to (3.1).

Proof. Let $(\bar{x}, \bar{z}) \in \mathbb{R} \times [z_{min}, z_{max}]$. Let $\phi \in C^2(\mathbb{R} \times [z_{min}, z_{max}])$ be a test function such that (recalling V_* represents the lower semicontinuous envelope of V)

$$0 = (V_* - \phi)(\bar{x}, \bar{z}) = \min_{(x,z) \in \mathbb{R} \times [z_{min}, z_{max}]} (V_* - \phi)(x, z) \tag{3.3}$$

Since V_* is defined to be $V_*(x,z) = \liminf_{(x',z') \to (x,z)} V(x', z')$ for each (x, z), there exists a sequence $(x_n, z_n) \subset \mathbb{R} \times [z_{min}, z_{max}]$ such that $(x_n, z_n) \to (\bar{x}, \bar{z})$ and $V(x_n, z_n) \to V_*(\bar{x}, \bar{z})$ as $n \to \infty$. Also, since ϕ is continuous, $\phi(x_n, z_n) \to \phi(\bar{x}, \bar{z})$. Thus,

$$\gamma_n := V(x_n, z_n) - \phi(x_n, z_n) \to 0 \quad \text{ as } n \to \infty$$

Let $\mathrm{u} \in [u_{min}, u_{max}]$ and set $u(t) \equiv \mathrm{u}$. Then, $u \in \mathcal{U}$. Denote the controlled process under this u starting at initial point (x_n, z_n) by $(X_r^{(n)}, Z_r^{(n)})$. Fix some $\rho > 0$, and let $\pi_n := \inf\{r \geq 0 \mid |Y_0^{(n)}(r) - \bar{x}| \geq \rho\}$. Let $(h_n) \subset \mathbb{R}^+$ be a sequence of positive numbers such that $h_n \to 0$ and $\frac{\gamma_n}{h_n} \to 0$ as $n \to \infty$.

Further, let $\xi_1 = \min\{t > 0 \mid L_1(t) - L_1(t-) \neq 0\}$ and $\xi_2 = \min\{t > 0 \mid L_2(t) - L_2(t-) \neq 0\}$, i.e. the first time a jump occurs in each of the Poisson processes driving the price process. Note that the jump processes L_1 and L_2 which drive the jumps in the price process X are independent of the starting price, that is these stopping times are the same for each n since they are independent of x_n. Let $\xi_3 = \min\{t > 0 \mid Z^{(n)}(t) - Z^{(n)}(t-) \neq 0\}$, and set $\xi = \xi_1 \wedge \xi_2 \wedge \xi_3$. Further, take two sequences of values in $[z_{min}, z_{max}]$, $\tilde{z}_n \to z_{min}$ and $\tilde{z}^n \to z_{max}$, and let $\tau_n := \inf\{t > 0 \mid Z^{(n)}(t) \notin (\tilde{z}_n, \tilde{z}^n)\}$. Without loss of generality, we can assume that $z_n \in (\tilde{z}_n, \tilde{z}^n)$ for all n. Finally, let $\theta_n := \pi_n \wedge h_n \wedge \xi \wedge \tau_n \wedge \tau$. Then (recalling the notational convention that $X_t = X(t)$ and $Z_t = Z(t)$), since $V(x, z)$ is defined as the supremum over all admissible controls and the supremum over all stopping times θ, applying the dynamic programming principle of Theorem 2.1, we have in particular that

$$V(x_n, z_n) \geq \mathbb{E}\left[\int_0^{\theta_n} e^{-\delta r} G(X_{r\wedge\tau}^{(n)}, Z_{r\wedge\tau}, \mathrm{u})\, dr + e^{-\delta\theta_n} V(X_{\theta_n}^{(n)}, Z_{\theta_n}^{(n)})\right].$$

Note that (3.3) implies that $V(x, z) \geq V_*(x, z) \geq \phi(x, z)$ for all $(x, z) \in \mathbb{R} \times [z_{min}, z_{max}]$. So, together with the previous inequality, we have

$$\begin{aligned}\phi(x_n, z_n) + \gamma_n &= \phi(x_n, z_n) + V(x_n, z_n) - \phi(x_n, z_n) = V(x_n, z_n)\\ &\geq \mathbb{E}\left[\int_0^{\theta_n} e^{-\delta r} G(X_{r\wedge\tau}^{(n)}, Z_{r\wedge\tau}^{(n)}, \mathrm{u})\, dr + e^{-\delta\theta_n} V(X_{\theta_n}^{(n)}, Z_{\theta_n}^{(n)})\right]\end{aligned}$$

$$\geq \mathbb{E}\left[\int_0^{\theta_n} e^{-\delta r} G(X_{r\wedge\tau}^{(n)}, Z_{r\wedge\tau}^{(n)}, \mathrm{u})\, dr + e^{-\delta\theta_n}\phi(X_{\theta_n}^{(n)}, Z_{\theta_n}^{(n)})\right].$$

Now, applying Itô's formula to $e^{-\delta\theta_n}\phi(X_{\theta_n}^{(n)}, Z_{\theta_n}^{(n)})$ and recalling the compensated jump processes $\tilde{L}_1$, $\tilde{L}_2$, and $\tilde{Z}$, which are local martingales, we get

$$\begin{aligned}\phi(x_n, z_n) + \gamma_n \geq \mathbb{E}\bigg[&\int_0^{\theta_n} e^{-\delta r} G(X_r^{(n)}, Z_r^{(n)}, \mathrm{u})\, dr + \phi(X_0^{(n)}, Z_0^{(n)}) \\ &+ \int_0^{\theta_n} e^{-\delta r}\left(A\phi(X_r^{(n)}, Z_r^{(n)}) - \delta\phi(X_r^{(n)}, Z_r^{(n)})\right) dr \\ &+ \int_0^{\theta_n} e^{-\delta r}\phi_x(X_r^{(n)}, Z_r^{(n)})\, dW_r + \int_0^{\theta_n} e^{-\delta r}\phi_x(X_r^{(n)}, Z_r^{(n)})\, \tilde{L}_1(dr) \\ &+ \int_0^{\theta_n} e^{-\delta r}\phi_x(X_r^{(n)}, Z_r^{(n)})\, \tilde{L}_2(dr) + \int_0^{\theta_n} e^{-\delta r}\phi_x(X_r^{(n)}, Z_r^{(n)})\, d\tilde{Z}_r^{(n)}\bigg].\end{aligned}$$

Since $X_0^{(n)} = x_n$ and $Z_0^{(n)} = z_n$, we can replace $\phi(X_0^{(n)}, Z_0^{(n)})$ by $\phi(x_n, z_n)$ which is a constant. Further, note that since ϕ is in $C^2(\mathbb{R}\times[z_{min}, z_{max}])$ and the choice of stopping time θ_n means that $X_r^{(n)}$ and $Z_r^{(n)}$ are both bounded, we have that both $\phi(X_r^{(n)}, Z_r^{(n)})$ and $\phi_x(X_r^{(n)}, Z_r^{(n)})$ are continuous and bounded, so the stochastic integrals all have mean 0. These observations yield

$$\begin{aligned}\phi(x_n, z_n) + \gamma_n \geq \mathbb{E}\bigg[&\int_0^{\theta_n} e^{-\delta r} G(X_r^{(n)}, Z_r^{(n)}, \mathrm{u})\, dr \\ &+ \int_0^{\theta_n} e^{-\delta r}\left(A\phi(X_r^{(n)}, Z_r^{(n)}) - \delta\phi(X_r^{(n)}, Z_r^{(n)})\right) dr\bigg] + \phi(x_n, z_n)\end{aligned}$$

which implies that

$$\gamma_n - \mathbb{E}\left[\int_0^{\theta_n} e^{-\delta r} G(X_r^{(n)}, Z_r^{(n)}, \mathrm{u})\, dr + \int_0^{\theta_n} e^{-\delta r}\left(A\phi(X_r^{(n)}, Z_r^{(n)} - \delta\phi(X_r^{(n)}, Z_r^{(n)})\right) dr\right] \geq 0 \tag{3.4}$$

Dividing by h_n on both sides, we get

$$\frac{\gamma_n}{h_n} + \mathbb{E}\left[\frac{1}{h_n}\int_0^{\theta_n} e^{-\delta r}\left(\delta\phi(X_r^{(n)}, Z_r^{(n)}) - A\phi(X_r^{(n)}, Z_r^{(n)}) - G(X_r^{(n)}, Z_r^{(n)}, \mathrm{u})\right) dr\right] \geq 0 \tag{3.5}$$

Now, we consider

$$0 \leq \liminf_{n\to\infty}\left(\frac{\gamma_n}{h_n} + \frac{1}{h_n}\mathbb{E}\left[\int_0^{\theta_n} e^{-\delta r}\left(\delta\phi(X_r^{(n)}, Z_r^{(n)}) - A\phi(X_r^{(n)}, Z_r^{(n)}) - G(X_r^{(n)}, Z_r^{(n)}, \mathrm{u})\right) dr\right]\right)$$

$$= \liminf_{n\to\infty} \mathbb{E}\left[\frac{1}{h_n}\int_0^{\theta_n} e^{-\delta r}\left(\delta\phi(X_r^{(n)}, Z_r^{(n)}) - A\phi(X_r^{(n)}, Z_r^{(n)}) - G(X_r^{(n)}, Z_r^{(n)}, \mathrm{u})\right)\, dr\right]$$

By definition of θ_n, we have that $|X_r^{(n)} - x_n| < \rho$ for all $r \in [0, \theta_n)$. Further, since $x_n \to \bar{x}$, for any $\delta_1 > 0$ there exists some N such that $|x_n - \bar{x}| \le \delta_1$ for all $n \ge N$. Thus, there exist x_* and x^* such that $x_n \in [x_* - \delta_1, x^* + \delta_1]$ for all n. That is, for all n, $X_r^{(n)} \in [x_* - \delta_1 - \rho, x^* + \delta_1 + \rho]$ for all $r \in [0, \theta_n)$. Also, by definition, for all n, $Z_r^{(n)} \in [z_{min}, z_{max}]$ for all r. Moreover, by definition of ξ, for all n the paired process $(X_\cdot^{(n)}, Z_\cdot^{(n)})$ is continuous for r in the time interval $[0, \theta_n)$. So, we have that the integrand above, $e^{-\delta r}(\delta\phi(x_r^{(n)}, Z_r^{(n)}) - A\phi(X_r^{(n)}, Z_r^{(n)}) - G(X_r^{(n)}, Z_r^{(n)}, u(r)))$, is continuous on $[0, \theta_n)$ and that there exists a compact set $C := [x_* - \delta_1 - \rho, x^* + \delta_1 + \rho] \times [z_{min}, z_{max}]$ independent of n such that $(X_\cdot^{(n)}, Z_\cdot^{(n)}) \in C$ for all n. That is, there exists a uniform bound $M < \infty$ such that $|e^{-\delta r}(\delta\phi(x_r^{(n)}, Z_r^{(n)}) - A\phi(X_r^{(n)}, Z_r^{(n)}) - G(X_r^{(n)}, Z_r^{(n)}, u(r)))| \le M$ for all n. So, we have for all n that

$$\left|\frac{1}{h_n}\int_0^{\theta_n} e^{-\delta r}(\delta\phi(X_r^{(n)}, Z_r^{(n)}) - A\phi(X_r^{(n)}, Z_r^{(n)}) - G(X_r^{(n)}, Z_r^{(n)}, u(r)))\, dr\right| \le \frac{1}{h_n}\int_0^{\theta_n} 2M\, dr = \frac{\theta_n}{h_n}(2M) \le 2M \tag{3.6}$$

since $\theta_n \le h_n$ for all n.

We continue the analysis of this inequality with the following lemma.

Lemma 3.2.

$$\lim_{n\to\infty} \frac{1}{h_n}\int_0^{\theta_n} e^{-\delta r}(\delta\phi(X_r^{(n)}, Z_r^{(n)}) - A\phi(X_r^{(n)}, Z_r^{(n)}) - G(X_r^{(n)}, Z_r^{(n)}, u))\, dr = \delta\phi(x_n, z_n) - A\phi(\bar{x}, \bar{z}) - G(\bar{x}, \bar{z}, u)$$

almost surely.

Proof. First, define a sequence of subsets of the sample space Ω of the paired process (X, Z) as follows

$$E_n := \{\omega \in \Omega \,|\, \theta_n = h_n\}.$$

Then, $\Omega = E_n \cup E_n^c$ for each n, and therefore,

$$\frac{1}{h_n}\int_0^{\theta_n} e^{-\delta r}(-A\phi(X_r^{(n)}, Z_r^{(n)}) - G(X_r^{(n)}, Z_r^{(n)}, \mathrm{u}))\, dr = \frac{1}{h_n}\int_0^{\theta_n} \mathbf{1}_{E_n} e^{-\delta r}(-A\phi(X_r^{(n)}, Z_r^{(n)}) - G(X_r^{(n)}, Z_r^{(n)}, u(r)))\, dr$$

$$+\frac{1}{h_n}\int_0^{\theta_n} \mathbf{1}_{E_n^c} e^{-\delta r}(-A\phi(X_r^{(n)}, Z_r^{(n)}) - G(X_r^{(n)}, Z_r^{(n)}, \mathrm{u}))\, dr$$

Note that by Lemma 5.2 and Lemma 5.3 in the appendix, we have that $\mathbf{1}_{E_n} \to 1$ a.s. and $\mathbf{1}_{E_n^c} \to 0$ a.s as $n \to \infty$. So, using this result and the fact that $(x_n, z_n) \to (\bar{x}, \bar{z})$, we conclude from the mean value theorem that

$$\begin{aligned}
&\lim_{n\to\infty} \frac{1}{h_n}\int_0^{\theta_n} e^{-\delta r}\left(\delta\phi(X_r^{(n)}, Z_r^{(n)}) - A\phi(X_r^{(n)}, Z_r^{(n)}) - G(X_r^{(n)}, Z_r^{(n)}, \mathrm{u})\right)\, dr \\
&\qquad = \lim_{n\to\infty} \frac{1}{h_n}\int_0^{\theta_n} e^{-\delta r}\left(\delta\phi(X_r^{(n)}, Z_r^{(n)}) - A\phi(X_r^{(n)}, Z_r^{(n)}) - G(X_r^{(n)}, Z_r^{(n)}, \mathrm{u})\right)\, dr \mathbf{1}_{E_n} \\
&\qquad\quad + \lim_{n\to\infty} \frac{1}{h_n}\int_0^{\theta_n} e^{-\delta r}\left(\delta\phi(X_r^{(n)}, Z_r^{(n)}) - A\phi(X_r^{(n)}, Z_r^{(n)}) - G(X_r^{(n)}, Z_r^{(n)}, \mathrm{u})\right)\, dr \mathbf{1}_{E_n^c} \\
&\qquad = \delta\phi(\bar{x}, \bar{z}) - A\phi(\bar{s}, \bar{z}) - G(\bar{x}, \bar{z}, \mathrm{u})
\end{aligned}$$

almost surely. □

Returning to the proof of Proposition (3.1) , we have from the dominated convergence theorem that

$$\begin{aligned}
0 &\le \liminf_{n\to\infty} \mathbb{E}\left[\frac{1}{h_n}\int_0^{\theta_n} e^{-\delta r}\left(\delta\phi(X_r^{(n)}, Z_r^{(n)}) - A\phi(X_r^{(n)}, Z_r^{(n)}) - G(X_r^{(n)}, Z_r^{(n)}, \mathrm{u})\right)\, dr\right] \\
&= \mathbb{E}\left[\lim_{n\to\infty} \frac{1}{h_n}\int_0^{\theta_n} e^{-\delta r}\left(\delta\phi(X_r^{(n)}, Z_r^{(n)}) - A\phi(X_r^{(n)}, Z_r^{(n)}) - G(X_r^{(n)}, Z_r^{(n)}, \mathrm{u})\right)\, dr\right] \\
&= \delta\phi(\bar{x}, \bar{z}) - A\phi(\bar{x}, \bar{z}) - G(\bar{x}, \bar{z}, \mathrm{u}) \ge 0. \qquad (3.7)
\end{aligned}$$

So, since u was arbitrary in $[u_{min}, u_{max}]$, we have that the value function $V(x, z)$ is a viscosity supersolution of (3.1). □

Proposition 3.3. *Suppose the value function*

$$V(x, z) = \sup_{u\in\mathcal{U}} \sup_{\theta\in\mathcal{T}} \mathbb{E}\left[\int_0^{\theta\wedge\tau} e^{-\delta r} G(X(r), Z(r), u(r))\, dr\right]$$

is locally bounded. Then, V is a viscosity subsolution to (3.1)*.*

Proof. Let $(\bar{x}, \bar{z}) \in \mathbb{R} \times [z_{min}, z_{max}]$. Let $\phi \in C^2(\mathbb{R} \times [z_{min}, z_{max}])$ be a test function such that

$$0 = (V_* - \phi)(\bar{x}, \bar{z}) = \max_{(x,z)\in\mathbb{R}\times[z_{min}, z_{max}]} (V_* - \phi)(x, z) \qquad (3.8)$$

We wish to show that $\delta\phi(\bar{x},\bar{z}) - A\phi(\bar{x},\bar{z}) - G(\bar{x},\bar{z},u) \leq 0$. In order to proceed by contradiction, we assume that

$$\delta\phi(\bar{x},\bar{z}) - A\phi(\bar{x},\bar{z}) - G(\bar{x},\bar{z},u) > 0. \tag{3.9}$$

Then, since ϕ is in $C^2(\mathbb{R} \times [z_{min}, z_{max}])$, there exist constants $\eta > 0$ and $\epsilon > 0$ such that

$$\delta\phi(x',z') - A\phi(x',z') - G(x',z',u) \geq \epsilon.$$

for all $(x',z') \in B((\bar{x},\bar{z}),\eta) = \{(x',z') \in \mathbb{R} \times [z_{min}, z_{max}] : \sqrt{(\bar{x}-x')^2 + (\bar{z}-z')^2} < \eta\}$. Since V^* is defined to be $V^*(x,z) = \limsup_{(x',z')\to(x,z)} V(x',z')$, there exists a sequence $(x_n,z_n) \subset \mathbb{R} \times [z_{min}, z_{max}]$ such that $(x_n,z_n) \to (\bar{x},\bar{z})$ and $V(x_n,z_n) \to V^*(\bar{x},\bar{z})$ as $n \to \infty$. Also, since ϕ is continuous, $\phi(x_n,z_n) \to \phi(\bar{x},\bar{z})$. Thus,

$$\gamma_n := V(x_n,z_n) - \phi(x_n,z_n) \to 0 \quad \text{ as } n \to \infty.$$

Let $(h_n) \subset \mathbb{R}^+$ be a sequence of positive numbers such that $h_n \to 0$ and $\frac{\gamma_n}{h_n} \to 0$ as $n \to \infty$. Define a sequence of stopping times $\theta_n := \pi_n \wedge h_n \wedge \xi \wedge \tau_n \wedge \tau$ with ξ and τ_n defined as in the proof of Proposition 3.1 above and $\pi_n := \inf\{r \geq 0 : |Y_0^{(n)}(r) - \bar{x}| \geq \eta'\}$ for some $0 < \eta' < \eta$ with η' chosen such that $B((x_n,z_n),\eta') \subset B((\bar{x},\bar{z}),\eta)$ and for $0 \leq r \leq \theta_n$, $(X_r^{(n)}, Z_r^{(n)}) \in B((\bar{x},\bar{z}),\eta)$. Then, according to Theorem 2.2, for each n there exists an $\frac{\epsilon h_n}{2}$-optimal control $\hat{u}^{(n)} \in \mathcal{U}$ such that

$$\begin{aligned} V(x_n,z_n) - \frac{\epsilon h_n}{2} &= \phi(x_n,z_n) + \gamma_n - \frac{\epsilon h_n}{2} \\ &\leq \mathbb{E}\left[\int_0^{\theta_n} e^{-\delta r} G(X_r^{(n)}, Z_r^{(n)}, \hat{u}^{(n)}(r))k\, dr + \phi(X_{\theta_n}^{(n)}, Z_{\theta_n}^{(n)})\right]. \end{aligned}$$

Note that if we choose $\eta' < \frac{\zeta}{2}$ (where ζ is the amount of coal delivered in a single shipment), $\phi_z(x,z)$ will be continuous. So, choosing η' in this way and applying Itô's formula, we have

$$\begin{aligned} &\phi(x_n,z_n) + \gamma_n - \frac{\epsilon h_n}{2} \\ &\quad \leq \mathbb{E}\Big[\phi(x_n,z_n) + \int_0^{\theta_n} e^{-\delta r}\left(A\phi(X_r^{(n)}, Z_r^{(n)}) - \delta\phi(X_r^{(n)}, Z_r^{(n)}) + G(X_r^{(n)}, Z_r^{(n)}, \hat{u}^{(n)})\right) dr\Big] \\ &\qquad + \mathbb{E}\left[\int_0^{\theta_n} e^{-\delta r}\phi_x(X_r^{(n)}, Z_r^{(n)})\, d\tilde{X}_r^{(n)} + \int_0^{\theta_n} e^{-\delta r}\phi_z(X_r^{(n)}, Z_r^{(n)})\, d\tilde{Z}_r^{(n)}\right]. \end{aligned} \tag{3.10}$$

By the choice of stopping time θ_n, the integrand in the stochastic integrals above is continuous and bounded on $[0, \theta_n)$, so the expectations of the stochastic integrals are zero. Moreover, by choice of η, for $0 \le s < \theta_m$

$$\delta\phi(X_r^{(n)}, Z_r^{(n)}) - A\phi(X_r^{(n)}, Z_r^{(n)}) - G(X_r^{(n)}, Z_r^{(n)}, \hat{u}(r)) \ge \epsilon,$$

and thus dividing by h_n everywhere in (3.10) shows that

$$\frac{\gamma_n}{h_n} - \epsilon\left(\frac{1}{2} - \frac{1}{h_n}\mathbb{E}[\theta_n]\right) \le 0. \tag{3.11}$$

We now consider

$$\mathbb{P}\left[\pi_n \wedge \xi \wedge \tau_n \wedge \tau \le h_n\right] \le \mathbb{P}[\pi_n \le h_n] + \mathbb{P}[\xi \le h_n] + \mathbb{P}[\tau_n \le h_n] + \mathbb{P}[\tau \le h_n]. \tag{3.12}$$

Using Lemma 5.1, Lemma 5.2, and Lemma 5.3 in the appendix which show that the first three terms of this sum approach 0 as $n \to \infty$ and noting that $\mathbb{P}\left[\tau_n \le h_n\right] \to 0$ implies that $\mathbb{P}\left[\tau \le h_n\right] \to 0$ (since $\tilde{z}_n > z_{min}$ implies that $\tau_n \le \tau$ for all n), we have from (3.12) that

$$\lim_{n\to\infty} \mathbb{P}\left[\pi_n \wedge \xi \wedge \tau_n \wedge \tau \le h_n\right] \le \lim_{n\to\infty} \mathbb{P}[\pi_n \le h_n] + \lim_{n\to\infty} \mathbb{P}[\xi \le h_n] + \lim_{n\to\infty} \mathbb{P}[\tau_n \le h_n] = 0. \tag{3.13}$$

Further, we have that

$$\mathbb{P}\left[\pi_n \wedge \xi \wedge \tau_n \wedge \tau > h_n\right] = \mathbb{E}[\mathbf{1}_{\{\pi_n \wedge \xi \wedge \tau_n \wedge \tau > h_n\}}] = \frac{1}{h_n}\mathbb{E}[h_n \mathbf{1}_{\{\pi_n \wedge \xi \wedge \tau_n \wedge \tau > h_n\}}].$$

By definition, $\theta_n := \pi_n \wedge h_n \wedge \xi \wedge \tau_n \wedge \tau$. So, on the event $\{\pi_n \wedge \xi \wedge \tau_n \wedge \tau > h_n\}$, we have $\theta_n = h_n$. Thus, $h_n \mathbf{1}_{\{\pi_n \wedge \xi \wedge \tau_n \wedge \tau > h_n\}} \equiv \theta_n \mathbf{1}_{\{\pi_n \wedge \xi \wedge \tau_n \wedge \tau > h_n\}}$, and we have

$$\mathbb{P}\left[\pi_n \wedge \xi \wedge \tau_n \wedge \tau > h_n\right] = \frac{1}{h_n}\mathbb{E}[\theta_n \mathbf{1}_{\{\pi_n \wedge \xi \wedge \tau_n \wedge \tau > h_n\}}] \le \frac{1}{h_n}\mathbb{E}[\theta_n]$$

Also note that, by definition, $\theta_n \le h_n$ for every $\omega \in \Omega$ for each n, so $\mathbb{E}[\theta_n] \le h_n$ which gives us

$$\mathbb{P}\left[\pi_n \wedge \xi \wedge \tau_n \wedge \tau > h_n\right] \le \frac{1}{h_n}\mathbb{E}[\theta_n] \le 1$$

So, letting $n \to \infty$, we have that $\frac{1}{h_n}\mathbb{E}[\theta_n] \to 1$. Finally, letting $n \to \infty$, that is $h_n \to 0$, in (3.11) gives the desired contradiction.

□

So, combining the results of Proposition 3.1 and Proposition 3.3, we have that the value function, V, is a viscosity solution to the HJB equation (3.1).

Chapter 4

Uniqueness of Viscosity Solution of the HJB Equation

As previously mentioned, the nonlocal behavior present in both the spot price (due to the discontinuous jumps caused by the compound Poisson terms in the spot price model) and the coal supply (due to the instantaneous nature of coal arrivals) implies that the HJB equation associated with this problem is in fact the second order partial integro-differential equation (PIDE) (3.1) rather than a second order PDE as is usually associated with a stochastic control problem. Thus, the traditional theory of viscosity solutions is not sufficient to determine the uniqueness of a solution to this HJB equation. We turn instead to the work of Barles and Imbert [2], who provide a set of sufficient conditions for the uniqueness of the viscosity solution to a second order PIDE. However, the mean-reversion term of our HJB equation (4.32) fails to satisfy assumption (A3-1) of Barles and Imbert. While mean-reverting processes are generally considered very well-behaved, the assumptions of Barles and Imbert place particularly strong restrictions on the interaction of the position of the process with the drift. Since mean reversion here imposes a drift with exponentially increasing magnitude as the process moves away from the mean pric, μ, there is a very strong interaction between position and drift over much of the domain.

An extension of the theory of Barles and Imbert which provides conditions which are satisfied by this spot price process is developed in this chapter. We first define a continuous function $F(\mathbf{x}, g, p, X, \ell)$ where $\mathbf{x} \in \mathbb{R}^d$, $g \in \mathbb{R}$, $p \in \mathbb{R}^d$, $X \in \mathbb{S}_d$, and $\ell \in \mathbb{R}$. The general form of a second order PIDE with nonlocal behavior is then

$$F\left(\mathbf{x}, f, \nabla f, D^2 f, \mathcal{I}[\mathbf{x}, f]\right) = 0 \tag{4.1}$$

where $\mathbf{x} \in \mathbb{R}^d$, f is a function on $\mathbb{R}^d$, ∇f is the gradient, $D^2 f$ is the Hessian, and $\mathcal{I}[\mathbf{x}, f]$ is an operator which collects all of the terms appearing due to nonlocal behavior.

In order to handle this nonlocal behavior, it is necessary to restrict the class of functions on which we will work. Given an upper-semicontinuous function $R : \mathbb{R}^d \to \mathbb{R}$, define $\mathcal{C}$ to be the space of functions f such that there exists a constant $\bar{c} > 0$ such that for all $\mathbf{x} \in \mathbb{R}^d$,

$$|f(\mathbf{x})| \leq \bar{c}(1 + R(\mathbf{x})).$$

Note the following important properties of the space $\mathcal{C}$:

- Any function $f \in \mathcal{C}$ is locally bounded.
- For any functions $f_1, f_2 \in \mathcal{C}$, $\max\{f_1, f_2\} \in \mathcal{C}$ and $\min\{f_1, f_2\} \in \mathcal{C}$.
- For any compact set $K \subset \mathbb{R}^d$ and function $\phi \in C^2(K)$, there exists a function $\psi \in \mathcal{C}$ such that $\psi = \phi$ on the interior of K.

1 Comparison Principle

1.1 Some Results from Barles and Imbert [2]

We quote here, without proof, a few of the results from the paper by Barles and Imbert [2] which will be used in the following section.

We begin by defining a modified version of the inf-convolution and sup-convolution that are commonly used in viscosity solution theory. For any upper-semicontinuous function $U : \mathbb{R}^m \to \mathbb{R}$ and $r \in \mathbb{R}^m$, we define

$$R^a[U](y, r) := \sup_{|Y-y| \leq 1} \left\{ U(Y) - r \cdot (Y - y) - \frac{|Y - y|^2}{2a} \right\}. \tag{4.2}$$

Similarly, for any lower-semicontinuous function $V : \mathbb{R}^m \to \mathbb{R}$, we define

$$R_a[V](y, r) := \inf_{|Y-y| \leq 1} \left\{ V(Y) - r \cdot (Y - y) + \frac{|Y - y|^2}{2a} \right\}. \tag{4.3}$$

Note that, similarly to the traditional inf/sup-convolutions, $R_a[V] = -R^a[-V]$. The next proposition gives some other useful properties of these functions.

Definition 4.1 (Superjet and Subjet)**.** Let $U : \mathbb{R}^d \to \mathbb{R}$ be an upper-semicontinuous function. A couple $(p, Y) \in \mathbb{R}^d \times \mathbb{S}_d$ is a superjet of U at $y \in \mathbb{R}^d$ if

$$U(y+z) \le U(y) + p \cdot z + \frac{1}{2} Y z \cdot z + o(|z|^2).$$

Let $V : \mathbb{R}^d \to \mathbb{R}$ be a lower-semicontinuous function. A couple $(p, Y) \in \mathbb{R}^d \times \mathbb{S}_d$ is a subjet of V at $y \in \mathbb{R}^d$ if

$$V(y+z) \ge V(y) + p \cdot z + \frac{1}{2} Y z \cdot z + o(|z|^2).$$

We denote by $J^+U(y)$ and $J^-U(y)$, respectively, the set of superjets and subjets of U and V at y and.

Proposition 4.2 (Proposition 3 of Barles and Imbert [2])**.** *For any upper-semicontinuous function $U : \mathbb{R}^m \to \mathbb{R}$ and any lower-semicontinuous function $V : \mathbb{R}^m \to \mathbb{R}$, the functions $R^a[U]$ and $R_a[V]$ satisfy the following properties:*

1. *For any $y, r \in \mathbb{R}^m$, $R^a[U](y,r) \ge U(y)$ and $R_a[V](y,r) \le V(y)$.*

2. *For any $y \in \mathbb{R}^m$ and $k > 0$, there exists $\bar{a} = \bar{a}(y, \bar{k})$ such that, for $0 < a \le \bar{a}$, $R^a[U](\cdot, r)$ is semi-convex in $B(y, \bar{k})$ (respectively, $R_a[V](\cdot, r)$ is semi-concave in $B(y, \bar{k})$).*

3. *Assume that $U \in C^2(\mathbb{R}^m)$ (respectively $V \in C^2(\mathbb{R}^m)$). For any $y \in \mathbb{R}^m$ and $\bar{k} > 0$, there exists $\bar{a} = \bar{a}(y, \bar{k})$ such that, for $0 < a \le \bar{a}$, $R^a[U]$ (respectively $R_a[V]$) is C^2 in $B(0, \bar{k})$. Moreover, $R^a[U]$ (respectively $R_a[V]$) converges towards U (respectively V) in $C^2(B(0, \bar{k})$ as $a \to 0$.*

4. *If $R^a[U](y,r) = U(\bar{y}) - r \cdot (\bar{y} - y) - \frac{|\bar{y}-y|^2}{2a}$ and if $|\bar{y} - y| < 1$, then*

$$(s, A) \in J^+R^a[U](y,r) \Rightarrow (s, A) \in J^+U(\bar{y}) \text{ and } s = r - \frac{\bar{y} - y}{a}, \tag{4.4}$$

$$(r, A) \in \overline{D}^{2,+} R^a[U](y,r) \Rightarrow (s, A) \in \overline{D}^{2,+} U(y). \tag{4.5}$$

Lemma 4.3 (Nonlocal Jensen-Ishii's Lemma, Lemma 1 of Barles and Imbert [2])**.** *Let u and v be respectively an upper-semicontinuous and a lower-semicontinuous function defined on $\mathbb{R}^d$, and let ϕ be a C^2 function defined on $\mathbb{R}^{2d}$. If $(\bar{x}, \bar{y}) \in \mathbb{R}^{2d}$ is a zero global maximum point of $u(x) - v(y) - \phi(x,y)$ and if $p := D_x\phi(\bar{x}, \bar{y})$, $q := D_y\phi(\bar{x}, \bar{y})$, then the following hold:*

$$u(x) - v(y) \le R^a[u](x,p) - R_a[v](y, -q) \le R^a[\phi]((x,y),(p,q)),$$

$$\begin{aligned} u(\bar{x}) &= R^a[u](\bar{x}, p), \\ v(\bar{y}) &= R_a[v](\bar{y}, -q), \\ \textit{and} \quad R^a[\phi]((\bar{x}, \bar{y}), (p, q)) &= \phi(\bar{x}, \bar{y}). \end{aligned}$$

Moreover, for any $\bar{k} > 0$, there exists $\bar{a}(\bar{k}) > 0$ such that, for any $0 \leq a \leq \bar{a}(\bar{k})$, we have that there exist sequences $x_k \to x$, $y_k \to y$, $p_k \to p$, and $q_k \to q$, matrices X_k and Y_k, and a sequence of functions $\phi_k \in C^2(B((\bar{x}, \bar{y}), \bar{k}))$ converging uniformly to $\phi_a := R^a[\phi]((x, y), (p, q))$ such that

$$\begin{aligned} (x_k, y_k) \textit{ is a global maximum point of } u - v - \phi_k \\ u(x_k) \to u(\bar{x}) \textit{ and } v(y_k) \to v(\bar{y}) \\ (p_k, X_k) \in J^+u(x_k) \\ (-q_k, Y_k) \in J^-v(y_k) \\ -\frac{1}{a}I \leq \begin{bmatrix} X_k & 0 \\ 0 & -Y_k \end{bmatrix} \leq D^2\phi_k(x_k, y_k). \end{aligned}$$

Moreover, $p_k = D_x\phi_k(x_k, y_k)$, $q_k = D_y\phi_k(x_k, y_k)$, $\phi_a(\bar{x}, \bar{y}) = \phi(\bar{x}, \bar{y})$, and $D\phi_a(\bar{x}, \bar{y}) = D\phi(x, y)$.

Finally, the main result needed for the comparison principle discussed in the next section is Corollary 1 from [2]. Here we assume that we can decompose the operator $\mathcal{I}$ into

$$\mathcal{I}[\mathbf{x}, f] = \mathcal{I}^{1,\delta}[\mathbf{x}, f] + \mathcal{I}^{2,\delta}[\mathbf{x}, \nabla f, f].$$

(This decomposition will be discussed in greater detail in the following section.)

Corollary 4.4 (Corollary 1 of Barles and Imbert [2])**.** *Let U be an upper-semicontinuous viscosity solution of (3.2), let V be a lower-semicontinuous viscosity solution of (3.2), and let $\phi \in C^2(\mathbb{R}^d)$. If $(\bar{x}, \bar{y}) \in \mathbb{R}^{2d}$ is a global maximum point of $U(x) - V(y) - \phi(x, y)$, then, for any $\delta > 0$, there exists $\bar{a}$ such that, for $0 < a < \bar{a}$, we have*

$$F(\bar{x}, U(\bar{x}), p, X, \mathcal{I}^{1,\delta}[\bar{x}, \phi_a(\cdot, \bar{y})] + \mathcal{I}^{2,\delta}[\bar{x}, p, U]) \leq 0 \tag{4.6}$$

$$F(\bar{y}, V(\bar{y}), q, Y, \mathcal{I}^{1,\delta}[\bar{y}, -\phi_a(\bar{x}, \cdot)] + \mathcal{I}^{2,\delta}[\bar{y}, q, V]) \geq 0 \tag{4.7}$$

where $p = \nabla_x\phi_a(\bar{x}, \bar{y})$, $q = -\nabla_y\phi(\bar{x}, \bar{y}) = \nabla_y\phi_a(\bar{x}, \bar{y})$, and

$$-\frac{1}{a}I \leq \begin{bmatrix} X & 0 \\ 0 & -Y \end{bmatrix} \leq D^2\phi_a(\bar{x}, \bar{y}) = D^2\phi(\bar{x}, \bar{y}) + o_a(1). \tag{4.8}$$

Remark 4.5. Barles and Imbert note in Section 5.2 of [2] that the conditions of their comparison principle require a certain interaction between the first derivative term in the generator and the location of the process. Because exponential mean reversion has a very strong interaction between these two components, the conditions in their paper are not met by the spot price process under consideration.

1.2 Specialized Comparison Principle

The result shown in this section is a specialization of that found in Barles and Imbert [2] which allows for a relaxation of one of the assumptions of the main result of that paper due to the greater regularity of the nonlocal operator used in our problem compared with that in the original paper. This relaxation allows the uniqueness result to be applied to mean-reverting processes, as well.

1.2.1 Assumptions

First, we have the following general ellipticity assumption on the function F,

- (E) For any $\mathbf{x} \in \mathbb{R}^d$; $g \in \mathbb{R}$; $p \in \mathbb{R}^d$; $M, N \in \mathbb{S}_d$; and $l_1, l_2 \in \mathbb{R}^d$,

$$F(x, g, p, M, l_1) \leq F(x, g, p, N, l_2) \quad \text{if} \quad M \geq N \text{ and } l_1 \geq l_2. \tag{4.9}$$

We also make a series of assumptions about a decomposition of the nonlocal term $\mathcal{I}[x, f]$, all of which are combined under assumption (NLT) below.

- (NLT) For any $\delta > 0$, there exist operators $\mathcal{I}^{1,\delta}[x, \phi]$ and $\mathcal{I}^{2,\delta}[x, \nabla\phi(x), \phi]$ which are well-defined for any $x \in \mathbb{R}^d$ and $\phi \in \mathcal{C} \cap C^2(\mathbb{R}^d)$ and which satisfy the following:
  - ◇ For any $x \in \mathbb{R}^d$ and $\phi \in \mathcal{C} \cap C^2(\mathbb{R}^d)$, $\mathcal{I}[x, \phi] = \mathcal{I}^{1,\delta}[x, \phi] + \mathcal{I}^{2,\delta}[x, \nabla\phi(x), \phi]$. Moreover, for any $a \in \mathbb{R}$, $\mathcal{I}^{1,\delta}[x, \phi + a] = \mathcal{I}^{1,\delta}[x, \phi]$ and $\mathcal{I}^{2,\delta}[x, \nabla\phi(x), \phi + a] = \mathcal{I}^{2,\delta}[x, \nabla\phi(x), \phi]$.
  - ◇ There exists $R_\delta > 0$ with $R_\delta \to 0$ as $\delta \to 0$ such that, if $\phi_1 = \phi_2$ on $B(x, R_\delta)$, then $\mathcal{I}^{1,\delta}[x, \phi_1] = \mathcal{I}^{1,\delta}[x, \phi_2]$ (and respectively if $\phi_1 = \phi_2$ on $\mathbb{R}^d \setminus B(x, R_\delta)$, $\mathcal{I}^{2,\delta}[x, \nabla\phi_1(x), \phi_1] = \mathcal{I}^{2,\delta}[x, \nabla\phi_2(x), \phi_2]$).
  - ◇ For any $\phi \in C^2(\mathbb{R}^d)$ and $g \in \mathcal{C}$ such that $g - \phi$ attains a maximum at x on $B(x, R_\delta)$, there exists a sequence $\phi_k \in \mathcal{C} \cap C^2(\mathbb{R}^d)$ such that $g - \phi_k$ attains a

global maximum at x,

$$\mathcal{I}^{1,\delta}[x,\phi_k] \to \mathcal{I}^{1,\delta}[x,\phi] \quad \text{as} \;\; k\to\infty,$$

and

$$\mathcal{I}^{2,\delta}[x,\nabla\phi_k,\phi_k] \to \mathcal{I}^{2,\delta}[x,\nabla\phi,\phi] \quad \text{as} \;\; k\to\infty.$$

◇ The operator $\mathcal{I}^{1,\delta}[x,\phi]$ is well-defined for any $x\in\mathbb{R}^d$ and $\phi\in C^2(B(x,r))\cap\mathcal{C}$ for any $r<R_\delta$. Moreover, $\mathcal{I}^{1,\delta}[x,\phi]\to 0$ as $\delta\to 0$ and $\mathcal{I}^{1,\delta}[x_k,\phi_k]\to\mathcal{I}^{1,\delta}[x,\phi]$ when $x_k\to x$ and $\phi_k\to\phi$ in $C^2(B(x,r))\cap C(\overline{B(x,r)})$.

◇ The operator $\mathcal{I}^{2,\delta}[x,p,\phi]$ is defined for any $x\in\mathbb{R}^d$ and $\phi\in\mathcal{C}$. Moreover, if $x_k\to x$, $p_k\to p$, and $(\phi_k)_k$ is a sequence of uniformly locally bounded functions such that $|\phi_k|\le\psi$ with $\psi\in\mathcal{C}$,

$$\limsup_{k\to\infty}\mathcal{I}^{2,\delta}[x_k,p_k,\phi_k]\le\mathcal{I}^{2,\delta}[x,p,\overline{\phi}] \quad \text{where} \;\; \overline{\phi}:=\limsup{}^*\phi_k$$

and

$$\liminf_{k\to\infty}\mathcal{I}^{2,\delta}[x_k,p_k,\phi_k]\ge\mathcal{I}^{2,\delta}[x,p,\underline{\phi}] \quad \text{where} \;\; \underline{\phi}:=\liminf{}_*\phi_k.$$

These two assumptions mirror precisely those in [2].

The remaining assumptions are nearly the same as those made by Barles and Imbert except for (A3*) where a relaxation is required for this problem and (A1*) where it is notationally convenient (though not actually necessary for the comparison result) to use a slightly stronger assumption since the process being considered in this paper is more regular than a general Lévy-Itô process. We assume here that the nonlocal operator $\mathcal{I}$ can be written in terms of a jump size function $j:\mathbb{R}^d\times\mathbb{R}^d\to\mathbb{R}^d$ and a jump measure μ in the following way:

$$\mathcal{I}[(x,z),g]=\int\left(g((x,z)+j((x,z),(y,w)))-g(x,z)\right)\mu(dy,dw). \tag{4.10}$$

We then make the following assumptions on μ, j, and F:

- (A1*) The measure $\mu(dy,dw)$ and the function $j(x,z)$ satisfy

$$\int_{\mathbb{R}^d}\mu(dy,dw)<\infty \;\;\text{and}\;\; \sup_{(x,z)\in\mathcal{D}}\int_{\mathbb{R}^d}\left|j((x,z),(y,w))\right|^2\,\mu(dy,dw)<\infty, \tag{4.11}$$

and there exists a constant $\bar{c}>0$ such that

$$\int_{\mathbb{R}^d}\left|j((x,z),(y,w))-j((\tilde{x},\tilde{z}),(y,w))\right|^2\mu(dy,dw)\le\bar{c}|(x,z)-(\tilde{x},\tilde{z})|^2 \;\;\text{and}$$

$$\int_{\mathbb{R}^d} |j((x,z),(y,w)) - j((\tilde{x},\tilde{z}),(y,w))| \; \mu(dy,dw) \le \bar{c}|(x,z) - (\tilde{x},\tilde{z})|. \tag{4.12}$$

(Note that the last two points of this version of the assumption are a simple consequence of the Cauchy-Schwartz Inequality given the modification to the first two points used here.)

- (A2) There exists $\gamma > 0$ such that for any $(x,z) \in \mathcal{D}$, $g, h \in \mathbb{R}$, $p \in \mathbb{R}^d$, $X \in \mathbb{S}_2$, and $l \in \mathbb{R}$,
$$F(x,g,p,X,l) - F(x,h,p,X,l) \ge \gamma(g-h) \quad \text{when} \quad g \ge h. \tag{4.13}$$

- (A3*) Given a fixed constant $\rho > 0$, for any $\beta > 0$, there exist moduli of continuity ω and ω_β such that, for any $|(x,z)|, |(\tilde{x},\tilde{z})| \le \frac{\rho}{\beta}$, $|h| \le \frac{\rho}{\beta}$, $l \in \mathbb{R}$, and for any $X, Y \in \mathbb{S}_2$ satisfying
$$\begin{bmatrix} X & 0 \\ 0 & -Y \end{bmatrix} \le \frac{1}{\epsilon} \begin{bmatrix} I & -I \\ -I & I \end{bmatrix} + r(\beta) \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \tag{4.14}$$
for some $\epsilon > 0$ and $r(\beta) \to 0$ as $\beta \to 0$, then, if $\frac{\rho}{\beta}s(\beta) \to 0$ as $\beta \to 0$, we have
$$\begin{aligned} F((\tilde{x},\tilde{z}), h, \tfrac{1}{\epsilon}((x,z) - (\tilde{x},\tilde{z})), Y, l) - F((x,z), h, \tfrac{1}{\epsilon}((x,z) - (\tilde{x},\tilde{z})) + s(\beta), X, l) \\ \le \omega(\beta) + \omega_\beta(|(x,z) - (\tilde{x},\tilde{z})| + \tfrac{1}{\epsilon}|(x,z) - (\tilde{x},\tilde{z})|^2). \end{aligned} \tag{4.15}$$

- (A4) $F((x,z), g, p, X, l)$ is Lipschitz continuous in l, uniformly with respect to all the other variables.

Remark 4.6. Note that assumption (A3*) is a subtle but significant relaxation of assumption (A3-1) in [2] where the same is required for all $s(\beta) \to 0$ rather than just those $s(\beta)$ with $\frac{\rho}{\beta}s(\beta) \to 0$. We also replace an arbitrary constant, R, used in the original by $\frac{\rho}{\beta}$. But since this is the only form used in the version of the original proof using (A3-1) in [2], which is the version we have modified here, this change does not amount to any further restriction than is actually put to use there.

1.2.2 The Specialized Comparison Principle

Theorem 4.7. *(Specialization of the Comparison Principle) Suppose that the PIDE (4.1) satisfies assumptions (A1*), (A2), (A3*), and (A4) as well as assumptions (E) and (NLT) above. If g is a bounded upper semi-continuous viscosity subsolution of (4.1) and h is a bounded lower semi-continuous viscosity supersolution of (4.1), then $g \le h$ on $\mathcal{D}$.*

Proof. For the sake of clarity in the longer expressions in this proof, we will use boldface letters to represent points in $\mathcal{D}$ or $\mathbb{R}^2$, as appropriate. For instance, where we have previously referred to points $(x, z) \in \mathcal{D}$, we will instead use $\mathbf{x} \in \mathcal{D}$, and we will notate a value $(y, w) \in \mathbb{R}^2$ representing a jump size by $\mathbf{y} \in \mathbb{R}^2$.

Define $M := \sup_{\mathcal{D}}(g - h)$. Assume by way of contradiction that $M > 0$.

We now use a dedoubling of variables technique similar to that in the classical viscosity solution theory to approximate M. First, we define $\mathcal{R} := (||g||_\infty + ||h||_\infty)$. Let $\psi : \mathbb{R} \to \mathbb{R}^+$ be a smooth function such that $\psi(\mathbf{x}) = 0$ for $|\mathbf{x}| \leq 1$, $\psi(\mathbf{x}) = \mathcal{R} + 1$ for $|\mathbf{x}| \geq 2$, and ψ, $\nabla\psi$, and $D^2\psi$ are all bounded. Then, for $\beta > 0$, define $\psi_\beta : \mathbb{R}^2 \to \mathbb{R}$ by $\psi_\beta(\mathbf{x}) = \psi(\beta^3\mathbf{x})$. We note the following important properties of such functions ψ_β. By the chain rule, $\frac{1}{\beta}\nabla\psi_\beta \to 0$ and $D^2\psi_\beta \to 0$ as $\beta \to 0$ uniformly on $\mathbb{R}^2$. Also, it can be shown that $\mathcal{I}[\psi_\beta, \mathbf{x}] \to 0$ as $\beta \to 0$ uniformly on $\mathbb{R}^2$.

We now approximate M by

$$M_{\epsilon,\beta} := \sup_{\mathbf{x}_1,\mathbf{x}_2 \in \mathcal{D}} \left\{ g(\mathbf{x}_1) - h(\mathbf{x}_2) - \frac{|\mathbf{x}_1 - \mathbf{x}_2|^2}{2\epsilon} - \psi_\beta(\mathbf{x}_1) \right\} \tag{4.16}$$

where ϵ and β are small parameters which will tend to 0. Since $\psi_\beta(\mathbf{x}_1) > \mathcal{R}$ when $|\mathbf{x}_1| \geq \frac{2}{\beta^3}$, the supremum above is achieved and is thus actually a maximum.

Consider any maximum point $(\bar{\mathbf{x}}_1, \bar{\mathbf{x}}_2)$ of the function

$$g(\mathbf{x}_1) - h(\mathbf{x}_2) - \frac{|\mathbf{x}_1 - \mathbf{x}_2|^2}{2\epsilon} - \psi_\beta(\mathbf{x}_1).$$

For ϵ and β sufficiently small,

$$0 < \frac{M}{2} \leq M_{\epsilon,\beta} \leq g(\bar{\mathbf{x}}_1) - h(\bar{\mathbf{x}}_2) \quad \text{and} \quad \frac{|\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2|}{\epsilon} \leq \frac{C}{\sqrt{\epsilon}} \quad \text{and} \quad \psi_\beta(\bar{\mathbf{x}}_1) \leq \mathcal{R} + 1.$$

Further, for any $\mathbf{d}, \mathbf{d}' \in \mathbb{R}^2$, we have

$$\begin{aligned} g(\bar{\mathbf{x}}_1 + \mathbf{d}) - h(\bar{\mathbf{x}}_2 + \mathbf{d}') - &\frac{|\bar{\mathbf{x}}_1 + \mathbf{d} - \bar{\mathbf{x}}_2 - \mathbf{d}'|^2}{2\epsilon} - \psi_\beta(\bar{\mathbf{x}}_1 + \mathbf{d}) \\ &\leq g(\bar{\mathbf{x}}_1) - h(\bar{\mathbf{x}}_2) - \frac{|\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2|^2}{2\epsilon} - \psi_\beta(\bar{\mathbf{x}}_1). \end{aligned}$$

In particular, let $\mathbf{y} \in \mathbb{R}^2$ be arbitrary, and set $d = j(\bar{\mathbf{x}}_1, \mathbf{y})$ and $d' = j(\bar{\mathbf{x}}_2, \mathbf{y})$. Then, rearranging terms in the above inequality, we have

$$g(\bar{\mathbf{x}}_1 + j(\bar{\mathbf{x}}_1, \mathbf{y})) - g(\bar{\mathbf{x}}_1) \leq h(\bar{\mathbf{x}}_2 + j(\bar{\mathbf{x}}_2, \mathbf{y})) - h(\bar{\mathbf{x}}_2) + \frac{|\bar{\mathbf{x}}_1 + j(\bar{\mathbf{x}}_1, \mathbf{y}) - \bar{\mathbf{x}}_2 - j(\bar{\mathbf{x}}_2, \mathbf{y})|^2}{2\epsilon}$$

$$
\begin{aligned}
&\qquad - \frac{|\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2|^2}{2\epsilon} + \psi_\beta(\bar{\mathbf{x}}_1 + j(\bar{\mathbf{x}}_1, \mathbf{y})) - \psi_\beta(\bar{\mathbf{x}}_1) \\
&\le h(\bar{\mathbf{x}}_2 + j(\bar{\mathbf{x}}_2, \mathbf{y})) - h(\bar{\mathbf{x}}_2) + \frac{|j(\bar{\mathbf{x}}_1, \mathbf{y}) - j(\bar{\mathbf{x}}_2, \mathbf{y})|^2}{2\epsilon} \\
&\qquad + \frac{[\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2] \cdot [(j(\bar{\mathbf{x}}_1, \mathbf{y}) - j(\bar{\mathbf{x}}_2, \mathbf{y})]}{\epsilon} \\
&\qquad + \psi_\beta(\bar{\mathbf{x}}_1 + j(\bar{\mathbf{x}}_1, \mathbf{y})) - \psi_\beta(\bar{\mathbf{x}}_1). \qquad (4.17)
\end{aligned}
$$

Define $\phi(\mathbf{x}_1, \mathbf{x}_2) := \frac{|\mathbf{x}_1 - \mathbf{x}_2|^2}{2\epsilon} + \psi_\beta(\mathbf{x}_1)$. Further, define $\phi_1(\mathbf{x}_1) := \phi(\mathbf{x}_1, \bar{\mathbf{x}}_2)$ and $\phi_2(\mathbf{x}_2) := \phi(\bar{\mathbf{x}}_1, \mathbf{x}_2)$. Then, returning for a moment to the notation $\bar{\mathbf{x}}_1 = (\bar{x}_1, \bar{z}_1)$, $\bar{\mathbf{x}}_2 = (\bar{x}_2, \bar{z}_2)$, and $\mathbf{y} = (y, w)$, we have

$$
\begin{aligned}
\mathcal{I}^{1,\delta}[\bar{\mathbf{x}}_1, \phi_1] &= \eta_1 \int_0^\delta \Bigg[\frac{|(\bar{x}_1 + y, \bar{z}_1) - (\bar{x}_2, \bar{z}_2)|^2}{2\epsilon} + \psi_\beta((\bar{x}_1 + y, \bar{z}_1)) \\
&\qquad\qquad - \frac{|(\bar{x}_1, \bar{z}_1) - (\bar{x}_2, \bar{z}_2)|^2}{2\epsilon} - \psi_\beta((\bar{x}_1, \bar{z}_1)) \Bigg] \beta_1 e^{-\beta_1 y} dy \\
&\quad + \eta_2 \int_{-\delta}^0 \Bigg[\frac{|(\bar{x}_1 + y, \bar{z}_1) - (\bar{x}_2, \bar{z}_2)|^2}{2\epsilon} + \psi_\beta((\bar{x}_1 + y, \bar{z}_1)) \\
&\qquad\qquad - \frac{|(\bar{x}_1, \bar{z}_1) - (\bar{x}_2, \bar{z}_2)|^2}{2\epsilon} - \psi_\beta((\bar{x}_1, \bar{z}_1)) \Bigg] \beta_2 e^{\beta_2 y} dy \\
&= \eta_1 \int_0^\delta \left[\frac{|(\bar{x}_1 + y, \bar{z}_1) - (\bar{x}_2, \bar{z}_2)|^2}{2\epsilon} - \frac{|(\bar{x}_1, \bar{z}_1) - (\bar{x}_2, \bar{z}_2)|^2}{2\epsilon} \right] \beta_1 e^{-\beta_1 y} dy \\
&\qquad\qquad + \eta_1 \int_0^\delta \Big[\psi_\beta((\bar{x}_1 + y, \bar{z}_1)) - \psi_\beta((\bar{x}_1, \bar{z}_1)) \Big] \beta_1 e^{-\beta_1 y} dy \\
&\quad + \eta_2 \int_{-\delta}^0 \left[\frac{|(\bar{x}_1 + y, \bar{z}_1) - (\bar{x}_2, \bar{z}_2)|^2}{2\epsilon} - \frac{|(\bar{x}_1, \bar{z}_1) - (\bar{x}_2, \bar{z}_2)|^2}{2\epsilon} \right] \beta_2 e^{\beta_2 y} dy \\
&\qquad\qquad + \eta_2 \int_{-\delta}^0 \Big[\psi_\beta((\bar{x}_1 + y, \bar{z}_1)) - \psi_\beta((\bar{x}_1, \bar{z}_1)) \Big] \beta_2 e^{\beta_2 y} dy \\
&= \eta_1 \int_0^\delta \left[\frac{|(\bar{x}_1 + y, \bar{z}_1) - (\bar{x}_2, \bar{z}_2)|^2}{2\epsilon} - \frac{|(\bar{x}_1, \bar{z}_1) - (\bar{x}_2, \bar{z}_2)|^2}{2\epsilon} \right] \beta_1 e^{-\beta_1 y} dy \\
&\quad + \eta_2 \int_{-\delta}^0 \left[\frac{|(\bar{x}_1 + y, \bar{z}_1) - (\bar{x}_2, \bar{z}_2)|^2}{2\epsilon} - \frac{|(\bar{x}_1, \bar{z}_1) - (\bar{x}_2, \bar{z}_2)|^2}{2\epsilon} \right] \beta_2 e^{\beta_2 y} dy \\
&\quad + \mathcal{I}^{1,\delta}[(\bar{x}_1, \bar{z}_1), \psi_\beta].
\end{aligned}
$$

Applying the triangle inequality to the above, we can further estimate

$$
\begin{aligned}
\mathcal{I}^{1,\delta}[\bar{\mathbf{x}}_1, \phi_1] &= \eta_1 \int_0^\delta \left[\frac{|(\bar{x}_1 + y, \bar{z}_1) - (\bar{x}_2, \bar{z}_2)|^2}{2\epsilon} - \frac{|(\bar{x}_1, \bar{z}_1) - (\bar{x}_2, \bar{z}_2)|^2}{2\epsilon} \right] \beta_1 e^{-\beta_1 y} dy \\
&\quad + \eta_2 \int_{-\delta}^0 \left[\frac{|(\bar{x}_1 + y, \bar{z}_1) - (\bar{x}_2, \bar{z}_2)|^2}{2\epsilon} - \frac{|(\bar{x}_1, \bar{z}_1) - (\bar{x}_2, \bar{z}_2)|^2}{2\epsilon} \right] \beta_2 e^{\beta_2 y} dy
\end{aligned}
$$

$$
\begin{aligned}
&+ \mathcal{I}^{1,\delta}[(\bar{x}_1,\bar{z}_1),\psi_\beta] \\
&= \eta_1 \int_0^\delta \left[\frac{|(\bar{x}_1,\bar{z}_1)+(y,0)-(\bar{x}_2,\bar{z}_2)|^2}{2\epsilon} - \frac{|(\bar{x}_1,\bar{z}_1)-(\bar{x}_2,\bar{z}_2)|^2}{2\epsilon} \right] \beta_1 e^{-\beta_1 y} dy \\
&\quad + \eta_2 \int_{-\delta}^0 \left[\frac{|(\bar{x}_1,\bar{z}_1)+(y,0)-(\bar{x}_2,\bar{z}_2)|^2}{2\epsilon} - \frac{|(\bar{x}_1,\bar{z}_1)-(\bar{x}_2,\bar{z}_2)|^2}{2\epsilon} \right] \beta_2 e^{\beta_2 y} dy \\
&\quad + \mathcal{I}^{1,\delta}[(\bar{x}_1,\bar{z}_1),\psi_\beta] \\
&\leq \eta_1 \int_0^\delta \left[\frac{|(\bar{x}_1,\bar{z}_1)-(\bar{x}_2,\bar{z}_2)|^2}{2\epsilon} + \frac{|(y,0)|^2}{2\epsilon} - \frac{|(\bar{x}_1,\bar{z}_1)-(\bar{x}_2,\bar{z}_2)|^2}{2\epsilon} \right] \beta_1 e^{-\beta_1 y} dy \\
&\quad + \eta_2 \int_{-\delta}^0 \left[\frac{|(\bar{x}_1,\bar{z}_1)-(\bar{x}_2,\bar{z}_2)|^2}{2\epsilon} + \frac{|(y,0)|^2}{2\epsilon} - \frac{|(\bar{x}_1,\bar{z}_1)-(\bar{x}_2,\bar{z}_2)|^2}{2\epsilon} \right] \beta_2 e^{\beta_2 y} dy \\
&\quad + \mathcal{I}^{1,\delta}[(\bar{x}_1,\bar{z}_1),\psi_\beta] \\
&= \eta_1 \int_0^\delta \left[\frac{y^2}{2\epsilon} \right] \beta_1 e^{-\beta_1 y} dy + \eta_2 \int_{-\delta}^0 \left[\frac{y^2}{2\epsilon} \right] \beta_2 e^{\beta_2 y} dy + \mathcal{I}^{1,\delta}[(\bar{x}_1,\bar{z}_1),\psi_\beta] \\
&= \frac{1}{\epsilon} \left(\frac{\eta_1}{2} \int_0^\delta y^2 \beta_1 e^{-\beta_1 y} dy + \frac{\eta_2}{2} \int_{-\delta}^0 y^2 \beta_2 e^{\beta_2 y} dy \right) + \mathcal{I}^{1,\delta}[(\bar{x}_1,\bar{z}_1),\psi_\beta].
\end{aligned}
$$

Finally, since $\int_0^\delta y^2 \beta_1 e^{-\beta_1 y} dy \to 0$ and $\int_{-\delta}^0 y^2 \beta_2 e^{\beta_2 y} dy \to 0$ as $\delta \to 0$, we have that

$$
\mathcal{I}^{1,\delta}[\bar{\mathbf{x}}_1,\phi_1] \leq \frac{1}{\epsilon} o_\delta(1) + \mathcal{I}^{1,\delta}[(\bar{x}_1,\bar{z}_1),\psi_\beta]. \tag{4.18}
$$

A similar computation shows that

$$
\mathcal{I}^{1,\delta}[\bar{\mathbf{x}}_2,-\phi_2] + \frac{1}{\epsilon} \int_0^\delta y^2 \beta_1 e^{-\beta_1 y} dy \geq \frac{1}{2\epsilon} \int_0^\delta y^2 \beta_1 e^{-\beta_1 y} dy \geq 0.
$$

Combining these inequalities with assumption (A1*), we have

$$
\begin{aligned}
\mathcal{I}^{1,\delta}[\bar{\mathbf{x}}_1,\phi_1] &\leq \mathcal{I}^{1,\delta}[\bar{\mathbf{x}}_2,-\phi_2] + \frac{1}{\epsilon} \int_0^\delta y^2 \beta_1 e^{-\beta_1 y} dy + \frac{1}{2\epsilon} \int_0^\delta y^2 \beta_1 e^{-\beta_1 y} dy + \mathcal{I}^{1,\delta}[(\bar{x},\bar{z}),\psi_\beta] \\
&\leq \mathcal{I}^{1,\delta}[\bar{\mathbf{x}}_2,-\phi_2] + \frac{1}{\epsilon} o_\delta(1) + o_\beta(1).
\end{aligned} \tag{4.19}
$$

We now develop a similar estimate on $\mathcal{I}^{2,\delta}[\bar{\mathbf{x}}_1, g]$. We again return to the two-coordinate notation for a moment (e.g. $\mathbf{x} = (x,z)$). First, recall that $j((x,z),(y,w)) = (y, w \wedge (z_{max} - z))$ for any $(x,z) \in \mathcal{D}$ and $(y,w) \in \mathbb{R}^2$ and thus

$$
\begin{aligned}
j(\bar{\mathbf{x}}_1,\mathbf{y}) - j(\bar{\mathbf{x}}_2,\mathbf{y}) &= \big(y - y, (w \wedge (z_{max} - \bar{z}_1)) - (w \wedge (z_{max} - \bar{z}_2))\big) \\
&= \big(0, (w \wedge (z_{max} - \bar{z}_1)) - (w \wedge (z_{max} - \bar{z}_2))\big).
\end{aligned}
$$

By direct calculation of each of the four possible cases, it can be shown that

$$
\begin{aligned}
[\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2] \cdot [(j(\bar{\mathbf{x}}_1, \mathbf{y}) - j(\bar{\mathbf{x}}_2, \mathbf{y})] &= (\bar{x}_1 - \bar{x}_2, \bar{z}_1 - \bar{z}_2) \\
&\qquad \cdot \big(0, (w \wedge (z_{max} - \bar{z}_1)) - (w \wedge (z_{max} - \bar{z}_2))\big) \\
&= 0 + (\bar{z}_1 - \bar{z}_2)[(w \wedge (z_{max} - \bar{z}_1)) - (w \wedge (z_{max} - \bar{z}_2))] \\
&\leq (\bar{z}_1 - \bar{z}_2)^2.
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
\frac{1}{\epsilon} \int_{\mathbb{R}^2 \setminus B(\bar{\mathbf{x}}_1, R_\delta)} [\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2] \cdot [(j(\bar{\mathbf{x}}_1, \mathbf{y}) - j(\bar{\mathbf{x}}_2, \mathbf{y})] \, \mu(dy, dw) &\leq \frac{1}{\epsilon} \int_{\mathbb{R}^2 \setminus B(\bar{\mathbf{x}}_1, R_\delta)} (\bar{z}_1 - \bar{z}_2)^2 \mu(dy, dw) \\
&= \frac{1}{\epsilon} (\bar{z}_1 - \bar{z}_2)^2 \int_{\mathbb{R}^2 \setminus B(\bar{\mathbf{x}}_1, R_\delta)} \mu(dy, dw) \\
&= \frac{1}{\epsilon} (\bar{z}_1 - \bar{z}_2)^2 \\
&\leq O\left(\frac{|\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2|^2}{\epsilon} \right). \qquad (4.20)
\end{aligned}
$$

By a similar calculation,

$$
\frac{1}{2\epsilon} \int_{\mathbb{R}^2 \setminus B(\bar{\mathbf{x}}_1, R_\delta)} |j(\bar{\mathbf{x}}_1, \mathbf{y}) - j(\bar{\mathbf{x}}_2, \mathbf{y})|^2 \mu(dy, dw) \leq \frac{1}{2\epsilon} \int_{\mathbb{R}^2 \setminus B(\bar{\mathbf{x}}_1, R_\delta)} (\bar{z}_1 - \bar{z}_2)^2 \mu(dy, dw),
$$

and thus

$$
\frac{1}{2\epsilon} \int_{\mathbb{R}^2 \setminus B(\bar{\mathbf{x}}_1, R_\delta)} |j(\bar{\mathbf{x}}_1, \mathbf{y}) - j(\bar{\mathbf{x}}_2, \mathbf{y})|^2 \mu(dy, dw) \leq O\left(\frac{|\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2|^2}{\epsilon} \right). \qquad (4.21)
$$

Now, integrating on both sides of inequality (4.17) on $\mathbb{R}^2 \setminus B(\bar{\mathbf{x}}_1, \delta)$ and applying inequalities (4.20) and (4.21), we get

$$
\begin{aligned}
\mathcal{I}^{2,\delta}[\bar{\mathbf{x}}_1, g] &\leq \mathcal{I}^{2,\delta}[\bar{\mathbf{x}}_2, h] + \frac{1}{2\epsilon} \int_{\mathbb{R}^2 \setminus B(\bar{\mathbf{x}}_1, \delta)} [\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2] \cdot [(j(\bar{\mathbf{x}}_1, \mathbf{y}) - j(\bar{\mathbf{x}}_2, \mathbf{y})] \, \mu(dy, dw) \\
&\qquad + \frac{1}{2\epsilon} \int_{\mathbb{R}^2 \setminus B(\bar{\mathbf{x}}_1, \delta)} |j(\bar{\mathbf{x}}_1, \mathbf{y}) - j(\bar{\mathbf{x}}_2, \mathbf{y})|^2 \mu(dy, dw) + \mathcal{I}^{2,\delta}[\bar{\mathbf{x}}_1, \psi_\beta] \\
&\leq \mathcal{I}^{2,\delta}[\bar{\mathbf{x}}_2, h] + O\left(\frac{|\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2|^2}{\epsilon} \right) + o_\beta(1). \qquad (4.22)
\end{aligned}
$$

Combining (4.19) and (4.22), we get an estimate for the nonlocal integral terms we wish to consider:

$$
\begin{aligned}
\ell := \mathcal{I}^{1,\delta}[\bar{\mathbf{x}}_1, \phi_1] + \mathcal{I}^{2,\delta}[\bar{\mathbf{x}}_1, g] \\
\leq \mathcal{I}^{1,\delta}[\bar{\mathbf{x}}_2, -\phi_2] + \mathcal{I}^{2,\delta}[\bar{\mathbf{x}}_2, h] + O\left(\frac{|\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2|^2}{\epsilon}\right) + o_\beta(1) + \frac{1}{\epsilon} o_\delta(1). \quad (4.23)
\end{aligned}
$$

Let $(p, -q) = \nabla\phi(\bar{x}, \bar{y})$ and define $\phi_a := R^a[\phi]((x,y),(p,q))$ as in (4.3). Then, by Corollary 4.4, for any $a > 0$, there exist matrices $X, Y \in \mathbb{S}_2$ such that (4.8) holds and

$$
F(\bar{\mathbf{x}}_1, g(\bar{\mathbf{x}}_1), p, X, \mathcal{I}^{1,\delta}[\bar{\mathbf{x}}_1, \phi_a(\cdot, \bar{\mathbf{x}}_2)] + \mathcal{I}^{2,\delta}[\bar{\mathbf{x}}_1, g])) \leq 0 \quad (4.24)
$$

$$
F(\bar{\mathbf{x}}_2, h(\bar{\mathbf{x}}_2), q, Y, \mathcal{I}^{1,\delta}[\bar{\mathbf{x}}_2, -\phi_a(\bar{\mathbf{x}}_1, \cdot)] + \mathcal{I}^{2,\delta}[\bar{\mathbf{x}}_2, h])) \geq 0 \quad (4.25)
$$

Now, applying Proposition 4.2, we get

$$
\begin{aligned}
F(\bar{\mathbf{x}}_1, g(\bar{\mathbf{x}}_1), p, X, \ell) \leq o_a(1) \\
F(\bar{\mathbf{x}}_2, h(\bar{\mathbf{x}}_2), q, Y, \mathcal{I}^{1,\delta}[\bar{\mathbf{x}}_2, -\phi_2] + \mathcal{I}[\bar{\mathbf{x}}_2, h]) \geq o_a(1)
\end{aligned}
$$

By assumption (A2), we have that there exists some $\gamma > 0$ such that

$$
\gamma\frac{M}{2} \leq F(\bar{\mathbf{x}}_1, g(\bar{\mathbf{x}}_1), p, X, \ell) - F(\bar{\mathbf{x}}_2, h(\bar{\mathbf{x}}_2), p, X, \ell) \quad (4.26)
$$

First, note that by definition, $p = q + \nabla\psi_\beta(\bar{\mathbf{x}}_1)$. So, the above inequality can be rewritten as

$$
\gamma\frac{M}{2} \leq F(\bar{\mathbf{x}}_1, g(\bar{\mathbf{x}}_1), p, X, \ell) - F(\bar{\mathbf{x}}_2, h(\bar{\mathbf{x}}_2), q + \nabla\psi_\beta(\bar{\mathbf{x}}_1), X, \ell) \quad (4.27)
$$

Now, applying inequalities (4.24) and (4.25) and adding and subtracting $F(\bar{\mathbf{x}}_2, h(\bar{\mathbf{x}}_2), q, Y, \ell)$, we get

$$
\begin{aligned}
\gamma\frac{M}{2} &\leq F(\bar{\mathbf{x}}_1, g(\bar{\mathbf{x}}_1), p, X, \ell) - F(\bar{\mathbf{x}}_2, h(\bar{\mathbf{x}}_2), q + \nabla\psi_\beta(\bar{\mathbf{x}}_1), X, \ell) \\
&\leq F(\bar{\mathbf{x}}_1, g(\bar{\mathbf{x}}_1), p, X, \ell) - F(\bar{\mathbf{x}}_2, h(\bar{\mathbf{x}}_2), q + \nabla\psi_\beta(\bar{\mathbf{x}}_1), X, \ell) \\
&\qquad + F(\bar{\mathbf{x}}_2, h(\bar{\mathbf{x}}_2), q, Y, \mathcal{I}^{1,\delta}[\bar{\mathbf{x}}_2, -\phi_a(\bar{\mathbf{x}}_1, \cdot)] + \mathcal{I}^{2,\delta}[\bar{\mathbf{x}}_2, h])) \\
&\qquad\quad - F(\bar{\mathbf{x}}_1, g(\bar{\mathbf{x}}_1), p, X, \mathcal{I}^{1,\delta}[\bar{\mathbf{x}}_1, \phi_a(\cdot, \bar{\mathbf{x}}_2)] + \mathcal{I}^{2,\delta}[\bar{\mathbf{x}}_1, g])) \\
&\qquad + F(\bar{\mathbf{x}}_2, h(\bar{\mathbf{x}}_2), q, Y, \ell) - F(\bar{\mathbf{x}}_2, h(\bar{\mathbf{x}}_2), q, Y, \ell) \quad (4.28)
\end{aligned}
$$

Rearranging the terms above, we will consider the inequality

$$
\begin{aligned}
\gamma\frac{M}{2} &\leq F(\bar{\mathbf{x}}_1, g(\bar{\mathbf{x}}_1), p, X, \ell) - F(\bar{\mathbf{x}}_2, h(\bar{\mathbf{x}}_2), q + \nabla\psi_\beta(\bar{\mathbf{x}}_1), X, \ell) \\
&\leq F(\bar{\mathbf{x}}_2, h(\bar{\mathbf{x}}_2), q, Y, \ell) - F(\bar{\mathbf{x}}_2, h(\bar{\mathbf{x}}_2), q + \nabla\psi_\beta(\bar{\mathbf{x}}_1), X, \ell) \\
&\qquad + F(\bar{\mathbf{x}}_2, h(\bar{\mathbf{x}}_2), q, Y, \mathcal{I}^{1,\delta}[\bar{\mathbf{x}}_2, -\phi_a(\bar{\mathbf{x}}_1, \cdot)] + \mathcal{I}^{2,\delta}[\bar{\mathbf{x}}_2, h])) - F(\bar{\mathbf{x}}_2, h(\bar{\mathbf{x}}_2), q, Y, \ell)
\end{aligned}
$$

$$+ F(\bar{\mathbf{x}}_1, g(\bar{\mathbf{x}}_1), p, X, \ell) - F(\bar{\mathbf{x}}_1, g(\bar{\mathbf{x}}_1), p, X, \mathcal{I}^{1,\delta}[\bar{\mathbf{x}}_1, \phi_a(\cdot, \bar{\mathbf{x}}_2)] + \mathcal{I}^{2,\delta}[\bar{\mathbf{x}}_1, g])) \tag{4.29}$$

We now examine each of the last two pairs of terms. By assumption (A4), we have that there exists a Lipschitz constant K_{Lip} such that

$$\begin{aligned}
&|F(\bar{\mathbf{x}}_1, g(\bar{\mathbf{x}}_1), p, X, \ell) - F(\bar{\mathbf{x}}_1, g(\bar{\mathbf{x}}_1), p, X, \mathcal{I}^{1,\delta}[\bar{\mathbf{x}}_1, \phi_a(\cdot, \bar{\mathbf{x}}_2)] + \mathcal{I}^{2,\delta}[\bar{\mathbf{x}}_1, g]))| \\
&\qquad \le K_{Lip} \, |\ell - (\mathcal{I}^{1,\delta}[\bar{\mathbf{x}}_1, -\phi_a(\cdot, \bar{\mathbf{x}}_2)] + \mathcal{I}^{2,\delta}[\bar{\mathbf{x}}_1, g])| \\
&\qquad = K_{Lip} |\mathcal{I}^{1,\delta}[\bar{\mathbf{x}}_1, \phi_x] - \mathcal{I}^{1,\delta}[\bar{\mathbf{x}}_2, -\phi_a(\bar{\mathbf{x}}_1, \cdot)]|.
\end{aligned}$$

Since Proposition 4.2 implies that $\mathcal{I}^{1,\delta}[\bar{\mathbf{x}}_2, -\phi_a(\bar{\mathbf{x}}_1, \cdot)] \to \mathcal{I}^{1,\delta}[\bar{\mathbf{x}}_1, \phi_x]$ as $a \to 0$, we have that

$$F(\bar{\mathbf{x}}_1, g(\bar{\mathbf{x}}_1), p, X, \ell) - F(\bar{\mathbf{x}}_1, g(\bar{\mathbf{x}}_1), p, X, \mathcal{I}^{1,\delta}[\bar{\mathbf{x}}_1, \phi_a(\cdot, \bar{\mathbf{x}}_2)] + \mathcal{I}^{2,\delta}[\bar{\mathbf{x}}_1, g])) \le o_a(1). \tag{4.30}$$

Turning to the other pair of terms from (4.29), we first note that, by (4.23) and the ellipticity assumption (E),

$$\begin{aligned}
F(\bar{\mathbf{x}}_2, h(\bar{\mathbf{x}}_2), q, Y, \ell) \ge F(\bar{\mathbf{x}}_2, h(\bar{\mathbf{x}}_2), q, Y, &\mathcal{I}^{1,\delta}[\bar{\mathbf{x}}_2, \phi_2] + \mathcal{I}^{2,\delta}[\bar{\mathbf{x}}_2, h] \\
&+ O\left(\frac{|\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2|^2}{\epsilon}\right) + o_\beta(1) + \frac{1}{\epsilon} o_\delta(1)).
\end{aligned}$$

Therefore,

$$\begin{aligned}
F(\bar{\mathbf{x}}_2, h(\bar{\mathbf{x}}_2), q, Y, &\mathcal{I}^{1,\delta}[\bar{\mathbf{x}}_2, -\phi_a(\bar{\mathbf{x}}_1, \cdot)] + \mathcal{I}^{2,\delta}[\bar{\mathbf{x}}_2, h])) - F(\bar{\mathbf{x}}_2, h(\bar{\mathbf{x}}_2), q, Y, \ell) \\
&\le F(\bar{\mathbf{x}}_2, h(\bar{\mathbf{x}}_2), q, Y, \mathcal{I}^{1,\delta}[\bar{\mathbf{x}}_2, -\phi_a(\bar{\mathbf{x}}_1, \cdot)] + \mathcal{I}^{2,\delta}[\bar{\mathbf{x}}_2, h])) \\
&\qquad - F(\bar{\mathbf{x}}_2, h(\bar{\mathbf{x}}_2), q, Y, \mathcal{I}^{1,\delta}[\bar{\mathbf{x}}_2, \phi_2] + \mathcal{I}^{2,\delta}[\bar{\mathbf{x}}_2, h] + O\left(\frac{|\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2|^2}{\epsilon}\right) \\
&\qquad + o_\beta(1) + \frac{1}{\epsilon} o_\delta(1))
\end{aligned}$$

Furthermore, again using assumption (A4),

$$\begin{aligned}
&\Big| F\left(\bar{\mathbf{x}}_2, h(\bar{\mathbf{x}}_2), q, Y, \mathcal{I}^{1,\delta}[\bar{\mathbf{x}}_2, -\phi_a(\bar{\mathbf{x}}_1, \cdot)] + \mathcal{I}^{2,\delta}[\bar{\mathbf{x}}_2, h]\right) \\
&\qquad - F\Big(\bar{\mathbf{x}}_2, h(\bar{\mathbf{x}}_2), q, Y, \mathcal{I}^{1,\delta}[\bar{\mathbf{x}}_2, \phi_2] + \mathcal{I}^{2,\delta}[\bar{\mathbf{x}}_2, h] \\
&\qquad\qquad + O\left(\frac{|\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2|^2}{\epsilon}\right) + o_\beta(1) + \frac{1}{\epsilon} o_\delta(1)\Big)\Big| \le K_{Lip} \Big| \mathcal{I}^{1,\delta}[\bar{\mathbf{x}}_2, -\phi_a(\bar{\mathbf{x}}_1, \cdot)] + \mathcal{I}^{2,\delta}[\bar{\mathbf{x}}_2, h] \\
&\qquad\qquad\qquad - \mathcal{I}^{1,\delta}[\bar{\mathbf{x}}_2, \phi_2] + \mathcal{I}^{2,\delta}[\bar{\mathbf{x}}_2, h] \\
&\qquad\qquad\qquad + O\left(\frac{|\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2|^2}{\epsilon}\right) + o_\beta(1)
\end{aligned}$$

$$
\begin{aligned}
&\qquad + \frac{1}{\epsilon} o_\delta(1))\Big| \\
&\le K_{Lip} \Big| \mathcal{I}^{1,\delta}[\bar{\mathbf{x}}_2, -\phi_a(\bar{\mathbf{x}}_1,\cdot)] + \mathcal{I}^{2,\delta}[\bar{\mathbf{x}}_2, h] \\
&\qquad - \mathcal{I}^{1,\delta}[\bar{\mathbf{x}}_2, \phi_2] + \mathcal{I}^{2,\delta}[\bar{\mathbf{x}}_2, h] \Big| \\
&\qquad + O\left(\frac{|\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2|^2}{\epsilon}\right) + o_\beta(1) \\
&\qquad + \frac{1}{\epsilon} o_\delta(1)).
\end{aligned}
$$

So, again applying Proposition 4.2 as we did to the previous pair of terms, we have

$$
\begin{aligned}
F(\bar{\mathbf{x}}_2, h(\bar{\mathbf{x}}_2), q, Y, \mathcal{I}^{1,\delta}[\bar{\mathbf{x}}_2, -\phi_a(\bar{\mathbf{x}}_1,\cdot)] + \mathcal{I}^{2,\delta}[\bar{\mathbf{x}}_2, h])) - F(\bar{\mathbf{x}}_2, h(\bar{\mathbf{x}}_2), q, Y, \ell) \\
\le o_a(1) + O\left(\frac{|\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2|^2}{\epsilon}\right) + o_\beta(1) + \frac{1}{\epsilon} o_\delta(1).
\end{aligned}
$$

So, finally we have

$$
\begin{aligned}
\gamma \frac{M}{2} \le F(\bar{\mathbf{x}}_2, h(\bar{\mathbf{x}}_2), q, Y, \ell) - F(\bar{\mathbf{x}}_1, h(\bar{\mathbf{x}}_2), q + \nabla\psi_\beta(\bar{\mathbf{x}}_1), X, \ell) + O\left(\frac{|\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2|^2}{\epsilon}\right) \\
+ o_\beta(1) + o_a(1) + \frac{1}{\epsilon} o_\delta(1). \qquad (4.31)
\end{aligned}
$$

In particular, (4.8) and the properties of ψ_β above imply that

$$
\begin{aligned}
\frac{1}{a}\begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \le \begin{bmatrix} X & 0 \\ 0 & -Y \end{bmatrix} &\le \frac{1}{\epsilon}\begin{bmatrix} I & -I \\ -I & I \end{bmatrix} + \begin{bmatrix} D^2\psi_\beta(\bar{\mathbf{x}}_1) & 0 \\ 0 & 0 \end{bmatrix} + o_a(1)\begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \\
&\le \frac{1}{\epsilon}\begin{bmatrix} I & -I \\ -I & I \end{bmatrix} + (o_a(1) + o_\beta(1))\begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}
\end{aligned}
$$

with $o_a(1)$ and $o_\beta(1)$ being uniform in ϵ.

We now apply assumption (A3*) taking $\rho = 2$ to get

$$
\begin{aligned}
\gamma \frac{M}{2} &\le F(\bar{\mathbf{x}}_2, h(\bar{\mathbf{x}}_2), q, Y, \ell) - F(\bar{\mathbf{x}}_1, h(\bar{\mathbf{x}}_2), q + \nabla\psi_\beta(\bar{\mathbf{x}}_1), X, \ell) + O\left(\frac{|\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2|^2}{\epsilon}\right) \\
&\qquad + o_\beta(1) + o_a(1) + \frac{1}{\epsilon} o_\delta(1) \\
&\le \omega\left(\frac{|\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2|^2}{\epsilon} + |\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2| + o_\beta(1)\right) + \omega_{R_\beta}\left(\frac{|\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2|^2}{\epsilon} + |\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2|\right) \\
&\qquad + O\left(\frac{|\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2|^2}{\epsilon}\right) + o_\beta(1) + o_a(1) + \frac{1}{\epsilon} o_\delta(1).
\end{aligned}
$$

Finally, letting in order δ, a, ϵ, and then β go to 0, we get that $M \le 0$ giving us the desired contradiction. □

1.3 Verification of the Assumptions

Define $R(x) := \min\{e^{(\beta_1-\epsilon)x}, e^{(\beta_2-\epsilon)x}\}$ for some $\epsilon > 0$ such that both $(\beta_1 - \epsilon)$ and $(\beta_2 - \epsilon)$ are positive. Then, the class of functions $\mathcal{C}$ is defined to be all functions $u : \mathcal{D} \to \mathbb{R}$ such that $u(\cdot, z)$ is continuous on $[z_{min}, z_{max}]$ and there exists a constant $\bar{c} > 0$ such that

$$u(x, z) \leq \bar{c}(1 + R(x)).$$

We will restrict our choices of proposed value function to these functions of sub-exponential growth in order to ensure that all terms in the HJB equation for this problem are integrable. Note that, in particular, the class $\mathcal{C}$ certainly contains all functions u which satisfy the polynomial growth condition for any degree $p \geq 0$.

1.3.1 The Nonlocal Term

Recall that in the HJB equation for this problem,

$$\begin{aligned}
0 = {} & F((x,z), f, Df, D^2 f, \mathcal{I}[(x,z), f]) && (4.32)\\
= {} & \delta f(x,z) - \frac{1}{\lambda}(\mu - x) f_x(x,z) - \frac{\sigma^2}{2} f_{xx}(x,z) - \sup_{u \in U} \{-u f_z(x,z) + G(x,z,u)\} \\
& - \eta_1 \int_0^\infty (f(x+y,z) - f(x,z))\, \beta_1 e^{-\beta_1 y}\, dy \\
& - \eta_2 \int_0^\infty (f(x-y,z) - f(x,z))\, \beta_2 e^{-\beta_2 y}\, dy \\
& - \alpha \big[p_1 (f(x, z+\zeta) - f(x,z)) + p_2 (f(x, z+2\zeta) - f(x,z)) \big], && (4.33)
\end{aligned}$$

the last three terms are due to the nonlocal behavior of the X and Z processes. That is, we can write the nonlocal operator for this PIDE as

$$\begin{aligned}
\mathcal{I}[(x,z), f] = {} & - \eta_1 \int_0^\infty (f(x+y,z) - f(x,z))\, \beta_1 e^{-\beta_1 y}\, dy \\
& - \eta_2 \int_{-\infty}^0 (f(x+y,z) - f(x,z))\, \beta_2 e^{\beta_2 y}\, dy \\
& - \alpha \big[p_1 (f(x, z+\zeta) - f(x,z)) + p_2 (f(x, z+2\zeta) - f(x,z)) \big]. && (4.34)
\end{aligned}$$

Following the form of Barles and Imbert, we separate this operator into

$$\mathcal{I}^{1,\delta}[(x,z), f] = - \eta_1 \int_0^{R_\delta} (f(x+y,z) - f(x,z))\, \beta_1 e^{-\beta_1 y}\, dy$$

$$
- \eta_2 \int_{-R_\delta}^{0} \left(f(x+y,z) - f(x,z)\right) \beta_2 e^{\beta_2 y} \, dy \tag{4.35}
$$

$$
\begin{aligned}
\mathcal{I}^{2,\delta}[(x,z),f] = & - \eta_1 \int_{R_\delta}^{\infty} \left(f(x+y,z) - f(x,z)\right) \beta_1 e^{-\beta_1 y} \, dy \\
& - \eta_2 \int_{-\infty}^{-R_\delta} \left(f(x+y,z) - f(x,z)\right) \beta_2 e^{\beta_2 y} \, dy \\
& - \alpha \big[p_1 \left(f(x,z+\zeta) - f(x,z)\right) + p_2 \left(f(x,z+2\zeta) - f(x,z)\right) \big].
\end{aligned} \tag{4.36}
$$

for any $\delta > 0$ where $R_\delta = \min\{\delta, \zeta\}$.

By construction of $\mathcal{I}^{1,\delta}$ and $\mathcal{I}^{2,\delta}$, it is clear that for any $(x,z) \in \mathcal{D}$ and $\phi \in \mathcal{C} \cap C^2(\mathcal{D})$,

$$
\mathcal{I}[(x,z),\phi] = \mathcal{I}^{1,\delta}[(x,z),\phi] + \mathcal{I}^{2,\delta}[(x,z),\phi],
$$

and further for any $a \in \mathbb{R}$,

$$
\begin{aligned}
\mathcal{I}^{2,\delta}[(x,z),\phi + a] = & - \eta_1 \int_{R_\delta}^{\infty} \left([\phi(x+y,z) + a] - [\phi(x,z) + a]\right) \beta_1 e^{-\beta_1 y} \, dy \\
& - \eta_2 \int_{-\infty}^{-R_\delta} \left([\phi(x+y,z) + a] - [\phi(x,z) + a]\right) \beta_2 e^{\beta_2 y} \, dy \\
& - \alpha \big[p_1 \left([\phi(x,z+\zeta) + a] - [\phi(x,z) + a]\right) \\
& \qquad + p_2 \left([\phi(x,z+2\zeta) + a] - [\phi(x,z) + a]\right) \big] \\
= & - \eta_1 \int_{R_\delta}^{\infty} \left(\phi(x+y,z) - \phi(x,z)\right) \beta_1 e^{-\beta_1 y} \, dy \\
& - \eta_2 \int_{-\infty}^{-R_\delta} \left(\phi(x+y,z) - \phi(x,z)\right) \beta_2 e^{\beta_2 y} \, dy \\
& - \alpha \big[p_1 \left(\phi(x,z+\zeta) - \phi(x,z)\right) + p_2 \left(\phi(x,z+2\zeta) - \phi(x,z)\right) \big] \\
= & \, \mathcal{I}^{2,\delta}[(x,z),\phi]
\end{aligned}
$$

and similarly $\mathcal{I}^{1,\delta}[(x,z),\phi + a] = \mathcal{I}^{1,\delta}[(x,z),\phi]$. Moreover, setting $R_\delta = \min\{\delta, \zeta\}$, we have that $R_\delta \to 0$ as $\delta \to 0$ and, for $\phi_1, \phi_2 \in \mathcal{C} \cap C^2(\mathcal{D})$,

$$
\phi_1 = \phi_2 \;\; \forall (x,z) \in B((x,z), R_\delta) \quad \Rightarrow \quad \mathcal{I}^{1,\delta}[(x,z),\phi_1] = \mathcal{I}^{1,\delta}[(x,z),\phi_2] \;\; \text{and}
$$

$$
\phi_1 = \phi_2 \;\; \forall (x,z) \in (\mathcal{D} \setminus B((x,z), R_\delta)) \cup \{(x,z)\} \quad \Rightarrow \quad \mathcal{I}^{2,\delta}[(x,z),\phi_1] = \mathcal{I}^{2,\delta}[(x,z),\phi_2].
$$

Now, let $\phi \in C^2(\mathcal{D})$ and $f \in \mathcal{C}$ be an upper semi-continuous function such that $f - \phi$ attains a maximum on $B((\hat{x},\hat{z}), R_\delta)$ at $(\hat{x},\hat{z})$ and $\phi(\hat{x},\hat{z}) = f(\hat{x},\hat{z})$. Denote the ball of radius $d > 0$ centered at $(\hat{x}, \hat{z})$ by $\hat{B}(d) := B((\hat{x}, \hat{z}), d)$, and denote the annulus with inner radius $d > 0$ and outer radius $D > 0$ centered at $(\hat{x}, \hat{z})$ by $\hat{A}(d, D) := \{(x, z) \in \mathcal{D} : d < |(x, z) - (\hat{x}, \hat{z})| < D\}$. We proceed with the following construction of a sequence of approximating functions which follows a very similar construction by Arisawa [1]. First, we note that in the case of viscosity solutions, it is sufficient to consider $f \in \mathcal{C} \cap C(\mathcal{D})$, since if f is not continuous, it can be approximated by the sequence of sup-convolutions, each of which is continuous. It is a classical result of viscosity solution theory that f is a viscosity subsolution of (4.1), if and only if each of the sequence of sup-convolutions approximating f is a viscosity subsolution of (4.1) (see, for example, Section 5.2 of the book by Pallaschke and Rolewicz [12]).

Consider $\hat{B}(s)$ for some $s > R_\delta$. Since, as noted earlier, the class $\mathcal{C}$ contains the polynomial functions, there certainly exists a sequence of functions $\psi_n \in C^2(\mathcal{D}) \times \mathcal{C}$ such that

$$\lim_{n\to\infty} \psi_n((x, z)) = f((x, z)) \text{ uniformly in } \hat{B}(s).$$

Define $\overline{\psi}_n((x, z)) := \psi_n((x, z)) + \|f - \psi_n((x, z))\|_{L^\infty(\hat{B}(s))}$. Then,

$$\overline{\psi}_n((x, z)) \geq f((x, z)) \ \ \forall\ (x, z) \in \hat{B}(s) \ \text{ and}$$

$$\lim_{n\to\infty} \overline{\psi}_n((x, z)) = f((x, z)) \text{ uniformly in } \hat{B}(s).$$

Without loss of generality, we may assume that $f - \phi$ takes a strict maximum at $(\hat{x}, \hat{z})$ (since if it does not, we can add a small positive quadratic function to ϕ), and therefore for any r such that $R_\delta < r < s$, there exists some $\sigma(r) > 0$ such that

$$\min_{(x,z)\in\hat{A}(r,s)} (\phi - f)(x, z) =: \sigma(r) > 0.$$

Let $\chi_r(x, z) \in C^2(\hat{B}(s))$ be a function such that $0 \leq \chi_r(x, z) \leq 1$ for all $(x, z) \in \hat{B}(s)$ and

$$\chi_r(x, z) := \begin{cases} 1 & \text{if } |(x, z) - (\hat{x}, \hat{z})| \leq R_\delta \\ 0 & \text{if } r \leq |(x, z) - (\hat{x}, \hat{z})| \leq s \end{cases}.$$

Then, define

$$\overline{\phi}_n(x, z) := \chi_r(x, z)\phi(x, z) + (1 - \chi_r(x, z))\overline{\psi}_n(x, z)$$

for all $(x, z) \in \hat{B}(s)$. That is, the function $\overline{\phi}_n$ takes on the same values as ϕ inside the ball $\hat{B}(R_\delta)$ and the same values as f on the annulus $(x, z) \in \hat{A}(r, s)$, making the transition between the two in a C^2 manner across the annulus $(x,z) \in \hat{A}(R_\delta, r)$. By construction of $\overline{\psi}_n$, and the fact that $\phi(x,z) \geq f(x,z)$ for all $(x,z) \in \hat{B}(s)$ by assumption, it is clear that $\overline{\phi}_n(x,z) \geq f(x,z)$ for all $(x,z) \in \hat{B}(s)$ and all $n = 1, 2, ...$ Further, since ϕ, $\overline{\psi}_n$, and χ_r are all in $C^2(\hat{B}(s))$, we also have that $\overline{\phi}_n \in C^2(\hat{B}(s))$. Also, by the definitions of χ_r and $\overline{\phi}_n$, it is clear that $\overline{\phi}_n(x,z) = \phi(x,z)$ for all (x,z) such that $|(x,z) - (\hat{x}, \hat{z})| \leq R_\delta$. Moreover, noting first that there exists some $n_1 \in \mathbb{N}$ such that

$$\begin{aligned} \phi(x,z) - \overline{\psi}_n(x,z) &= \phi(x,z) - f(x,z) + f(x,z) - \overline{\psi}_n(x,z) \\ &\geq \sigma(R_\delta) + f(x,z) - \overline{\psi}_n(x,z) > 0 \end{aligned}$$

for any $n \geq n_1$ and $(x,z) \in \hat{A}(R_\delta, s)$, it is clear that

$$\phi(x,z) - \overline{\phi}_n = (1 - \chi_r(x,z))(\phi(x,z) - \overline{\psi}_n(x,z)) \geq 0$$

for $n \geq n_1$ and (x,z) such that $R_\delta \leq |(x,z) - (\hat{x},\hat{z})| \leq s$. That is, for $n \geq n_1$, $\overline{\phi}_m(x,z) \leq \phi(x,z)$ for $(x,z) \in \hat{A}(R_\delta, s)$. Since the fact that $\overline{\psi}_n \to f$ uniformly gives us that $\overline{\phi}_n \to f$ uniformly on $(x,z) \in \hat{A}(r,s)$, there exist $r > R_\delta$ sufficiently small and $n_2 \in \mathbb{N}$ such that for $n \geq n_2$ and for all $(x,z) \in \hat{A}(R_\delta, s)$,

$$|\overline{\phi}_n(x,z) - f(x,z)| \leq \frac{1}{s}.$$

Take $n' := \max\{n_1, n_2\}$. Also note that since $\overline{\phi}_n(x,z) = \phi(x,z)$ for $(x,z) \in \hat{B}(R_\delta)$ and, as argued earlier, $\overline{\phi}_n(x,z) \leq \phi(x,z)$ for $(x,z) \in \hat{A}(R_\delta, s)$ for $n \geq n'$, we have

$$f(x,z) \leq \overline{\phi}_n(x,z) \leq \phi(x,z) \quad \text{for all } (x,z) \in \hat{B}(s).$$

Then we can extend this $\overline{\phi}_{n'}$ to $\phi_1 \in C^2(\mathcal{D})$ such that

$$\phi_1(x,z) = \overline{\phi}_{n'}(x,z) \quad \text{for } (x,z) \in \hat{B}(s),$$

$$f(x) \leq \phi_1(x,z) \leq \phi(x,z) \quad \text{for all } x \in \mathcal{D},$$

and $f - \phi_1$ takes a strict maximum at $(\hat{x}, \hat{z})$.

Now, repeat the above construction with ϕ replaced by ϕ_1 and s replaced by $2s$. This will yield a function $\phi_2 \in C^2(\mathcal{D})$ such that

$$f(x) \leq \phi_2(x,z) \leq \phi_1(x,z) \leq \phi(x,z) \quad \text{for all } x \in \mathcal{D},$$

$$|\phi_2(x,z) - f(x,z)| \le \frac{1}{2s} \text{ for } (x,z) \in \{(x,z) \in \mathcal{D} : R_\delta \le |(x,z) - (\hat{x},\hat{z})| \le 2s\}$$

and $f - \phi_2$ takes a strict maximum at $(\hat{x},\hat{z})$. Continuing iteratively in this manner, we can construct a sequence of functions $\phi_k \in \mathcal{C} \cap C^2(\mathcal{D})$ such that

$$f(x) \le \phi_k(x,z) \le \phi_{(k-1)}(x,z) \le ... \le \phi_1(x,z) \le \phi(x,z) \text{ for all } x \in \mathcal{D},$$

$$|\phi_k(x,z) - f(x,z)| \le \frac{1}{ks} \text{ for } (x,z) \in \{(x,z) \in \mathcal{D} : R_\delta \le |(x,z) - (\hat{x},\hat{z})| \le ks\}$$

and $f - \phi_k$ takes a strict maximum at $(\hat{x},\hat{z})$.

Since $\phi_k((x,z)) = \phi((x,z))$ for all $(x,z) \in \hat{B}(R_\delta)$ for every $k = 1,2,...$, clearly

$$\mathcal{I}^{1,\delta}[(\hat{x},\hat{z}),\phi_k] \to \mathcal{I}^{1,\delta}[(\hat{x},\hat{z}),\phi] \text{ as } k \to \infty.$$

Further, since each function ϕ_k is in the class $\mathcal{C}$,

$$\left|\eta_1 \int_{R_\delta}^{\infty} (\phi_k(\hat{x}+y,\hat{z}) - \phi_k(\hat{x},\hat{z}))\, \beta_1 e^{-\beta_1 y}\, dy\right| + \left|\eta_2 \int_{-\infty}^{-R_\delta} (\phi_k(\hat{x}+y,\hat{z}) - \phi_k(\hat{x},\hat{z}))\, \beta_2 e^{\beta_2 y}\, dy\right| < \infty,$$

and ϕ_k is a monotone decreasing sequence converging to f outside $\hat{B}(R_\delta)$, the Monotone Convergence Theorem implies that

$$\mathcal{I}^{2,\delta}[(\hat{x},\hat{z}),\phi_k] \to \mathcal{I}^{2,\delta}[(\hat{x},\hat{z}),f].$$

Clearly, $\mathcal{I}^{1,\delta}[(x,z),\phi]$ is well-defined for $(x,z) \in \mathcal{D}$ and $\phi \in C^2(B((x,x),r)) \cap \mathcal{C}$ for any $r \in (0,R_\delta)$. Further, when $\delta \to 0$, $R_\delta = \min\{\delta,\zeta\} \to 0$ and thus $\mathcal{I}^{1,\delta}[(x,z),\phi] \to 0$. Now, suppose $(x_k,z_k) \to (x,z)$ and $\{\phi_k\} \subset C^2(B((x,z),r)) \cap C(\overline{B((x,z),R_\delta)})$ with $\phi_k \to \phi$. Then, since the integrands are continuous and thus bounded for $y \in [0,R_\delta)$, we have

$$\begin{aligned}
\lim_{k\to\infty} \mathcal{I}^{1,\delta}[(x_k,z_k),\phi_k] = &- \eta_1 \lim_{k\to\infty} \int_0^{R_\delta} (\phi_k(x_k+y,z_k) - \phi_k(x_k,z_k))\, \beta_1 e^{-\beta_1 y}\, dy \\
&- \eta_2 \lim_{k\to\infty} \int_{-R_\delta}^{0} (\phi_k(x_k+y,z_k) - \phi_k(x_k,z_k))\, \beta_2 e^{\beta_2 y}\, dy \\
= &- \eta_1 \int_0^{R_\delta} \lim_{k\to\infty} (\phi_k(x_k+y,z_k) - \phi_k(x_k,z_k))\, \beta_1 e^{-\beta_1 y}\, dy \\
&- \eta_2 \int_{-R_\delta}^{0} \lim_{k\to\infty} (\phi_k(x_k-y,z_k) - \phi_k(x_k,z_k))\, \beta_2 e^{-\beta_2 y}\, dy
\end{aligned}$$

$$
\begin{aligned}
&= -\eta_1 \int_0^{R_\delta} \left(\phi(x+y,z) - \phi(x,z)\right) \beta_1 e^{-\beta_1 y}\, dy \\
&\quad - \eta_2 \int_{-R_\delta}^{0} \left(\phi(x+y,z) - \phi(x,z)\right) \beta_2 e^{\beta_2 y}\, dy \\
&= \mathcal{I}^{1,\delta}[(x,z),\phi]
\end{aligned}
$$

Since all functions $\phi \in \mathcal{C}$ are locally bounded, $\mathcal{I}^{2,\delta}[(x,z),\phi]$ is well-defined for any $(x,z) \in \mathcal{D}$ and any $\phi \in \mathcal{C}$. Let (ϕ_k) be a sequence of uniformly locally bounded functions such that $|\phi_k| \le \psi$ for some function $\psi \in \mathcal{C}$. Recalling the definition $\overline{\phi} := \limsup^* \phi_k$, we then have by a standard extension of Fatou's lemma that

$$
\begin{aligned}
\limsup_{k\to\infty} \mathcal{I}^{2,\delta}[(x_k,z_k),\phi_k] &= -\eta_1 \limsup_{k\to\infty} \int_{R_\delta}^{\infty} \left(\phi(x_k+y,z_k) - \phi(x_k,z_k)\right) \beta_1 e^{-\beta_1 y}\, dy \\
&\quad - \eta_2 \limsup_{k\to\infty} \int_{-\infty}^{-R_\delta} \left(\phi_k(x_k+y,z_k) - \phi_k(x_k,z_k)\right) \beta_2 e^{\beta_2 y}\, dy \\
&\quad - \alpha \limsup_{k\to\infty} \big[p_1\left(\phi_k(x_k,z_k+\zeta) - \phi_k(x_k,z_k)\right) \\
&\qquad\qquad + p_2\left(\phi_k(x_k,z_k+2\zeta) - \phi_k(x_k,z_k)\right)\big] \\
&\le -\eta_1 \int_{R_\delta}^{\infty} \limsup_{k\to\infty} \left(\phi(x_k+y,z_k) - \phi(x_k,z_k)\right) \beta_1 e^{-\beta_1 y}\, dy \\
&\quad - \eta_2 \int_{-\infty}^{-R_\delta} \limsup_{k\to\infty} \left(\phi_k(x_k+y,z_k) - f(x_k,z_k)\right) \beta_2 e^{\beta_2 y}\, dy \\
&\quad - \alpha\big[p_1 \left(\limsup_{k\to\infty} \phi_k(x_k,z_k+\zeta) - \limsup_{k\to\infty} \phi_k(x_k,z_k)\right) \\
&\qquad + p_2 \left(\limsup_{k\to\infty} \phi_k(x_k,z_k+2\zeta) - \limsup_{k\to\infty} \phi_k(x_k,z_k)\right)\big] \\
&= \mathcal{I}^{2,\delta}[(x,z),\overline{\phi}].
\end{aligned}
$$

By a symmetric argument and the definition $\underline{\phi} := \liminf^* \phi_k$, we also have that $\liminf_{k\to\infty} \mathcal{I}[(x_k,z_k),\phi_k] \ge \mathcal{I}[(x,z),\underline{\phi}]$.

Thus, the nonlocal operator $\mathcal{I}$ satisfies the conditions of assumption (NLT).

1.3.2 Assumptions of the Comparison Principle

Again mirroring the form of Barles and Imbert, we define the following function j representing the sizes of the jumps and Lévy measure μ representing the action of the nonlocal components of the paired process (X,Z):

$$
j((x,z),(y,w)) := (y, w \wedge (z_{max} - z)) \tag{4.37}
$$

$$\mu(y,w) := -\eta_1 \mathbf{1}_{(0,\infty)}(y)\beta_1 e^{-\beta_1 y} dy \delta_{\{0\}}(dw) - \eta_2 \mathbf{1}_{(-\infty,0)}(y)\beta_2 e^{\beta_2 y} dy \delta_{\{0\}}(dw)$$
$$-\alpha\big[p_1\delta_{\{\zeta\}}(dw) + p_2\delta_{\{2\zeta\}}(dw)\big]\delta_{\{0\}}(dy) \tag{4.38}$$

where $\mathbf{1}_A(\cdot)$ is the indicator of the set $A \in \mathbb{R}$ and $\delta_{\{a\}}(\cdot)$ is the point mass at $a \in \mathbb{R}$. Note that in the measure μ, the point masses guarantee that μ will only have non-zero mass if at most one of the processes X or Z jumps at any given time. We can make this simplifying assumption since these two processes are independent and each have exponentially distributed times between jumps, implying that

$$\mathbb{P}\left(|X(t) - X(t-)| \neq 0\,,\, |Z(t) - Z(t-)| \neq 0\right) = 0 \quad \forall t \geq 0.$$

Recall that we defined above the following function j representing the sizes of the jumps and Lévy measure μ representing the action of the nonlocal components of the paired process (X, Z):

$$j((x,z),(y,w)) := (y, w \wedge (z_{max} - z)) \tag{4.39}$$

$$\mu(y,w) := \eta_1 \mathbf{1}_{(0,\infty,)}(y)\beta_1 e^{-\beta_1 y} dy \delta_{\{0\}}(w) + \eta_2 \mathbf{1}_{(-\infty,0)}(y)\beta_2 e^{\beta_2 y} dy \delta_{\{0\}}(w)$$
$$+\alpha\big[p_1\delta_{\{\zeta\}}(w) + p_2\delta_{\{2\zeta\}}(w)\big]\delta_{\{0\}}(y) \tag{4.40}$$

where $\mathbf{1}_A(\cdot)$ is the indicator of the set $A \in \mathbb{R}$ and $\delta_{\{a\}}(\cdot)$ is the point mass at $a \in \mathbb{R}$.

Lemma 4.8. *The HJB equation (4.32) with the paired process (X, Z) as defined above satisifies assumption (A1*) of the comparison principle.*

Proof. First, we note that

$$\begin{aligned}
\int_{\mathbb{R}^2} \mu(dy,dw) = &\int_{\mathbb{R}}\int_{\mathbb{R}} \eta_1 \mathbf{1}_{(0,\infty,)}(y)\beta_1 e^{-\beta_1 y} dy \delta_{\{0\}}(dw) \\
&+ \int_{\mathbb{R}}\int_{\mathbb{R}} \eta_2 \mathbf{1}_{(-\infty,0)}(y)\beta_2 e^{\beta_2 y} dy \delta_{\{0\}}(dw) \\
&+ \int_{\mathbb{R}}\int_{\mathbb{R}} \alpha\big[p_1\delta_{\{\zeta\}}(dw) + p_2\delta_{\{2\zeta\}}(dw)\big]\delta_{\{0\}}(dy) \\
= &\int_0^\infty \eta_1\beta_1 e^{-\beta_1 y} dy + \int_{-\infty}^0 \eta_2\beta_2 e^{\beta_2 y} dy + p_1\zeta + 2p_2\zeta \\
= &\;\eta_1 + \eta_2 + \alpha[p_1(\zeta \wedge (z_{max} - z)) + p_2(2\zeta \wedge (z_{max} - z))].
\end{aligned}$$

So clearly $\int_{\mathbb{R}^2 \setminus B} \mu(dy,dw) \leq \eta_1 + \eta_2 + p_1\zeta + 2p_2\zeta < \infty$ for any open ball $B \subset \mathcal{D}$. Further, since $|j((x,z),(y,w))|^2 = y^2 + (w \wedge (z_{max} - z))^2 \leq y^2 + w^2 = |(y,w)|^2$ for all (x,z) in $\mathcal{D}$,

$$\sup_{(x,z)\in\mathcal{D}} \int_{\mathbb{R}} |j((x,z),(y,w))|^2 \mu(dy,dw) \leq \int_{\mathbb{R}^2} |(y,w)|^2 \mu(dy,dw)$$

$$
\begin{aligned}
&= \int_{\mathbb{R}^2} (y^2 + w^2)\mu(dy, dw) \\
&= \int_0^{\infty} y^2 \eta_1 \beta_1 e^{-\beta_1 y} dy + \int_{-\infty}^{0} y^2 \eta_2 \beta_2 e^{\beta_2 y} dy \\
&\qquad + \alpha[(\zeta)^2 p_1(\zeta) + (2\zeta)^2 p_2(2\zeta)] \\
&= \eta_1 \frac{2}{{\beta_1}^2} + \eta_2 \frac{2}{{\beta_2}^2} + \alpha[p_1 \zeta^3 + 4p_2 \zeta^3]
\end{aligned}
$$

and thus $\sup_{(x,z)\in\mathcal{D}} \int_{\mathbb{R}^2} |j((x,z),(y,w))|^2 \mu(dy,dw) < \infty$. Now, note that

$$
\begin{aligned}
|j((x,z),(y,w)) - j((\tilde{x},\tilde{z}),(y,w))| &= |(y - y, (w \wedge (z_{max} - z)) - (w \wedge (z_{max} - \tilde{z})))| \\
&= |(w \wedge (z_{max} - z)) - (w \wedge (z_{max} - \tilde{z}))| \\
&\le |z - \tilde{z}|
\end{aligned}
$$

for any (x,z) and $(\tilde{x},\tilde{z})$ in $\mathcal{D}$. So, noting that $\mu(dy,dw)$ does not depend on z and

$$
\begin{aligned}
|j((x,z),(y,w)) - j((\tilde{x},\tilde{z}),(y,w))|^2 &= |(y, w \wedge (z - z_{max})) - (y, w \wedge (\tilde{z} - z_{max})|^2 \\
&= (y-y)^2 + (z - \tilde{z})^2 = (z - \tilde{z})^2,
\end{aligned}
$$

we have

$$
\begin{aligned}
&\int_{\mathbb{R}^2} |j((x,z),(y,w)) - j((\tilde{x},\tilde{z}),(y,w))|^2 \mu(dy,dw) \\
&\qquad = \int_{\mathbb{R}^2} |((w \wedge (z - z_{max})) - (w \wedge (\tilde{z} - z_{max}))|^2 \mu(dy,dw) \\
&\qquad \le \int_{\mathbb{R}^2} (z - \tilde{z})^2 \mu(dy,dw) \\
&\qquad = (z - \tilde{z})^2 \int_{\mathbb{R}^2} \mu(dy,dw) \\
&\qquad \le \bar{c}(z - \tilde{z})^2 \\
&\qquad \le \bar{c}|(x,z) - (\tilde{x},\tilde{z})|^2
\end{aligned}
$$

for any $\bar{c} \ge \eta_1 + \eta_2 + \alpha[p_1(\zeta \wedge (z_{max} - z)) + p_2(2\zeta \wedge (z_{max} - z))]$ and any (x,z) and $(\tilde{x},\tilde{z})$ in $\mathcal{D}$. Similarly,

$$
\int_{\mathbb{R}^2} |j((x,z),(y,w)) - j((\tilde{x},\tilde{z}),(y,w))| \mu(dy,dw) \le \bar{c}|z - \tilde{z}| \le \bar{c}|(x,z) - (\tilde{x},\tilde{z})|
$$

for any constant $\bar{c} \ge \eta_1 + \eta_2 + \alpha[p_1(\zeta \wedge (z_{max} - z)) + p_2(2\zeta \wedge (z_{max} - z))]$ and any (x,z) and $(\tilde{x},\tilde{z})$ in $\mathcal{D}$. Thus we have that assumption (A1) is satisfied. $\square$

In order to show the remaining assumptions, we need to consider the PIDE (4.32) in terms of constant arguments rather than functions. That is, for $(x, z) \in \mathcal{D}$, $a \in \mathbb{R}$, $b = (b_1, b_2) \in \mathbb{R}^2$, $X \in \mathbb{S}_2$, and $l \in \mathbb{R}$,

$$F((x,z), a, b, X, l) = \delta a - \frac{1}{\lambda}(\mu - x)b_1 - \frac{\sigma^2}{2}X_{1,1} - \sup_{u \in U}\{-ub_2 + G(x,z,u)\} + l \tag{4.41}$$

Then clearly

$$F((x,z), a, b, X, l) - F((x,z), \tilde{a}, b, X, l) = \delta(a - \tilde{a}),$$

and thus assumption (A2) is satisfied. Also, F is clearly Lipschitz continuous in l, and thus assumption (A4) is satisfied.

It remains only to show that assumption (A3-1*) is satisfied.

Lemma 4.9. *The HJB equation (4.32) with the paired process (X, Z) as defined above satisifies assumption (A3*) of the comparison principle.*

Proof. First, we note that the assertion that the matrices $X, Y \in \mathbb{S}_2$ satisfy

$$\begin{bmatrix} X & 0 \\ 0 & -Y \end{bmatrix} \le \frac{1}{\epsilon}\begin{bmatrix} I & -I \\ -I & I \end{bmatrix} + r(\beta)\begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \tag{4.42}$$

for some $\epsilon > 0$ and some $r(\beta)$ with $r(\beta) \to 0$ as $\beta \to 0$ really means that the matrix

$$\begin{aligned} M &:= \frac{1}{\epsilon}\begin{bmatrix} I & -I \\ -I & I \end{bmatrix} + r(\beta)\begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} - \begin{bmatrix} X & 0 \\ 0 & -Y \end{bmatrix} \\ &= \begin{bmatrix} \frac{1}{\epsilon} + r(\beta) - X_{1,1} & X_{1,2} & -\frac{1}{\epsilon} & 0 \\ -X_{1,2} & \frac{1}{\epsilon} + r(\beta) - X_{2,2} & 0 & \frac{1}{\epsilon} \\ -\frac{1}{\epsilon} & 0 & \frac{1}{\epsilon} + r(\beta) + Y_{1,1} & Y_{1,2} \\ 0 & -\frac{1}{\epsilon} & Y_{1,2} & \frac{1}{\epsilon} + r(\beta) + Y_{2,2} \end{bmatrix} \end{aligned}$$

is positive semi-definite. That is, for any non-zero column vector $v \in \mathbb{R}^4$, $v^T M v \ge 0$. So, in particular,

$$\begin{aligned} (1,0,1,0)M\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} &= \begin{pmatrix} \frac{1}{\epsilon} + r(\beta) - X_{1,1} - \frac{1}{\epsilon} \\ X_{1,2} \\ -\frac{1}{\epsilon} + \frac{1}{\epsilon} + r(\beta) + Y_{1,1} \\ Y_{1,2} \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} \\ &= \frac{1}{\epsilon} + r(\beta) - X_{1,1} - \frac{1}{\epsilon} - \frac{1}{\epsilon} + \frac{1}{\epsilon} + r(\beta) + Y_{1,1} = 2r(\beta) - X_{1,1} + Y_{1,1} \ge 0, \end{aligned}$$

implying that $Y_{1,1} - X_{1,1} \geq -2r(\beta)$. So, for $s(\beta) = (s_1(\beta), s_2(\beta))$ where $s(\beta) \to (0,0)$ as $\beta \to 0$ and $|x,z)|$, $|(\tilde{x},\tilde{z})| \leq \frac{\rho}{\beta} =: R$,

$$
\begin{aligned}
&F((\tilde{x},\tilde{z}), a, \frac{1}{\epsilon}((x,z)-(\tilde{x},\tilde{z})), Y, l) - F((x,z), a, \frac{1}{\epsilon}((x,z)-(\tilde{x},\tilde{z})) + s(\beta), X, l) \\
&\quad = \left[\delta a - \frac{1}{\lambda}(\mu - \tilde{x})\left(\frac{1}{\epsilon}(x-\tilde{x})\right) - \frac{\sigma^2}{2}Y_{1,1} - \sup_{u\in U}\left\{-u\left(\frac{1}{\epsilon}(z-\tilde{z})\right) + G(\tilde{x},\tilde{z},u)\right\} + l\right] \\
&\qquad\qquad - \left[\delta a - \frac{1}{\lambda}(\mu - x)\left(\frac{1}{\epsilon}(x-\tilde{x}) + s_1(\beta)\right) - \frac{\sigma^2}{2}X_{1,1}\right. \\
&\qquad\qquad\qquad \left. - \sup_{u\in U}\left\{-u\left(\frac{1}{\epsilon}(z-\tilde{z}) + s_2(\beta)\right) + G(x,z,u)\right\} + l\right] \\
&\quad = \frac{1}{\lambda}(\mu - x)\left(\frac{1}{\epsilon}(x-\tilde{x}) + s_1(\beta)\right) - \frac{1}{\lambda}(\mu-\tilde{x})\left(\frac{1}{\epsilon}(x-\tilde{x})\right) - \frac{\sigma^2}{2}(Y_{1,1} - X_{1,1}) \\
&\qquad\qquad\qquad - \sup_{u\in U}\left\{-u\left(\frac{1}{\epsilon}(z-\tilde{z})\right) + G(\tilde{x},\tilde{z},u)\right\} \\
&\qquad\qquad\qquad + \sup_{u\in U}\left\{-u\left(\frac{1}{\epsilon}(z-\tilde{z}) + s_2(\beta)\right) + G(x,z,u)\right\} \\
&\quad = \frac{1}{\lambda\epsilon}(x-\tilde{x})(\tilde{x}-x) + \frac{1}{\lambda}(\mu - x)s_1(\beta) - \frac{\sigma^2}{2}(Y_{1,1} - X_{1,1}) \\
&\qquad\qquad\qquad - \sup_{u\in U}\left\{-u\left(\frac{1}{\epsilon}(z-\tilde{z})\right) + G(\tilde{x},\tilde{z},u)\right\} \\
&\qquad\qquad\qquad + \sup_{u\in U}\left\{-u\left(\frac{1}{\epsilon}(z-\tilde{z}) + s_2(\beta)\right) + G(x,z,u)\right\} \\
&\quad = -\frac{1}{\lambda\epsilon}(x-\tilde{x})^2 + \frac{1}{\lambda}(\mu - x)s_1(\beta) - \frac{\sigma^2}{2}(Y_{1,1} - X_{1,1}) \\
&\qquad\qquad\qquad - \sup_{u\in U}\left\{-u\left(\frac{1}{\epsilon}(z-\tilde{z})\right) + G(\tilde{x},\tilde{z},u)\right\} \\
&\qquad\qquad\qquad + \sup_{u\in U}\left\{-u\left(\frac{1}{\epsilon}(z-\tilde{z}) + s_2(\beta)\right) + G(z,x,u)\right\} \\
&\quad \leq \frac{1}{\lambda}(\mu - x)s_1(\beta) + \frac{\sigma^2}{2}(2r(\beta)) - \sup_{u\in U}\left\{-u\left(\frac{1}{\epsilon}(z-\tilde{z})\right) + G(\tilde{x},\tilde{z},u)\right\} \\
&\qquad\qquad\qquad + \sup_{u\in U}\left\{-u\left(\frac{1}{\epsilon}(z-\tilde{z}) + s_2(\beta)\right) + G(x,z,u)\right\} \\
&\quad \leq \frac{1}{\lambda}\mu s_1(\beta) + \frac{\rho}{\lambda\beta}|s_1(\beta)| + \sigma^2 r(\beta) - \sup_{u\in U}\left\{-u\left(\frac{1}{\epsilon}(z-\tilde{z})\right) + G(\tilde{x},\tilde{z},u)\right\} \\
&\qquad\qquad + \sup_{u\in U}\left\{-u\left(\frac{1}{\epsilon}(z-\tilde{z})\right) + G(x,z,u)\right\} + \sup_{u\in U}\{u|s_2(\beta)|\} \\
&\quad = \frac{1}{\lambda}\mu s_1(\beta) + \frac{\rho}{\lambda\beta}|s_1(\beta)| + \sigma^2 r(\beta) + u_{max}|s_2(\beta)| - \sup_{u\in U}\left\{-u\left(\frac{1}{\epsilon}(z-\tilde{z})\right) + G(\tilde{x},\tilde{z},u)\right\}
\end{aligned}
$$

$$
\begin{aligned}
&+ \sup_{u\in U}\left\{-u\left(\frac{1}{\epsilon}(z-\tilde{z})\right) + G(x,z,u) + G(\tilde{x},\tilde{z},u) - G(\tilde{x},\tilde{z},u)\right\} \\
\le\ & \frac{1}{\lambda}\mu s_1(\beta) + \frac{\rho}{\lambda\beta}|s_1(\beta)| + \sigma^2 r(\beta) + u_{max}|s_2(\beta)| - \sup_{u\in U}\left\{-u\left(\frac{1}{\epsilon}(z-\tilde{z})\right) + G(\tilde{x},\tilde{z},u)\right\} \\
&+ \sup_{u\in U}\left\{-u\left(\frac{1}{\epsilon}(z-\tilde{z})\right) + G(\tilde{x},\tilde{z},u)\right\} + \sup_{u\in U}\left\{G(x,z,u) - G(\tilde{x},\tilde{z},u)\right\} \\
=\ & \frac{1}{\lambda}\mu s_1(\beta) + \frac{\rho}{\lambda\beta}|s_1(\beta)| + \sigma^2 r(\beta) + u_{max}|s_2(\beta)| + \sup_{u\in U}\left\{G(x,z,u) - G(\tilde{x},\tilde{z},u)\right\} \\
\le\ & \frac{1}{\lambda}\mu s_1(\beta) + \frac{\rho}{\lambda\beta}|s_1(\beta)| + \sigma^2 r(\beta) + u_{max}|s_2(\beta)| + K_R|(x,z) - (\tilde{x},\tilde{z})|
\end{aligned}
$$

for some $K_R > 0$ (where the last inequality holds due to the continuity of G and the fact that (x,z) and $(\tilde{x},\tilde{z})$ both lie in a compact subset of $\mathcal{D}$). So, taking $\omega(\beta) = \frac{1}{\lambda}\mu s_1(\beta) + \frac{\rho}{\beta}s_1(\beta) + \sigma^2 r(\beta) + u_{max}|s_2(\beta)|$ and $\omega_R(|(x,z)-(\tilde{x},\tilde{z})|) = K_R|(x,z)-(\tilde{x},\tilde{z})|$, we have that assumption (A3*) is satisfied since we assume that $\frac{1}{\beta}s(\beta) \to 0$ as $\beta \to 0$ and therefore $s_1(\beta) \to 0$ as $\beta \to 0$, as well. $\square$

Chapter 5

Conclusions and Future Work

1 Conclusions

We first developed a model which accurately captures both components of the environment in which the plant operates: the arrival of coal shipments and the spot price of electricity in the open market. We then formulated a control problem with the goal of maximizing the expected revenue of the plant until the first time the coal supply reached the shutdown level z_{min} and formally dervied the HJB equation associated with this problem

$$\begin{aligned} 0 = \delta f(x,z) - \frac{1}{\lambda}(\mu - x) f_x(x,z) - \frac{\sigma^2}{2} f_{xx}(x,z) - \sup_{u \in U} \{-u f_z(x,z) + G(x,z,u)\} \\ - \eta_1 \int_0^\infty \left(f(x+y,z) - f(x,z)\right) \beta_1 e^{-\beta_1 y}\, dy \\ - \eta_2 \int_{-\infty}^0 \left(f(x-y,z) - f(x,z)\right) \beta_2 e^{-\beta_2 y}\, dy \\ - \alpha\big[p_1\left(f(x,z+\zeta) - f(x,z)\right) + p_2\left(f(x,z+2\zeta) - f(x,z)\right)\big] \end{aligned}$$

where $\zeta = \tilde{\zeta} \wedge z_{max}$ and $2\zeta = 2\tilde{\zeta} \wedge z_{max}$ represent the amounts of coal actually added by a single and double train arrival, respectively. In Chapter 3, we proved that the value function (the maximum expected revenue we sought) is a viscosity solution to the HJB equation. Finally, we proved in Chapter 4 that the viscosity solution to this HJB equation is unique on any bounded domain in $\mathcal{D}$.

So, combining all of these results, we have shown that by finding the solution to the HJB equation above, we can solve the optimal control problem for the plant as it was posed in Chapter 1. As a consequence, we know that any burn rate policy u which achieves this optimal value is an optimal policy for the plant operator (though the policy is not necessarily unique).

2 Future Work

2.1 Numerical Solution Using Real Market Data

A particularly useful area of continued research is the application of these principles to real production plants and markets. Finding an approximate numerical solution for the value function and optimal burn policy using real spot price and production data would provide a tangible connection to the theoretical work presented here. Finding such a numerical approximation involves two stages. First, real data must be used to fit the parameters of the spot price model. Then, those parameters can be used to implement the Markov chain approximation methods of Kushner and Dupuis [9] to find the desired approximation to the solution of the HJB equation. We oultine in the following subsections a standard approach to this problem.

2.1.1 Parameter Fitting of the Spot Price Model

Gonzalez, Moriarty, and Palczewski [7] provide a method for parameter fitting in the spot price model used here. They provide results showing that the model provides a good approximation in two European energy markets for daily average prices. Oliver Meister showed in his MS thesis [10] that these methods can be applied to an American market, the MISO market, with equally good results. However, both of these works use only aggregated daily average prices, while a power plant operator needs to work on a shorter time scale, usually in the range of 15-30 minutes. Further examination of possible parameter fitting schemes on this shorter time scale would show whether the pricing model used here continues to be valid.

2.1.2 Production Function for a Coal Plant

In particular, electricity production at a coal plant is generally modeled by a linear relationship with some constant k kilowatts produced per ton of coal burned. With the spot price x being paid per kilowatt, this gives us a revenue function

$$R(x, z, u) = xku.$$

Such plants generally have long-term contracts with a coal supplier in which a constant amount is paid per month or year for a certain number of deliveries. That is, the cost rate for such a plant is constant, and we have cost function

$$C(x, z, u) = C.$$

So, we have a payoff function

$$G(x, z, u) = xku - C.$$

2.1.3 Finite Difference Scheme

Having parameters for the spot price process and a particular form for the payoff function, a numerical approximation to the value function and optimal control policy for our control problem could be sought using the method of Markov chain approximation as developed in the text by Kushner and Dupuis [9]. The particular HJB equation for this problem is

$$\begin{aligned} 0 = \delta v(x,z) - \frac{1}{\lambda}(\mu - x)v_x(x,z) - \frac{\sigma^2}{2}v_{xx}(x,z) - \sup_{u\in U}\{-uv_z(x,z) + xku - C\} \\ - \eta_1 \int_0^\infty (v(x+y,z) - v(x,z))\,\beta_1 e^{-\beta_1 y}\,dy \\ - \eta_2 \int_{-\infty}^0 (v(x+y,z) - v(x,z))\,\beta_2 e^{-\beta_2 y}\,dy \\ - \alpha\big[p_1\,(v(x,z+\zeta) - v(x,z)) + p_2\,(v(x,z+2\zeta) - v(x,z))\big] \end{aligned} \tag{5.1}$$

on the domain $\mathcal{D} = \mathbb{R} \times [z_{min}, z_{max}]$.

Discretizing the domain, we use a regular mesh in each direction. That is, we take the bounded domain $\tilde{\mathcal{D}} = [x_{min}, x_{max}] \times [z_{min}, z_{max}]$ with x_{min} and x_{max} chosen so that all observed spot prices in the dataset fall within $\tilde{\mathcal{D}}$. We then choose $N_x, N_z \in \mathbb{Z}^+$ and set $h_x := \frac{(x_{max}-x_{min})}{N_x}$ and $h_z := \frac{(z_{max}-z_{min})}{N_z}$. We partition $\tilde{\mathcal{D}}$ into a regular grid with mesh size h_x in the x-direction and mesh size h_z in the z-direction. So, we have the discretized domain $\tilde{\mathcal{D}}_{N_x,N_z} := \{(x_i, z_j)\,|\, i = 0, ..., N_x\,;\, j = 0, ..., N_z\}$ where $x_i = x_{min} + ih_x$ for each i and $z_j = z_{min} + jh_z$ for each j.

For clarity of notation, define $v_{i,j} := v(x_i, z_j)$, and define $\tilde{\zeta}$ to be the closest integer to $\frac{\zeta}{h_z}$. Further, we take a discretization of the probabilities of the exponential jump sizes to allow us to rewrite the two integrals in this HJB equation as sums by defining for $\ell \in \mathbb{Z}^+$

$$p_\ell^{(1)} := \int_{(\ell-1)h_x}^{\ell h_x} \beta_1 e^{-\beta_1 y} dy \qquad \text{and} \qquad p_m^{(2)} := \int_{(m-1)h_x}^{m h_x} \beta_2 e^{-\beta_2 y} dy,$$

and then defining, for some small $\epsilon > 0$, $N^{(1)}$ to be the smallest integer such that $p^{(1)}_{N^{(1)}} < \epsilon$ and $N^{(2)}$ to be the smallest integer such that $p^{(2)}_{N^{(2)}} < \epsilon$. Finally, take as a discretized form of the distributions of jump sizes in the spot price

$$\pi^{(1)}_\ell = \begin{cases} p^{(1)}_\ell, & \text{for } 1 \le \ell < N^{(1)} \\ 1 - \sum_{\ell=0}^{N^{(1)}-1} p^{(1)}_\ell, & \text{for } \ell = N^{(1)} \end{cases} \quad \text{and} \quad \pi^{(2)}_m = \begin{cases} p^{(2)}_m, & \text{for } 1 \le m < N^{(2)} \\ 1 - \sum_{m=0}^{N^{(2)}-1} p^{(2)}_m, & \text{for } m = N^{(2)}. \end{cases}$$

Using the forward difference approximation for the first derivatives and the central approximation for the second derivative, we then write the finite-difference approximation to the HJB (5.1) at each $(x_i, z_j) \in \tilde{\mathcal{D}}_{N_x,N_z}$ as

$$\begin{aligned} 0 = \delta v_{i,j} - \frac{1}{\lambda}(\mu - x_i)\frac{v_{i+1,j} - v_{i,j}}{h_x} - \frac{\sigma^2}{2}\frac{v_{i+1,j} - 2v_{i,j} + v_{i-1,j}}{{h_x}^2} \\ - \sup_{u \in U}\left\{-u\frac{v_{i,j+1} - v_{i,j}}{h_z} + x_i k u - C\right\} \\ - \eta_1 \sum_{\ell=1}^{N^{(1)}} \left(v_{i+\ell,j} - v_{i,j}\right)\pi^{(1)}_\ell - \eta_2 \sum_{m=1}^{N^{(2)}} \left(v_{i-m,j} - v_{i,j}\right)\pi^{(2)}_m \\ - \alpha\big[p_1(v_{i,(j+\tilde{\zeta})\wedge N_z} - v_{i,j}) + p_2(v_{i,(j+2\tilde{\zeta})\wedge N_z} - v_{i,j})\big]. \end{aligned} \tag{5.2}$$

In all boundary cases, we use a sticky boundary. That is, we take for $i < 0$, $v_{i,j} = v_{0,j}$; for $j < 0$, $v_{i,j} = v_{i,0}$; for $i > N_x$, $v_{i,j} = v_{N_x,j}$; and for $j > N_z$, $v_{i,j} = v_{i,N_z}$. Further, the boundary condition due to the termination of the process at the first time the coal supply reaches zero gives us $v(x_i, z_0) = 0$ for all i.

This finite-difference approximation is then used to find the value function at each iteration of the policy iteration method of Kushner and DuPuis. The optimal policy at each iteration is found by the simple maximization

$$u_{i,j} = \operatorname*{arg\,max}_{\{u_{min}, u_{max}\}} \left\{-u\frac{v_{i,j+1} - v_{i,j}}{h_z} + x_i k u\right\}$$

where $u_{i,j} := u(x_i, z_j)$ is the discretized version of the policy as with the value function above. That is,

$$u_{i,j} = \begin{cases} u_{min} & \text{if } \frac{v_{i,j+1} - v_{i,j}}{h_z} \ge x_i k \\ u_{max} & \text{if } \frac{v_{i,j+1} - v_{i,j}}{h_z} < x_i k \end{cases}.$$

Unfortunately, preliminary tests showed that this standard numerical method results in highly numerically unstable behavior. An upwind/downwind scheme in which the approximation of v_x is made using a forward difference when the coefficient on v_x, $-\frac{1}{\lambda}(\mu - x_i)$, is positive and a backwards difference when it is negative produces a more stable solution, but it is still not satisfactory. Further investigation into the cause of this instability and more appropriate numerical schemes is necessary.

2.2 Alternative Supply Process and Payoff Function

Another productive area of research would be replacing the particular forms of the supply process and/or payoff function used here. The action of a coal plant with the simple payoff function described in the previous section was the model driving the research in this paper. However, the methods used could be extended in several directions. Directly using these methods, but with different production and cost functions leading to a different form of the payoff function, but one that is still continuous, could lead to better numerical results than those found thus far. Working with a different supply process, whether continuous or discontinuous, could also prove worthwhile. More ambitiously, one could examine whether the restriction to continuous payoff functions could be relaxed, as we have not proven that this condition is necessary, merely sufficient.

2.3 Multiple Nodes

The problem examined in this paper was, for analytical tractability, limited to a single spot price at a single node on the national electricity grid. However, in reality, plants are able to sell to dozens of different nodes, each with its own spot price process. This extension would require a significantly different payoff structure as well as an examination of the correlation of the spot price at different nodes (as a cursory examination shows that these prices are far from independent). Even an extension to two nodes would require another component of the control representing the distribution of the power being produced between the available nodes. So, while this area of research would be of great interest for practical applications of the model presented here, it presents a great deal of increased complexity to be dealt with.

Bibliography

[1] Arisawa, M. A new definition of viscosity solutions for a class of second-order degenerate elliptic integro-differential equations. *Ann. Inst. H. Poincaré Anal. Non Linéaire 23*, 5 (2006), 695–711.

[2] Barles, G., and Imbert, C. Second-order elliptic integro-differential equations: viscosity solutions theory revisited. *Annales de lInstitut Henri Poincare (C) Non Linear Analysis 25*, 3 (2008), 567585.

[3] Benth, F. E., Kallsen, J., and MeyerBrandis, T. A nongaussian ornstein-uhlenbeck process for electricity spot price modeling and derivatives pricing. *Applied Mathematical Finance 14*, 2 (2007), 153169.

[4] Benth, F. E., Saltyte., B. J., and Koekebakker, S. *Stochastic modelling of electricity and related markets.* World Scientific, 2012.

[5] Bertsekas, D. P., and Shreve, S. E. *Stochastic optimal control: the discrete time case: the discrete time case.* Elsevier, 1978.

[6] Gerber, H. U., and Shiu, E. S. W. On optimal dividend strategies in the compound poisson model. *North American Actuarial Journal 10*, 2 (2006), 7693.

[7] Gonzalez, J., Moriarty, J., and Palczewski, J. *Bayesian Calibration and Number of Jump Components in Electricity Spot Price Models.* PhD thesis, 2016.

[8] InvestExcel.net. http://investexcel.net/brent-oil-price-history/.

[9] Kushner, H. J., and Dupuis, P. *Numerical methods for stochastic control problems in continuous time.* Springer, 2001.

[10] Meister, O. Calibration of a stochastic price model for american electricity markets. Master's thesis, 2018.

[11] MISO. *http://www.misoenergy.org/*. MISO Energy Homepage.

[12] Pallaschke, D., and Rolewicz, S. *Foundations of mathematical optimization*. Polish Committe for Scientific Research, 1997.

[13] Pham, H. *Continuous-time stochastic control and optimization with financial applications*. Springer, 2010.

Appendix

We collect here the proofs to three lemmas concerning the ordering of the stopping times used in Chapter 3.

Lemma 5.1. *Let* π_n*,* h_n *be defined as in Proposition* 3.1*. Then,* $\mathbb{P}[\pi_n \leq h_n] \to 0$ *as* $n \to \infty$*.*

Proof. For each n, $Y_0^{(n)}(t)$ is defined by the SDE

$$dY_0^{(n)}(t) = -\frac{1}{\lambda}(Y_0^{(n)}(t) - \mu) + dW(t) \quad , \quad Y_0^{(n)}(0) = x_n. \tag{5.3}$$

Let $\bar{Y}_0(t)$ be defined by the same SDE with $\bar{Y}_0(0) = \bar{x}$. Since $x_n \to \bar{x}$, for any $\epsilon > 0$, there exists some N such that $x_n \in (\bar{x}-\epsilon, \bar{x}+\epsilon)$ for all $n \geq N$. Let $\pi_n^- := \inf\{r \geq 0 \,\big|\, Y_0^{(n)}(r) < \bar{x} - \rho\}$ and $\pi_n^+ := \inf\{r \geq 0 \,\big|\, Y_0^{(n)}(r) > \bar{x} + \rho\}$. Note that for every n, the random behavior of $Y_0^{(n)}$ is determined by the same Brownian motion process $W(t)$. Define $Y_0^+(t)$ as in (5.3) with $Y_0^+(0) = \bar{x} + \epsilon$. So, for every $n \geq N$, $Y_0^+(t) \geq Y_0^{(n)}(t)$ for all t for each $\omega \in \Omega$, and thus $\pi_n^+ \geq \pi^+ := \inf\{r \geq 0 \,\big|\, Y_0^+(r) \geq \bar{x} + \rho\}$. Since the process Y_0^+ is just a mean-reverting diffusion process and $h_n \to 0$, $\lim_{n\to\infty} \mathbb{P}[\pi^+ \leq h_n] = 0$ and therefore, since $\pi_n^+ \geq \pi^+$, $\lim_{n\to\infty} \mathbb{P}[\pi_n^+ \leq h_n] = 0$. By defining $Y_0^-(t)$ as in (5.3) with $Y_0^-(0) = \bar{x} - \epsilon$ and making a similar argument, it can be shown that for $n \geq N$, $\pi_n^- \geq \pi^-$ and $\mathbb{P}[\pi^- \leq h_n] = 0$, and therefore $\lim_{n\to\infty} \mathbb{P}[\pi_n^- \leq h_n] = 0$. Concluding from the fact that $\pi_n = \pi_n^- \wedge \pi_n^+$, we have the desired result. □

Lemma 5.2. *Let* ξ*,* h_n *be defined as in Proposition* 3.1*. Then,* $\mathbb{P}[\xi \leq h_n] \to 0$ *as* $n \to \infty$*.*

Proof. Recall that $\xi = \xi_1 \wedge \xi_2 \wedge \xi_3$ where $\xi_1 = \min\{t > 0 \,\big|\, L_1(t) - L_1(t-) \neq 0\}$, $\xi_2 = \min\{t > 0 \,\big|\, L_2(t) - L_2(t-) \neq 0\}$, and $\xi_3 = \min\{t > 0 \,\big|\, Z^{(n)}(t) - Z^{(n)}(t-) \neq 0\}$. So,

$$\mathbb{P}[\xi \leq h_n] = \mathbb{P}[\xi_1 \wedge \xi_2 \wedge \xi_3 \leq h_n] \leq \mathbb{P}[\xi_1 \leq h_n] + \mathbb{P}[\xi_2 \leq h_n] + \mathbb{P}[\xi_3 \leq h_n]. \tag{5.4}$$

We first consider $\mathbb{P}\left[\xi_1 \le h_n\right]$. By the definition of L_1 as a Poisson process with intensity λ_1, the first arrival time, ξ_1 is an exponential random variable with rate λ. So,

$$\mathbb{P}[\xi_1 \le h_n] = 1 - e^{-\lambda_1 h_n},$$

and since $h_n \to 0$ as $n \to \infty$, we have that $\mathbb{P}[\xi_1 \le h_n] \to 0$ as $n \to \infty$.

Since ξ_2 and ξ_3 are also exponentially distributed first arrival times with a constant intensity, a similar argument shows that $\mathbb{P}[\xi_2 \le h_n] \to 0$ and $\mathbb{P}[\xi_3 \le h_n] \to 0$ as $n \to \infty$. So, we have from (5.4) that

$$\lim_{n\to\infty} \mathbb{P}[\xi \le h_n] = 0.$$

□

Lemma 5.3. *Let τ_n, h_n be defined as in Proposition 3.1. Then, $\mathbb{P}[\tau_n \le h_n] \to 0$ as $n \to \infty$.*

Proof. Recall that $\tau_n := \min\{t > 0 \mid Z^{(n)} \not\in (\tilde{z}_n, \tilde{z}^n)\}$. We consider two possible cases. First, let $\overline{\tau_n} := \min\{t > 0 \,|\, Z^{(n)}(t) \ge \tilde{z}^n\}$. Since we assume that $z_n < \tilde{z}^n$, $Z^{(n)}(t)$ can only be greater than or equal to $\tilde{z}^n$ if at least one coal shipment has arrived by time t, that is if $\xi_3 \le t$. As shown in Lemma 5.2, $\mathbb{P}[\xi_3 \le h_n] \to 0$ as $n \to \infty$. Therefore, $\mathbb{P}[\overline{\tau_n} \le h_n] \to 0$ as $n \to \infty$.

Next, we consider $\underline{\tau_n} := \min\{t > 0 \,|\, Z^{(n)}(t) \le \tilde{z}_n\}$. Since, by construction, $z_n > \tilde{z}_n$ and the control u is constant, $\underline{\tau_n} \ge \frac{z_n - \tilde{z}_n}{u} > 0$. So, $\mathbb{P}[\underline{\tau_n} \le h_n] \to 0$. We therefore have that

$$\lim_{n\to\infty} \mathbb{P}[\tau_n \le h_n] = \lim_{n\to\infty} \mathbb{P}[\,\underline{\tau_n} \wedge \overline{\tau_n} \le h_n] = 0.$$

□

CURRICULUM VITAE

Charles Beer

EDUCATION

- PhD., Mathematical Sciences, University of Wisconsin-Milwaukee — May 2019
  Dissertation Title: *A Stochastic Control Model for Energy Producers*
  GPA: 3.80/4.00
- M.S., Mathematical Sciences, University of Wisconsin-Milwaukee — 2013
- B.S., Mathematics, University of Nebraska-Lincoln (Honors Program) — 2011
  Minors in Computer Science, Physics, and Economics
  GPA: 3.65/4.00

EXPERIENCE

- **Research/Project Assistant, UW-Milwaukee** — 2014-2016
  - Developed a stochastic control model for a local energy producer
  - Created a Markov chain model for arrival of coal supply
  - Fit parameters in a multifactor stochastic model for energy prices

  Supervisor: Dr. Anthony Ross, UW-Milwaukee – Lubar School of Business
  Rockwell Automation Endowed Chair in Supply Chain Management
  Founding Director, Supply Chain Management Institute
- **Graduate Teaching Assistant, UW-Milwaukee** — 2011-2014 & 2015-present
  - Independently taught 2-3 courses of approx. 30 students per semester
  - Communicated understanding of mathematics and statistics to non-experts
  - Developed teaching materials for use by a large group of instructors
  - Provided guidance to less experienced teaching assistants
  - Courses Taught: Intermediate Algebra, College Algebra, Trigonometry, Pre-Calculus, Calculus II, Intro to Statistics, and Finite Math

  Supervisor: Dr. Allen Bell, UW-Milwaukee
  Mathematical Sciences Department Chair
- **Undergraduate Research Assistant, Los Alamos National Labs (in physics)** — Summer 2005
  - Assisted with running and maintaining a magnetic plasma containment experiment (extensive electronics and vacuum system work)
  - Programmed, ran, and maintained a computer simulation system
  - Assisted in the compilation, writing, and submission of official reports

  Supervisors: Dr. Thomas Intrator and Dr. Glen Wurden

SELECTED COURSEWORK

Stochastic Control and Financial Applications, Stochastic Analysis, Stochastic Processes, Optimal Stopping, Dynamic Pricing, Stochastic Games and Backward SDEs, Mathematical Statistics (PhD level), Measure Theoretic Probability, Linear Optimization

SELECTED HONORS AND AWARDS

- Research Excellence Award, UW-Milwaukee 2016
- GAANN Fellowship, UW-Milwaukee 2015
- Chancellor's Award, UW-Milwaukee 2011-2015
- Putnam Exam Team, University of Nebraska-Lincoln (Personal Score: 10) 2011
- Regents' Scholarship, UN-Lincoln (full tuition merit-based scholarship) 2003-2007
- Kiewit Scholarship, UN-Lincoln (full cost of living merit-based scholarship) 2003-2007

INVITED PRESENTATIONS

- **"A Stochastic Control Model for Energy Producers"** Jun 2018
  A Symposium on Optimal Stopping (In Memory of Larry Shepp)
  Rice University, Houston, TX
- **"Markov Chain Monte Carlo Methods"** Mar 2017
  UW-Milwaukee Math Graduate Student Colloquium
  UW-Milwaukee, Milwaukee, Wisconsin
- **"Electricity Pricing Models Including Jump Components"** Nov 2016
  UW-Milwaukee Economics Research Seminar
  UW-Milwaukee, Milwaukee, Wisconsin

RESEARCH INTERESTS

- Interdisciplinary and Applied Mathematics
- Stochastic Analysis and Control
- Probability Theory
- Statistics