Dept. of Math. Univ. of Oslo
Pure Mathematics No. 4
ISSN 0806–2439 January 2008

VISCOSITY SOLUTIONS THEORY FOR TOLLING AGREEMENTS AND SWING OPTIONS

O. WALLIN

Abstract. We investigate a constrained stochastic control problem connected to a financial contract representing a virtual factory. Commonly known as tolling agreements, these contracts are traded in free energy markets and include exercise flexibility in volume as well as in timing. Allowing for very general models for jump-diffusion processes with possibly time-dependant jump intensity, we study the control problem under the dynamic programming framework. After rigorously proving the dynamic programming principle, we define viscosity solutions of the associated Hamilton-Jacobi Bellman equation, and show the value is the unique solution of the equation. In fact, we give an original proof of a strong comparison principle using the maximum principle for semicontinuous functions that avoids some of the problems connected with unbounded Lévy measures that have been investigated in recent research of several authors.

1. Introduction

During the last decade or so, gradual liberalization of the energy markets has greatly impacted the industry. Market based pricing and inelasticity of demand under rising prices, especially on electricity, has lead to price spikes in situation where the production is insufficient or interrupted. To counter the uncertainties and transfer a part of the price risk several exotic contracts are nowadays traded over the counter between market participants. Although some of the contracts are similar to contracts from financial markets, many have features that are unique to the energy sector. Namely, many contracts include volumetric and timing optionality.

A *tolling agreement* is a financial contract which imitates the returns from a production plant. A typical example is a gas fired power plant. The holder of the contract has the right, within specified limits, at any moment to control how much electricity is being produced. There is also a maximum amount $\overline{C}$ that can be produced during the contracts lifetime. The holder has to buy the gas from the market with the current price S_G, and receives in turn the price S_E for the end product. Here, the price of the raw material (gas) should correspond to the amount that is needed to produce one unit of the end product (MWh of electricity). For gas and electricity, this is the price of gas times some heating rate, which is specified in the contract. We wish to emphasize that we are interested in a financially settled contract, where no real production takes place. Thus we do not need to consider complicating factors such as ramp up times.

Date: January 22, 2008.

The author thanks Fred Espen Benth for suggesting the topic and several discussions regarding this work.

Furthermore, we can simply take as starting point the definition of the payoff of the contract. In short, the payoff of a tolling agreement is simply defined as

$$\int_0^T \nu(t)(S_E(t) - S_G(t))dt,$$

where the production intensity $\nu(\cdot)$ and the total volume produced $C_\nu(t) = \int_0^t \nu(s)ds$ satisfy the bounds $\underline{u} \leq \nu(t) \leq \overline{u}$ and $0 \leq C_\nu(T) \leq \overline{C}$. We assume the contract is settled at time T, and we will therefore disregard discounting throughout this article. As a definition of value we take the value function of the control problem which seeks to maximize the expected profits under a given probability measure. In the case of electricity markets, arbitrage arguments become difficult to apply to a contract with this much flexibility due to non-storability of electricity and the nature and illiquidity of other derivatives traded in the market. Furthermore, utility based pricing methods could well be justified. However, we will not delve into this, and simply assume we are given a probability measure under which we seek to maximize the expectation. We refer the interested reader to the sources given below for more information.

Our goal here is to develop a mathematically sound framework for the pricing of such contracts that simultaneously is general enough to include the wild, unpredictable behavior of market prices and simple enough so that at least numerical results can be obtained with sufficient efficiency and accuracy. Especially, we wish to allow for time-dependent jump intensity of the price processes as this is what is observed in several electricity markets, see Benth, Kallsen and Meyer-Brandis [5] for a related model. In addition to the theoretical value this work has in itself, our results make ground for proofs of convergence of such numerical approximations as exemplified in [3], [11] for the diffusion case and [7] for integrodifferential equations.

One the main features of our work is the proof of the comparison principle given in section 4. To the best of our knowledge, the proof is new in the framework of integrodifferential equations. Our hope is that our approach, exemplified here for the Hamilton-Jacobi-Bellman equation of our problem, settles some questions concerning the applicability for integrodifferential equations of the well known Ishii's lemma which is the standard apparatus in the pure differential setting. Especially, we carefully study integrability properties that can cause problems when the associated Lévy measures are unbounded. We mention that previous work paying special attention on this issue include [2], [1], [4] and [15].

To be more precise, let us briefly compare the approach developed here to those presented in the articles mentioned above. First, we do not need any new definitions of a viscosity solution to begin with. While another characterization for a viscosity solutions is given in terms of semijets, this is a quite natural extension of the same concept from the diffusion setting. Previous papers in this field use splitting of the integral operator in two parts with test functions used near the singularity and semicontinuous sub- and supersolutions away from it. Here, we do not apply such splitting but use semicontinuous functions and viscosity sub- and supersolution properties. Second, the original "maximum principle for semicontinuous functions", also known as Ishii's lemma is used here, so no new maximum principles is needed. All in all, it turns out that integro-differential equations are not too different from the pure differential case, after all.

As our intention is to study the pricing of the contract described above as a given control problem, we do not give a full review of energy markets here. We refer the interested reader to the books by Eydeland and Wolyniec [10], and Geman [12]. See also the recent PhD theses by Kluge [18], Ludkovski [20], Vehviläinen [26] and Wilhelm [27]. Mathematical papers dealing specifically with tolling agreements, but some in a slightly deferent setting, include [8] and [23]. One should note that our framework will include as a special continuous-time *swing options* previously studied by Hambly, Howison and Kluge [13], Keppo [17], Lund and Ollmar [21]. A swing option is a purchase contract on power with similar flexibility but no obligation to deliver the raw material, see the references for more information. Finally, standard references for stochastic control and viscosity solutions theory in the diffusion case include [9], [11].

The rest of this article is organized in the following way. In the next section, we define the value of a tolling agreement in terms of a control problem. We then proceed to give full proof of the dynamic programming principle for the problem. Here we borrow ideas from [16] and [25]. One should note that we work under the strong formulation of the control problem, i.e. on a given probability space. After stating the related Hamilton-Jacobi Bellman equation, we define viscosity solutions of the equation and show that the value satisfies the viscosity property in section 3. Finally, after a few important definitions we explain some of the difficulties associated with integro-differential equations in section 4, and prove the comparison principle in a way that avoids these problems. The appendix contains some of the more technical results used in this paper.

2. Stochastic control and dynamic programming principle

In this section, we first give specific but quite general assumptions on the class of models for price processes we allow, and state the control problem. We donate the price processes of the raw material and the end product by $S_1 = (S_1(t))_{t\in[0,T]}$ and $S_2 = (S_2(t))_{t\in[0,T]}$, respectively, and by $S = (S_1, S_2)$ the vector of the joint dynamics. In addition to these (uncontrolled) price processes, the state of our control system includes the total volume $C_\nu = (C_\nu(t))_{t\in[0,T]}$ of production.

More precisely, let $(\Omega, \mathcal{F}, \mathbb{P})$ be a given complete probability space, and let W be a standard Brownian motion in $\mathbb{R}^2$. Let J be a (possibly nonhomogeneous) Poisson random measure on $[0,T]\times\mathbb{R}^2$ with intensity measure $m(dt, dz) = m_t(dz)dt = \eta(t)m(dz)dt$ where η is a time density satisfying

$$|\eta(t) - \eta(t')| \le K|t - t'|,$$

for some $K > 0$ and m is a measure satisfying

$$\int (1 \wedge z^2)m(dz) < \infty.$$

Furthermore, we denote with $\overline{J}(du, dz) = J(du, dz) - \rho(u)m(dz)du$ the compensated jump measure of J. Such jump measures correspond to a process with independent increments and absolutely continuous characteristics (PIIAC), see [14] or [19] for precise definitions and properties. Now, let $S^{t,s} = (S^{t,s}(u))_{u\in[t,T]} \in \mathbb{R}^2$ be the solution of the stochastic differential equation

$$dS(u) = b(u, S(u))du + \sigma(u, S(u))dW(u) + \int_{\mathbb{R}^2} \gamma(u, S(u), z)\overline{J}(du, dz),$$

on this probability space, given the initial condition $S(t) = s = (s_1, s_2)$. For each (t, s), we further let $S^{t,s}(u) = 0$ for $u \in [0, t)$ so that $S^{t,s}$ is defined on the whole time interval $[0, T]$. Moreover, we let

$$C_\nu^{t,c}(u) = c + \int_t^u \nu(r)dr.$$

We assume that the coefficients of S satisfy some standard Lipschitz and growth properties. More precisely, we assume that $b : [0,T]\times\mathbb{R}^2 \to \mathbb{R}^2$, $\sigma : [0,T]\times\mathbb{R}^2 \to \mathbb{R}^{2\times2}$, $\gamma : [0,T]\times\mathbb{R}^2\times\mathbb{R}^2 \to \mathbb{R}^{2\times2}$ satisfy

$$|b(t, s) - b(t, s')| + |\sigma(t, s) - \sigma(t, s')| \le K|s - s'|, \tag{2.1}$$

$$|\gamma(t, s, z) - \gamma(t, s', z)| \le \rho(z)|s - s'|, \tag{2.2}$$

$$|\gamma(t, s, z)| \le \rho(z)(1 + |s|). \tag{2.3}$$

Global Lipschitz conditions and continuity of b, σ with respect to (t, s) yield the global linear growth condition

$$|b(t, s)| + |\sigma(t, s)| \le K(1 + |s|). \tag{2.4}$$

Then, it is well known (see Jacod and Shiryaev [14], Theorem III.2.32) that the equation has a pathwise unique strong solution on the given probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Throughout this article we let $\mathbb{F} = \mathbb{F}^{W,J}$ be the completed natural filtration generated by the driving Brownian motion and jump measure. We note that it would be more natural from a practical point of view to assume the information is generated by the prices rather than the driving process and take $\mathbb{F} = \mathbb{F}^S$, as prices are what is truly observed. However, one can see that this case could be worked out here as well by using the fact that $\mathbb{F}^S \subset \mathbb{F}^{W,J}$.

Let us denote by $U = [\underline{u}, \overline{u}]$ the state space of the control ν, so that our control set is

$$\mathcal{V} := \{\nu \text{ progressively measurable } : \nu(t) \in U \text{ for a.e. } t \in [0,T]\}.$$

It is trivial that $\mathcal{V}$ is a subset of $L^p([0,T] \times \mathbb{R})$ for any $1 \leq p \leq \infty$. We shall equip the space of controls with the norm topology of $L^2([0,T] \times \mathbb{R})$, but note that by boundedness of $\nu \in \mathcal{V}$, the L^p-norms are equivalent on the control space. Let us denote by

$$\tau_{\overline{C}} := \inf\{t > 0 | C_\nu(u) = \overline{C}\}$$

the first time we have used up all of our production rights, and let $\mathcal{T}_{t,T}$ denote the set of stopping times θ such that $t \leq \theta \leq \tau_{\overline{C}} \wedge T$. Apart from some compactness to hold on the space $(\mathcal{V}, \|\cdot\|_{\mathcal{V}})$, the following properties of the control space are crucial for dynamic programming to work:

Property 1 Stability under concatenation: if $\nu_1, \nu_2 \in \mathcal{V}$ and $\theta \in \mathcal{T}_{t,T}$, then $\bar{\nu} := \nu_1 1_{[0,\theta)} + \nu_1 1_{[\theta,T]}$ is in $\mathcal{V}$.

Property 2 Stability under measurable selection: for any $\theta \in \mathcal{T}_{t,T}$ and any measurable map

$$\phi : (\Omega, \mathcal{F}(\theta)) \mapsto (\mathcal{V}, \mathcal{B}_{\mathcal{V}}),$$

there exists $\nu \in \mathcal{A}$ such that

$$\phi = \nu \quad \text{on } [\theta, T] \times \Omega, dt \times \mathbb{P} \text{ almost everywhere.}$$

That Property 1 holds is trivial. To see that Property 2 is in effect, it is sufficient to verify that $(\mathcal{V}, \|\cdot\|_{L^2})$ is a separable metric space (see Soner and Touzi [25]). But since $(\mathcal{V}, \|\cdot\|_{L^2})$ is a closed subset of L^2, this follows from the separablility of L^2. For this, it is enough that the underlying sigma-algebra $\mathcal{F} = \sigma(\mathcal{F}^W \times \mathcal{F}^J)$ is countably generated. But it is well known that this is the case for Brownian motion, and for the jump measure J associated with a PIIAC pure jump process this follows from the fact that there is at most a countable number of jumps. At this stage, we note that Property 2 is needed to ensure it is possible to define an admissible control by

$$\bar{\nu}(t,\omega) := (\phi(\omega))(t,\omega) 1_{t \geq \theta}(\omega) + \nu(t,\omega) 1_{t < \theta}(\omega) \tag{2.5}$$

for some $\nu \in \mathcal{V}$ and ϕ as above.

In this paper, we study a slightly generalized control problem, mainly for notational convenience, compared to the tolling agreement mentioned in the introduction. For this, let $f : \mathbb{R}^2 \to \mathbb{R}$ be a function satisfying

$$|f(s) - f(s')| \leq K|s - s'|, \tag{2.6}$$

$$|f(s)| \leq K(1 + |s|), \tag{2.7}$$

and define the *reward functional* by

$$R(t,s,c,\nu) := \mathbb{E}\Big[\int_t^{T \wedge \tau_{\overline{C}}} \nu(u) f(S^{t,s}(u)) du\Big] \tag{2.8}$$

$$= \mathbb{E}\Big[\int_t^T \nu(u) H(C_\nu^{t,c}(u)) f(S^{t,s}(u)) du\Big] \tag{2.9}$$

where $H(x) = 1_{x \leq \overline{C}}(x)$. This definition includes the tolling agreement with $f(s) = (s_1 - s_2)$ and the swing option with $f(s) = (s_1 - K_1)$ for some constant $K > 0$. We note that as the contract is settled at the end of the period, there is no loss of generality in considering the case with zero interest rate. We consider the classical stochastic control problem of finding V and $\nu^* \in \mathcal{V}$ such that

$$v(t,s,c) := \sup_{\nu \in \mathcal{V}} R(t,s,c,\nu) = R(t,s,c,\nu^*).$$

We denote by $\overline{\mathcal{O}_T} = [0, T] \times \mathbb{R}^2 \times [0, \overline{C}]$ the domain of definition for v, and $\mathcal{O}_T = [0, T) \times \mathbb{R}^2 \times [0, \overline{C})$. Note here that

$$v(t, s, \overline{C}) = 0, \quad (t, s) \in [0, T] \times \mathbb{R}^2$$

by definition. Our first goal is to give a full proof of the following strong form of the dynamic programming principle.

2.10. **Theorem.** *(DP1) For all $\nu \in \mathcal{V}$ and $\theta \in \mathcal{T}_{t,T}$, set of stopping times in $[t, T \wedge \tau_{\overline{C}}]$:*

$$v(t, s, c) \leq \sup_{\nu \in \mathcal{V}} \mathbb{E}\Big[\int_t^\theta \nu(u) f(S^{t,s}(u)) du + v(\theta, S^{t,s}(\theta))\Big]. \tag{2.11}$$

(DP2) For all $\epsilon > 0$, there exists $\nu^\epsilon \in \mathcal{V}$ such that for all $\theta \in \mathcal{T}_{t,T}$:

$$v(t, s, c) + \epsilon \geq \mathbb{E}\Big[\int_t^\theta \nu^\epsilon(u) f(S^{t,s}(u)) du + v(\theta, S^{t,s}(\theta))\Big]. \tag{2.12}$$

Combining (DP1) and (DP2), we obtain what is more commonly called the (weak form of) the dynamic programming principle:

$$v(t, s, c) = \sup_{\nu \in \mathcal{V}} \mathbb{E}\Big[\int_t^\theta \nu(u) f(S^{t,s}(u)) du + v(\theta, S^{t,s}(\theta))\Big] \tag{2.13}$$

Let $Y = (Y(t))_{t \in [0,T]}$ be a measurable stochastic process with sample paths that are right continuous and have left limits (RCLL). Here, it is useful to think of Y as an infinite dimensional random variable taking values in the *Skorohod space* $\mathbb{D}([0, T])$ of RCLL functions. The Skorohod space equipped with the *Skorohod metric* is a Polish space, see Jacod and Shiryaev [14] or Partasarathy [22]. Moreover, the Borel sets $\mathcal{B}(D([0, T]))$ are generated by the *cylinder sets*, i.e sets of the form

$$\mathcal{A} = \{\omega \in \mathbb{D}([0, T]) : (\omega(t_1), ..., \omega(t_n)) \in A\},$$

where $t_i \in [0, T]$, $i = 1, ..., n$ and $A \in \mathcal{B}(\mathbb{R}^n)$, see Partasarathy [22], Chapter VII, Theorem 7.1. For any such function $\omega \mapsto Y(\cdot, \omega)$, define the function Y^s by association with the stopped process $t \mapsto Y^s(t) := Y(t \wedge s)$. Following Kabanov and Klüppelberg [16], we shall need the following facts about conditional distributions. Let ξ and ζ be two random variables taking values in Polish spaces $\mathcal{X}$ and $\mathcal{Y}$, with Borel sigma-algebras $\mathcal{B}_\mathcal{X}$, $\mathcal{B}_\mathcal{Y}$. Then ξ admits a regular conditional probability distribution given $\zeta = y$ denoted by $\mathbb{P}_{\xi|\zeta}(A, y)$ and

$$\mathbb{E}[f(\xi, \zeta)|\zeta] = \int f(x, y)(dx, y)\Big|_{y=\zeta} \qquad \text{(a.s.)} \tag{2.14}$$

See section I.1.3 in Da Prato and Zabczyck [24] as well. Now, let Z be the unique pure jump PIIAC process corresponding to the jump measure J, so that $\mathcal{F}^J = \mathcal{F}^Z$. Denoting $Y(t) = (W(t), Z(t))$, we shall apply the above relation to the random variables $\zeta = Y^\theta$ and $\xi = Y - Y^\theta$. Then, $\mathbb{P}_{\xi|\zeta}(A, y)$ admits a version which is independent of y and coincides with $\mathbb{P}$. Note furthermore, that $Y = \zeta + \xi$.

2.15. **Lemma.** *Let $(t, s) \in [0, T] \times \mathbb{R}^2$ and $\nu \in \mathcal{V}$. Then, for any $\theta \in \mathcal{T}_{t,T}$,*

$$\mathbb{E}\Big[\int_\theta^T H(C_\nu^{t,c}(u)) \nu(u) f(S^{t,s}(u)) du \big| \zeta\Big] = R(\theta, S_\nu^{t,s}(\theta), C_\nu^{t,c}(\theta), \nu), \ \text{a.s.} \tag{2.16}$$

Proof. Let $\theta \in \mathcal{T}_{t,T}$. We start by noting that by pathwise uniqueness of the solutions of S, for $u > \theta$

$$S^{t,s}(u) = S^{\theta, S^{t,s}(\theta)(u)} + \int_\theta^u b(r, S^{\theta, S^{t,s}(\theta)}(\theta)(r)) dr + \int_\theta^u \sigma(r, S^{\theta, S^{t,s}(\theta)}(r)) dW(r) \tag{2.17}$$

$$+ \int_\theta^u \int_{\mathbb{R}^2} \gamma(r, S^{\theta, S^{t,s}(\theta)}(r), z) \overline{J}(dr, dz) \tag{2.18}$$

$$C_\nu^{t,c}(u) = C_\nu^{\theta, C_\nu^{t,c}(\theta)}(u). \tag{2.19}$$

Now, let X^S, X^C be random variables taking values in $\mathbb{R}^2$ and $[0, M]$, respectively. Considering cylinder sets it can be noted that the paths $C_\nu^{\theta,X^C}(\cdot), S^{\theta,X^S}(\cdot)$ and $\nu(\cdot)$ are Y-measurable. By Dynkin's lemma there exist functions $\tilde{S}, \tilde{C}, \tilde{\nu} : [0,T] \times D([0,T]) \to \mathbb{R}$ such that

$$\begin{aligned} S^{\theta,X^S}(\cdot) &= \tilde{S}.(Y) = \tilde{S}.(\zeta + \xi), \\ C_\nu^{\theta,X^C}(\cdot) &= \tilde{C}.(Y) = \tilde{C}.(\zeta + \xi) \\ \nu(\cdot) &= \tilde{\nu}.(Y) = \tilde{\nu}.(\zeta + \xi), \end{aligned}$$

almost everywhere under $dt \times \mathbb{P}$ where $\zeta = Y^\theta$, $\xi = Y - Y^\theta$ are as in the discussion preceding the lemma. We now combine the above putting $X^S = S^{t,s}(\theta)$, $X^C = C_\nu^{t,c}(\theta)$ with the fact that these are measurable with respect to Y^θ, so there are functions $\hat{S}^\theta, \hat{C}^\theta$ such that

$$\begin{aligned} S^{t,s}(\theta) &= \hat{S}^\theta(\zeta), \\ C_\nu^{t,c}(\theta) &= \hat{C}^\theta(\zeta) \end{aligned}$$

Again, with the help of the cylinder sets and the π-system criterion it can be verified that ζ and ξ are independent as discussed above. Then, the claim follows by conditional Fubini's theorem and (2.14): see also Propositions 1.11 and 1.12 in Chapter 1 of Da Prato and Zabzyck [24]. $\square$

The following lemma states some standard estimates for the two-dimensional price process S that will be useful for us. The proofs are omitted, since these are given in Pham for the case where the jump measure has a time-independent Lévy measure m. The adaptation to our setting is straightforward since the time density η is continuous, and therefore also bounded on $[0, T]$.

2.20. **Lemma.** *For any $k \in [0, 2]$, there exists $K > 0$ such that for all $0 \le t < t' \le T$, $s, s' \in \mathbb{R}^2$, and stopping times $t \le \theta \le t' \wedge \tau_{\overline{C}}$, we have*

$$\mathbb{E}[|S^{t,s}(\theta)|^k] \le K(1 + |s|^k) \tag{2.21}$$

$$\mathbb{E}[|S^{t,s}(\theta) - s|^k] \le K(1 + |s|^k)|t - t'|^{k/2} \tag{2.22}$$

$$\mathbb{E}[\sup_{t \le \theta \le t'} |S^{t,s}(\theta) - s|]^k \le K(1 + |s|^k)|t - t'|^{k/2} \tag{2.23}$$

$$\mathbb{E}[|S^{t,s}(\theta) - S^{t,s'}(\theta)|^k] \le K|s - s'|^k \tag{2.24}$$

$$\mathbb{E}[|S^{t,s}(u) - S^{t',s}(u)|^k] \le K(1 + |s|^k)|t - t'|^k \quad \text{for } u \ge t'. \tag{2.25}$$

Especially, we have from global Lipschitz properties for f that the value function also satisfies

$$|v(t, s, c) - v(t, s', c)| \le K|s - s'| \tag{2.26}$$

for some $K > 0$.

Measurability properties play a crucial role in proving dynamic programming principles. Fortunately, despite the constrained nature of our problem we have that the reward function R is even continuous, in the sense of the following lemma.

2.27. **Lemma.** *The map*

$$(t, s, c, \nu) \in \overline{\mathcal{O}_T} \times \mathcal{V} \mapsto R(t, s, c, \nu)$$

is continuous, where continuity in $\nu \mapsto R(t, s, c, \nu)$ is under the $L^2([0,T] \times \mathbb{R})$-norm on the space $\mathcal{V}$.

Proof. Let us prove continuity for each component separately. First, let $t < t'$ and let s and ν be fixed. Then

$$|R(t,s,c,\nu) - R(t',x,c,\nu)|^2 \leq \overline{C}^2 \int_t^{t'} \mathbb{E}[|f(S^{t,s}(u))|^2]du + \overline{C}^2 \int_{t'}^T \mathbb{E}[|f(S^{t,s}(u)) - f(S^{t',s}(u))|^2]du$$
$$\leq K_1(1+|s|^2)|t-t'|^2 + K_2 \int_{t'}^T \mathbb{E}[|S^{t,s}(u) - S^{t',s}(u)|^2]du$$

where we have used the upper bound on ν, growth condition for f and moment estimate (2.21). Now inequality (2.25) implies

$$|R(t,s,c,\nu) - R(t',x,c,\nu)|^2 \leq K(1+|s|^2)|t-t'|^2. \tag{2.28}$$

Next, to prove continuity in s alone, we note with t, c, and ν fixed that

$$|R(t,s,c,\nu) - R(t,s',c,\nu)|^2 \leq \mathbb{E}\Big[\int_t^T |f(S^{t,s}(u)) - f(S^{t,s'}(u))|^2 du\Big]$$
$$\leq K_1 \int_t^T \mathbb{E}[|S^{t,s}(u) - S^{t,s'}(u)|^2]du$$

where in the last inequality we used the Lipschitz-continuity of f. Then estimate (2.25) proves global Lipschitz-continuity in s. Next, let t, s and ν be fixed, and $c < c' \leq \underline{C}$. In this case, an application of Hölder's inequality yields

$$|R(t,s,c,\nu) - R(t,s,c',\nu)| \leq \mathbb{E}\Big[\int_t^T |\nu(u)f(S^{t,s}(u))(H(C_\nu^{t,c}(u)) - H(C_\nu^{t,c'}(u)))|du\Big]$$
$$\leq \mathbb{E}\Big[\int_t^T |f(S^{t,s}(u))|^2 du\Big]\mathbb{E}\Big[\int_t^T |\nu(u)(H(C_\nu^{t,c}(u)) - H(C_\nu^{t,c'}(u)))|^2 du\Big]$$
$$\leq K_1(1+|s|^2)\mathbb{E}\Big[\int_t^T |\nu(u)(H(C_\nu^{t,c}(u)) - H(C_\nu^{t,s'}(u)))|^2 du\Big]$$

where we again used the growth condition on f and moment estimate for S. By noting that

$$|\nu(u)(H(C_\nu^{t,c}(u)) - H(C_\nu^{t,c'}(u)))|^2 \leq \overline{C}\nu(u)(H(C_\nu^{t,c}(u)) - H(C_\nu^{t,c'}(u))),$$

we then have

$$|R(t,x,c,\nu) - R(t,x,c',\nu)|^2 \leq \overline{C}K_1(1+|s|^2)\mathbb{E}\int_t^T \nu(u)(H(C_\nu^{t,s}(u)) - H(C_\nu^{t,c'}(u)))du]$$
$$\leq \overline{C}K_1(1+|s|^2)\mathbb{E}[C_\nu^{t,c}(T) - C_\nu^{t,c'}(T))]$$
$$= K(1+|s|^2)|c-c'|$$

with $K = K_1\overline{C}$. It remains to prove that for fixed initial data (t,s,c), the map $\nu \mapsto R(t,s,c,\nu)$ is continuous, uniformly on compacts. For $\nu_1, \nu_2 \in \mathcal{V}$, we estimate as before that

$$|R(t,s,c,\nu_1) - R(t,s,c,\nu_2)| \leq \mathbb{E}\Big[\int_t^T |H(C_{\nu_1}^{t,c}(t))\nu_1(t) - H(C_{\nu_2}^{t,c}(t))\nu_2(t)||f(S^{t,s}(u))|du\Big]$$
$$\leq K(1+|s|^2)\mathbb{E}\Big[\int_t^T |H(C_{\nu_1}^{t,c}(u))\nu_1(u) - H(C_{\nu_2}^{t,c}(u))\nu_2(u)|^2 du\Big].$$

By the equivalence of the norms $\|\cdot\|^2_{L^1([0,T]\times\Omega)}$, $\|\cdot\|^2_{L^2([0,T]\times\Omega)}$ we have

$$H(C_{\nu_1}^{t,c}(u)) \to H(C_{\nu_2}^{t,c}(u))$$

in L^1 as $\nu_1 \to \nu_2$ in $L^2([0,T]\times\Omega)$, for all $t \in [0,T]$. Now the claim follows since

$$\begin{aligned}&\mathbb{E}[\int_t^T |H(C^{t,c}_{\nu_1}(u))\nu_1(u) - H(C^{t,c}_{\nu_2}(u))\nu_2(s)|^2 ds]\\ &\leq \|H(C^{t,c}_{\nu_1}(u)) - H(C^{t,c}_{\nu_2}(u))\|^2_{L^2([0,T]\times\Omega)} + \|\nu_1-\nu_2\|^2_{L^2([0,T]\times\Omega)}\\ &\leq K\|H(C^{t,c}_{\nu_1}(u)) - H(C^{t,c}_{\nu_2}(u))\|^2_{L^1([0,T]\times\Omega)} + \|\nu_1-\nu_2\|^2_{L^2([0,T]\times\Omega)},\end{aligned}$$

where we again have used the equivalence of norms. □

Note that by the lemma above, we also have that the value function v is measurable. After proving the dynamic programming principle, where we need the measurability of R and v, we will see that v is in fact continuous. At this stage we note that by the estimates in the proof, we also have continuity in c:

$$|V(t,s,c) - V(t,s,c')|^2 \leq \sup_{\nu\in\mathcal{V}} |R(t,s,c,\nu) - R(t,s,c',\nu)|^2 \leq K(1+|s|^2)|c-c'|. \tag{2.29}$$

In what follows, we denote by

$$\mathcal{G}^\epsilon(t,s,c) := \{\nu \in \mathcal{V} : R(t,x,c,\nu) \geq v(t,s,c) - \epsilon\}, \qquad \epsilon > 0$$

the set of ϵ-optimal controls for the initial data $(t,s,c) \in \overline{\mathcal{O}_T}$. The next lemma verifies that the selection of ϵ-optimal controls at random initial values can be done in a way that maintains measurability. The proof follows very closely the proof of Lemma 3.1 by Soner and Touzi in [25].

2.30. **Theorem** (Measurable selection). *For any probability measure μ defined on $\overline{\mathcal{O}_T}$, and for each $\epsilon > 0$, there exists a Borel-measurable function $\phi^\epsilon_\mu : (\overline{\mathcal{O}_T}, \mathcal{B}_{\overline{\mathcal{O}_T}}) \mapsto (\mathcal{V}, \mathcal{B}_\mathcal{V})$ such that*

$$\phi^\epsilon_\mu(t,s,c) \in \mathcal{G}^\epsilon(t,s,c) \quad \text{for } \mu - a.e.\ (t,s,c) \in \overline{\mathcal{O}_T}$$

Proof. Clearly, $\overline{\mathcal{O}_T}$ and $\mathcal{V}$ are Borel spaces. Set

$$B^\epsilon := \{(t,s,c,\nu) \in \overline{\mathcal{O}_T} \times \mathcal{V} : \nu \in \mathcal{G}^\epsilon(t,s,c)\}.$$

1. Clearly B^ϵ is a Borel subset of $\overline{\mathcal{O}_T} \times \mathcal{V}$, since

$$B^\epsilon = \{(t,s,c,\nu) \in \overline{\mathcal{O}_T} \times \mathcal{V} : R(t,s,c,\nu) + \epsilon \geq V(t,s,c)\}$$

and the functions R and V are measurable.

2. Borel sets are analytic, so the Jankov-von Neumann theorem implies the existence of an analytically measurable function $\phi : \mathcal{O}_T \to \mathcal{V}$ such that $Gr(\phi) \subset B^\epsilon$, i.e. $\phi(t,s,c)$ is an admissible control in $\mathcal{G}^\epsilon(t,s,c)$ for all (t,s,c).

3. The final step is to construct a Borel measurable map ϕ^ϵ_μ that agrees μ-a.e. with the analytical function ϕ from step 2. However, this step is word for word the same as in step 3. of the proof of Lemma 3.1 by Soner and Touzi in [25], and is therefore omitted. □

Now, we are finally set to prove the dynamic programming principle.

Proof for dynamic programming. Let $\theta \in \mathcal{T}_{t,T}$. By the definition of v and the tower property of conditional expectations

$$\begin{aligned} v(t,s,c) &= \sup_{\nu\in\mathcal{V}} \mathbb{E}\Big[\int_t^{T\wedge\tau_{\overline{C}}} \nu(u) f(S^{t,s}(u))du\Big]\\ &= \sup_{\nu\in\mathcal{V}} \mathbb{E}\Big[\int_t^\theta \nu(u) f(S^{t,s}(u))du + \int_\theta^{T\wedge\tau_{\overline{C}}} \nu(u) f(S^{t,s}(u))du\Big]\\ &= \sup_{\nu\in\mathcal{V}} \mathbb{E}\Big[\int_t^\theta \nu(u) f(S^{t,s}(u))du + \mathbb{E}[\int_\theta^{T\wedge\tau_{\overline{C}}} \nu(u) f(S^{t,s}(u))du|\zeta]\Big]. \qquad (2.31)\end{aligned}$$

where ζ is as in Lemma 2.15. By the concatenation property of the control class $\mathcal{V}$, $\nu 1_{[\theta,T]} \in \mathcal{V}$ so

$$\mathbb{E}\Big[\int_\theta^{T\wedge\tau_{\overline{C}}} \nu(u) f(S^{t,s}(u))du|\zeta\big] \leq v(\theta, S^\nu(\theta), C^{t,c}(\theta))$$

by Lemma 2.15. Substituting the above inequality into (2.31) we obtain (DP1).

Let us fix arbitrary $\epsilon > 0$ and $\nu \in \mathcal{V}$. Now, let μ be a probability measure on $\overline{\mathcal{O}_T}$ induced by $(\theta, S^{t,s}(\theta), C_\nu^{t,c}(\theta))$, and let ϕ_μ^ϵ be the Borel-measurable map given by Theorem 2.30. By stability under measurable selection, there exists an admissible control $\nu_1 \in \mathcal{V}$ such that, for μ - almost every (s, y),

$$P^{\phi_\mu^\epsilon(s,y)} = P^{\nu_1} \text{ on } [\theta, T].$$

Then, since the class of admissible controls is also stable under concatenation, we have that the process defined by

$$\nu^\epsilon(t) := \nu(t) 1_{[0,\theta)}(t) + \nu_1(t) 1_{[\theta,T]}(t)$$

is an admissible control which is ϵ-optimal at time θ, i.e.

$$R(\theta, S^{t,s}(\theta), C^{t,c}(\theta), \nu) \geq v(\theta, S^{t,s}(\theta), C^{t,c}(\theta)) - \epsilon.$$

Then, again using the tower property and similar considerations as in proving (DP1), we obtain

$$\begin{aligned} v(t,s,c) &\geq \mathbb{E}\Big[\int_t^{T\wedge\tau_{\overline{C}}} \nu(u) f(S^{t,s}(u))du|\zeta\Big] \\ &= \mathbb{E}\Big[\int_t^\theta \nu(u) f(S^{t,s}(u))du + \mathbb{E}[\int_\theta^{T\wedge\tau_{\overline{C}}} \nu(u) f(S^{t,s}(u))du|\zeta]\Big] \\ &\geq \mathbb{E}\Big[\int_t^\theta \nu(u) f(S^{t,s}(u))du + v(\theta, S^{t,s}(\theta), C^{t,c}(\theta))\Big] - \epsilon. \end{aligned}$$

Since this holds for every ϵ, the proof is complete. □

Now, the dynamic programming principle can be applied to prove continuity of the value function in time in a rather standard fashion.

2.32. **Proposition.** *For all $t, t' \in [0, T)$, $s \in \mathbb{R}^2$ and $c \in [0, \overline{C}]$,*

$$|v(t,s,c) - v(t',s,c)| \leq C(1 + |s|)|t - t'|^{\frac{1}{2}}. \tag{2.33}$$

Proof. Let $0 \leq t < t' \leq T$. Applying the dynamic programming principle (2.13) with $\theta = t'$ such that we still have $C_\nu^{t,c}(t') < \overline{C}$, we have

$$\begin{aligned} |v(t,s,c) - v(t',s,c)| &= \Big|\sup_{\nu\in\mathcal{V}} \mathbb{E}\Big[\int_t^{t'} \nu(u) f(S^{t,s}(u))du + v(t', S^{t,s}(t'))\Big] - v(t', s)\Big| \\ &\leq \sup_{\nu\in\mathcal{V}} \mathbb{E}\Big[\int_t^{t'} |\nu(u) f(S^{t,s}(u))|du + |v(t', S(t'), C_\nu^{t,c}(t')) - v(t', s, c)|\Big]. \end{aligned}$$

By the growth condition (2.7) and Lipschitz continuity of V with respect to s and c,

$$|v(t,s,c) - v(t',s,c)| \leq \int_t^s (1 + \mathbb{E}[|S^{t,s}(u)|])du + K \sup_{\nu\in\mathcal{V}} \mathbb{E}[|S^{t,s}(t') - s| + |C_\nu^{t,c}(t') - c|].$$

We conclude with the estimates (2.21), (2.24) and noticing that

$$|C_\nu^{t,c}(t') - c| \leq \overline{C}|t - t'|.$$

□

Combining the continuity estimates for v we have

2.34. **Corollary.** *The value function is continuous. In fact,*

$$|v(t,s,c) - v(t',s',c')| \le C\Big[(1+|s|)(|t-t'|^{\frac{1}{2}} + |c-c'|^{1/2}) + |s-s'|\Big]. \tag{2.35}$$

3. Hamilton-Jacobi-Bellman equation and viscosity solutions

If we assume that the value is sufficiently smooth, it is classical to use dynamic programming, Itô's formula and martingale property of the stochastic integrals to derive that the value function v should satisfy

$$\partial_t v(t,s,c) + G(t,s,D_s v(t,s), D_s^2 v(t,s)) + B(t,s,c,v) + \sup_{u\in U}\{u(\partial_c v(t,s,c) + f(s))\} = 0, \tag{3.1}$$

where

$$G(t,s,D_s v(t,s,c), D_s^2 v(t,s,c)) = b(t,s)\cdot Dv(t,s,c) + tr\Big(\frac{1}{2}\sigma\sigma^T(t,s)D_s^2 v(t,s,c))\Big) \tag{3.2}$$

$$= \sum_{i=1}^{2} b_i(t,s)\partial_{s_i} v(t,s,c) + \sum_{i,j=1}^{2} a_{ij}(t,s)\partial^2_{s_i s_j} v(t,s,c) \tag{3.3}$$

where $a_{ij} = (\frac{1}{2}\sigma\sigma^T)_{ij}$ and

$$B(t,s,c,v) = \int_{\mathbb{R}^2} [v(t,s+\gamma(t,s),c) - v(t,s,c) - \gamma(t,s,z)\cdot D_s v(t,s,c)]m_t(dz) \tag{3.4}$$

$$= \sum_{i=1}^{2}\int_{\mathbb{R}} [v(t,s+\gamma^{(i)}(t,s),c) - v(t,s,c) - \gamma^{(i)}(t,s,z_i)\cdot D_s v(t,s,c)\ ]m_t^{(i)}(dz_i). \tag{3.5}$$

In addition, we have the terminal and boundary values

$$v(T,s,c) = 0, \qquad v(t,s,\overline{C}) = 0.$$

We combine these in the statement

$$v|_{\partial\mathcal{O}_T} = 0. \tag{3.6}$$

However, it is known that value functions of control problems are not generally smooth enough to allow to interpret the above equation in the classical sense. The concept of viscosity solutions has become the standard tool for handling theory for Hamilton-Jacobi Bellman equations of control problems, deterministic and stochastic alike. Besides considering solutions of the equation, it is standard to consider more generally sub- and supersolutions of the equation, and demand only semicontinuity from these. This more general framework is extremely useful when proving convergence of various numerical solution methods for the equation. In the following, we give the rigorous definition for the tolling agreement.

We recall that, in general for every locally bounded function $h : [0,T]\times\mathbb{R}^n \to \mathbb{R}$, its *upper* and *lower semicontinuous envelopes*, denoted by h^* and h_* respectively, are defined as

$$h^*(t,x) := \limsup_{(s,y)\to(t,x)} h(s,y), \quad h_*(t,x) := \liminf_{(s,y)\to(t,x)} h(s,y).$$

A locally bounded function $h : [0,T]\times\mathbb{R}^n \to \mathbb{R}$ is said to be *upper semicontinuous* if $h^* \le h$ and *lower semicontinuous* if $h_* \ge h$. A function h is lower semicontinuous if and only if the set

$$\{(t,x) : h(t,x) > a\}$$

is open for all $a\in\mathbb{R}$, and upper semicontinuous if and only if

$$\{(t,x) : h(t,x) < a\}$$

is open for all $a\in\mathbb{R}$. If h is both upper and lower semicontinuous then it is continuous. We denote the sets of upper and lower semicontinuous functions by $USC([0,T]\times\mathbb{R}^n)$ and $LSC([0,T]\times\mathbb{R}^n)$, respectively.

In addition we denote by $USC_1([0,T]\times\mathbb{R}^n)$ ($LSC_1([0,T]\times\mathbb{R}^n)$) the class of non-negative functions h belonging to $USC([0,T]\times\mathbb{R}^n)$ ($LSC([0,T]\times\mathbb{R}^n)$) and satisfying

$$h(t,x)\le K(1+|x|)\qquad t\in[0,T]$$

for some $K>0$.

3.7. **Definition.** (i) A function $v\in USC_1(\overline{\mathcal{O}_T})$ is a *viscosity subsolution* of (3.1) if and only if

$$v\le 0 \text{ on } \partial\mathcal{O}_T$$

and for all $\phi\in C_1^{1,2}(\mathcal{O}_T)$ such that $v\le\phi$ everywhere we have:

$$\partial_t v(t,s,c)+G(t,s,D_s v(t,s),D_s^2 v(t,s))+B(t,s,c,v)+\sup_{u\in U}\{u(\partial_c v(t,s,c)+f(s))\}\ge 0 \tag{3.8}$$

whenever $\phi(t,s,c)=v(t,s,c)$.

(ii) A function $v\in LSC_1(\overline{\mathcal{O}_T})$ is a *viscosity supersolution* of (3.1) if and only if

$$v\ge 0 \text{ on } \partial\mathcal{O}_T$$

and for all $\phi\in C_1^{1,2}(\mathcal{O}_T)$ such that $v\ge\phi$ everywhere we have:

$$\partial_t v(t,s,c)+G(t,s,D_s v(t,s),D_s^2 v(t,s))+B(t,s,c,v)+\sup_{u\in U}\{u(\partial_c v(t,s,c)+f(s))\}\le 0. \tag{3.9}$$

whenever $v(t,s,c)=\phi(t,s,c)$.

(iii) A function $v\in C_1(\overline{\mathcal{O}_T})$ is a *viscosity solution* of (3.1) if and only if it is simultaneously a sub- and supersolution of (3.1).

We now prove by combining the above definitions with the dynamic programming principle that the definition is sensible in that it includes the value function.

3.10. **Theorem.** *The value function is a viscosity solution of (3.1).*

Proof. Continuity of the value function was already given in Corollary 2.34, and the growth condition follows directly from estimate (2.21) in Lemma 2.20. Also, that V satisfies the terminal and boundary conditions follows directly from the definition. Now, let $(t,s,c)\in\overline{\mathcal{O}_T}$ and $\phi\in C^{1,2}(\mathcal{O}_T)$ be such that

$$0=(v-\phi)(t,s,c)=\min_{(t,s,c)\in\mathcal{O}_T}(v-\phi)(t,s,c).$$

For $n\in\mathbb{N}$, let θ_n denote the exit time of $(r,S^{t,s}(r),C_\nu^{t,c}(r))_{r\in[0,T]}$ from a ball with radius $1/n$ and center at (t,s,c). From the dynamic programming principle (2.13), we have that

$$\begin{aligned}\phi(t,s,c)=v(t,s,c)&=\sup_{\nu\in\mathcal{V}}\mathbb{E}\Big[\int_t^{\theta_n}\nu(r)f(S^{t,s}(r))dr+v(\theta_n,S^{t,s}(\theta_n),C_\nu^{t,c}(\theta_n))\Big]\\&\ge\sup_{\nu\in\mathcal{V}}\mathbb{E}\Big[\int_t^{\theta_n}\nu(u)f(S^{t,s}(r))dr+\phi(\theta_n,S^{t,s}(\theta_n),C_\nu^{t,c}(\theta_n))\Big].\end{aligned}$$

Applying Itô's formula along with the martingale property of the stochastic integrals yields

$$\begin{aligned}0&\ge\sup_{\nu\in\mathcal{V}}\Big(\mathbb{E}\Big[\int_t^{\theta_n}\nu(r)f(S^{t,s}(r))dr+\phi(\theta_n,S^{t,s}(\theta_n),C_\nu^{t,c}(\theta_n))\Big]-\phi(t,s,c)\Big)\\&=\sup_{\nu\in\mathcal{V}}\mathbb{E}\Big[\int_t^{\theta_n}\nu(r)f(S^{t,s}(r))+\partial_r\phi(r,S^{t,s}(r),C_\nu^{t,c}(r))\\&+G(r,S^{t,s}(r),D_s\phi(r,S^{t,s}(r),C_\nu^{t,c}(r)),D_s^2\phi(r,S^{t,s}(r),C_\nu^{t,c}(r)))\\&+B(r,S^{t,s}(r),C_\nu^{t,c}(r),\phi)+\partial_c\phi(r,S^{t,s}(r),C_\nu^{t,c}(r))\Big]\end{aligned}$$

Dividing by $\mathbb{E}[\theta_n]$, setting $\nu \equiv u$ and sending $n \to \infty$, we have

$$\partial_t v(t,s,c) + G(t,s,D_s v(t,s), D_s^2 v(t,s)) + B(t,s,v) + u(\partial_c v(t,s,c) + f(s)) \le 0.$$

Since the inequality holds for arbitrary $u \in U$, this proves the supersolution property of the value function. The proof of the subsolution property follows similar ideas using inequality (2.12), and is left to the reader. □

4. Comparison

As was mentioned in the introduction, integral terms with possibly unbounded measures resulting from the jumps of a process have proven difficult to handle in this framework. We begin by giving the central definitions, and have a closer look on integrability issues in the semijet formulation for viscosity solutions. After briefly describing the fundamental problems, we describe carefully how they may be overcome.

In this section we treat the variables s and c simultaneously, and denote $x \in \mathbb{R}^2 \times [0, \overline{C}]$, $x = (s,c)$. We then work with the functions

$$V(t,x) := v(t,s,c).$$

Let us first recall the notion of parabolic semijets. In general, for a domain $D \subset \mathbb{R}^n$ and a function $V \in USC([0,T] \times D)$ (resp. $LSC([0,T]\times D)$) and $(t,x) \in [0,T] \times D$, we define the parabolic superjet (resp. subjet):

$$\begin{aligned}\mathcal{P}^{2,+(-)}V(t,x) := \quad & \Big\{(q,p,P) \in \mathbb{R} \times D \times \mathcal{S}_n | V(u,y) \le (\ge) V(t,x) + q(s-t) \\ & + p \cdot (y-x) + \frac{1}{2} P(y-x)\cdot(y-x) \\ & + o(|s-t| + |y-x|^2) \text{ as } (s,y) \to (t,x)\Big\}\end{aligned}$$

and its closure:

$$\begin{aligned}\overline{\mathcal{P}}^{2,+(-)}V(t,x) := \quad & \Big\{(q,p,P) = \lim_{n\to\infty}(q^{(n)}, p^{(n)}, P^{(n)}) \\ & \text{with } (q^{(n)}, p^{(n)}, P^{(n)}) \in \mathcal{P}^{2,+(-)}V(t_n,x_n) \\ & \textit{and } \lim_{n\to\infty}(t_n, x_n, V(t_n,x_n)) = (t,x,V(t,x))\Big\}.\end{aligned}$$

To apply results concerning semijets in our integro-differential setting, we denote

$$G(t,x,p,P) = b(t,x)\cdot p + tr\Big(\frac{1}{2}\sigma\sigma^T(t,x)P\Big)$$

and the integral now by

$$B(t,x,V,p) = \int_{\mathbb{R}^2} [V(t,x+\gamma(t,x,z)) - V(t,x) - \gamma(t,x,z)\cdot D_x V(t,x)] m_t(dz)$$

With slight abuse of notation, we have above identified $b(t,x) = b(t,s)$, $\sigma(t,x) = \sigma(t,s)$, $\gamma(t,x,z) = \gamma(t,s,z)$ and $f(x) = f(s)$.

The first observation is that the integral $B(t,x,V,p)$ is well defined when $V \in USC(\mathcal{O}_T)$ and $(q,p,P) \in \mathcal{P}^{2,+}V(t,x)$. Indeed, in this case we have (by the choice $s = t$, $y = x + \gamma(t,x,z)$ in the definition of the superjet)

$$V(t,x+\gamma(t,x,z)) - V(t,x) + p\cdot\gamma(t,x,z) \le \frac{1}{2}P\gamma(t,x,z)\cdot\gamma(t,x,z) + o(|\gamma(t,x,z)|^2).$$

Here the right-hand side is integrable by assumptions, so the positive part of the integrand is integrable with a finite integral value. Thus the integral is well defined in the usual measure theoretic sense

$$\int f(z)m(dz) = \int f^+(z)m(dz) - \int f^-(z)m(dz)$$

where f^+ and f^- are the positive and negative parts of a given function f, respectively, and m is a measure. A priori, we could have

$$\int f(z)m(dz) = -\infty.$$

This does not really matter for the purposes of this article, but it is interesting to note that the supersolution property actually implies a finite value for the integral B whenever $(q,p,P) \in \mathcal{P}^{2,+}V(t,x)$. For this, pick $\bar{\phi} \in C^{1,2}(\mathcal{O}_T)$ such that $\bar{\phi} \geq v$, $u(t,x) = \bar{\phi}(t,x)$, $\partial_t V(t,x) = \partial_t \bar{\phi}(t,x)$, $\partial_x V(t,x) = \partial_x \bar{\phi}(t,x)$, and $\partial_x^2 V(t,x) = \partial_x^2 \bar{\phi}(t,x)$. This can be done by the construction of Evans, see [28, Proposition 4.5.4]. Moreover, let $\{V_k\}_{k=1}^{\infty} \subset C_1^{\infty}(\mathcal{O}_T)$ such that $V_k \downarrow v$ almost everywhere as $k \to \infty$. Let $\mathcal{X}_k$ be a smooth function such that $0 \leq \mathcal{X}_k \leq 1$, $\mathcal{X}_k = 1$ in a ball with radius $1/2k$ and center at (t,x), and $\mathcal{X}_k = 0$ outside a ball with radius $1/k$ and center at (t,x). Then

$$\phi_k(r,y) := \mathcal{X}_k(y)\bar{\phi}(r,y) + (1 - \mathcal{X}_k(y))u_k(r,y)$$

defines a sequence of test functions such that

$$\phi_k(t,x) = V(t,x), \quad \partial_t \phi_k(t,x) = q, \quad \partial_x \phi_k(t,x) = p, \quad \partial_x \phi_k(t,x) = P$$

and $\phi^k \downarrow v$ everywhere as $k \to \infty$. The supersolution property implies

$$B(t,x,\phi_k,p) \geq q + G(t,x,p,P) + \sup_{u \in U}\{u(p_3 + f(x))\}$$

for all k. On the one hand, the right hand side here is constant and the left hand side is a decreasing sequence, so the sequence on the left hand has a finite limit. On the other hand, by monotone convergence

$$\lim_{k \to \infty} B(t,x,\phi_k) = B(t,x,V).$$

Thus the integral has a well defined, finite value and we have moreover proved the following

4.1. **Lemma.** *Let $\underline{V} \in USC_1(\overline{\mathcal{O}_T})$ be a viscosity subsolution, $\overline{V} \in LSC_1(\overline{\mathcal{O}_T})$ supersolution of (2.11). Then*

$$q + G(t,x,p,P) + B(t,x,\underline{V},p) + \sup_{u \in U}\{u(p_3 + f(x))\} \geq 0, \tag{4.2}$$

for $(q,p,P) \in \mathcal{P}^{2,+}\underline{V}(t,x)$ and

$$q + G(t,x,p,P) + B(t,x,\overline{V},p) + \sup_{u \in U}\{u(p_3 + f(x))\} \leq 0. \tag{4.3}$$

for $(q,p,P) \in \mathcal{P}^{2,-}\overline{V}(t,x)$.

Let us now explain some of the problems one may run into when trying to apply directly the methods from theory for second order (pure) differential equations, see also [15]. In the classical setting, one can easily verify that sub- and supersolutions satisfy inequalities of the form (4.2), (4.3) on the *closures* of the semijets. Then one may apply the maximum principle for semicontinuous functions (see the Appendix or [9]) to prove the comparison principle. In the integrodifferential setting, we can not directly take the limit to a point that is in the closure but possibly not in the semijet as it is not clear whether the integrals remain well defined. The early works on the topic use the fact that, for every point in a semijet there is a smooth test function, and separate the integral term into two parts B_ϵ and B^ϵ where B_ϵ is an integral operator integrating the test function in a small neighborhood of $z = 0$, and B^ϵ integrates the semicontinuous function v away from zero. It is then suggested that at the limit we regain a well defined inequality with B replaced by the sum $B_\epsilon + B^\epsilon$. The point that is missed in this argumentation is that, for every n the test function corresponding to the points (t_n, x_n) is different, so one actually has a sequence of test functions which may or may not converge to an integrable limit. Our point here is that (i) one does not need test functions but should rather use the sub- or supersolutions in the integrals, and (ii) the estimates typically made for the integrals B^ϵ is what insures there is a finite value for the integral differences exactly at the points where this is needed in the proof of comparison.

It is important to notice that, at this stage, we do not claim the integrals in our lemma above are well defined on the *closures* of the semijets, since this is not in general the case. In our proof of the comparison principle we do need the full closure, but some rather specific points of it. It turns out that we can make estimates that insure integrability at the limit.

4.4. **Lemma.** *Let $\underline{V} \in USC_1(\overline{\mathcal{O}_T})$ be a subsolution, and $\overline{V} \in LSC_1(\overline{\mathcal{O}_T})$ a supersolution. Let $(\hat{t}, \hat{x}, \hat{y})$, (b, p, P) and (c, q, Q) be such that*

$$(b, p, P) \in \overline{\mathcal{P}}^{2,+}\underline{V}(\hat{t}, \hat{x}),$$
$$(c, q, Q) \in \overline{\mathcal{P}}^{2,-}\overline{V}(\hat{t}, \hat{y}).$$

Define

$$\begin{aligned} f(t, x, y, p, q, z) :=&\underline{V}(t, x + \gamma(t, x, z)) - \underline{V}(t, x) - p \cdot \gamma(t, x, z) \\ &- (\overline{V}(t, y + \gamma(t, y, z)) - \overline{V}(t, y) - q \cdot \gamma(t, y, z)), \end{aligned}$$

and suppose that there is a continuous function $h = h(t, x, y, z)$ such that

$$f(\hat{t}, \hat{x}, \hat{y}, p, q, z) \leq h(\hat{t}, \hat{x}, \hat{y}, z) \quad \text{and} \quad \int_{\mathbb{R}^2} h(\hat{t}, \hat{x}, \hat{y}, z) m_{\hat{t}}(dz) < \infty.$$

Then

$$\begin{aligned} 0 \leq& b - c + G(\hat{t}, \hat{x}, p, P) - G(\hat{t}, \hat{y}, q, Q) + \sup_{u \in U}\{u(p_3 + f(\hat{x}))\} - \sup_{u \in U}\{u(q_3 + f(\hat{y}))\} \\ &+ \int_{\mathbb{R}^2} h(\hat{t}, \hat{x}, \hat{y}, z) m_{\hat{t}}(dz). \end{aligned}$$

Proof. By assumptions there exists sequences $(\hat{t}^{(n)}, \hat{x}^{(n)}, \hat{y}^{(n)})$ and $(b^{(n)}, p^{(n)}, P^{(n)})$, $(c^{(n)}, q^{(n)}, Q^{(n)})$ satisfying $(b^{(n)}, p^{(n)}, P^{(n)}) \in \mathcal{P}^{2,+}\underline{V}(\hat{t}^{(n)}, \hat{x}^{(n)})$, $(b^{(n)}, q^{(n)}, Q^{(n)}) \in \mathcal{P}^{2,-}\overline{V}(\hat{t}^{(n)}, \hat{y}^{(n)})$ such that

$$(\hat{t}^{(n)}, \hat{x}^{(n)}, \hat{y}^{(n)}) \to (\hat{t}, \hat{x}, \hat{y})$$

and

$$\lim_{n\to\infty} (b^{(n)}, p^{(n)}, P^{(n)}) = (b, p, P), \qquad \lim_{n\to\infty} (c^{(n)}, q^{(n)}, Q^{(n)}) = (c, q, Q).$$

By Lemma 4.1, we have

$$\begin{aligned} 0 \leq& b^{(n)} - c^{(n)} + G(\hat{t}^{(n)}, \hat{x}^{(n)}, p^{(n)}, P^{(n)}) - G(\hat{t}, \hat{y}, q^{(n)}, Q^{(n)}) \\ &+ \sup_{a \in A}\{a(p_3^{(n)} + \hat{x}_1 - \hat{x}_2)\} - \sup_{a \in A}\{a(q_3^{(n)} + \hat{y}_1 - \hat{y}_2)\} \\ &+ B(\hat{t}^{(n)}, \hat{x}^{(n)}, \underline{V}, p^{(n)}) - B(\hat{t}^{(n)}, \hat{y}^{(n)}, \overline{V}, q^{(n)}). \end{aligned}$$

Here

$$B(\hat{t}^{(n)}, \hat{x}^{(n)}, \underline{V}, p^{(n)}) - B(\hat{t}^{(n)}, \hat{y}^{(n)}, \overline{V}, q^{(n)}) = \int_{\mathbb{R}} f(\hat{t}^{(n)}, \hat{x}^{(n)}, \hat{y}^{(n)}, p^{(n)}, q^{(n)}, z) m(dz).$$

By the assumptions and upper semicontinuity of f,

$$\limsup_{n\to\infty} f(\hat{t}^{(n)}, \hat{x}^{(n)}, \hat{y}^{(n)}, p^{(n)}, q^{(n)}, z) \leq h(\hat{t}, \hat{x}, \hat{y}, z).$$

Now the claim follows by taking $n \to \infty$ and using (generalized) Fatou's lemma. □

The following theorem states that, not only is the solution of the HJB equation unique, but sub- and supersolutions also satisfy a comparison principle. As stated in the introduction, this stronger result is useful for the study of convergence results (among other things). The proof here is actually quite similar to the pure diffusion case, after noting there is a suitable function h so we can apply the above lemma.

4.5. **Theorem** (**Strong comparison principle**). *Suppose* $\overline{V} \in LSC_1(\overline{\mathcal{O}_T})$ *is a viscosity supersolution of (3.1) and* $\underline{V} \in USC_1(\overline{\mathcal{O}_T})$ *is a viscosity subsolution of (3.1), and suppose that*

$$\underline{V} \leq \overline{V}, \quad \text{on } \partial\mathcal{O}_T$$

Then

$$\underline{V} \leq \overline{V} \text{ on } \overline{\mathcal{O}_T}. \tag{4.6}$$

Proof. : Let $\mu > 0$, define $\overline{V}^\mu(t,x) := \overline{V}(t,x) + \mu(T-t)$. First, it is straightforward to verify that if $\overline{V}$ is a supersolution, then $\overline{V}^\mu$ is a supersolution of

$$\partial_t V(t,x) + G(t,x,D_xV(t,x),D_x^2V(t,x)) + B(t,x,V) + \sup_{u\in U}\{u(\partial_{x_3}v(t,x,c)+f(x))\} = -\mu.$$

We prove the claim holds for $\overline{V}^\mu$, and the main claim follows by taking $\mu \to 0$. Suppose there exists $\delta > 0$ and $(\bar{t},\overline{x}) \in [0,T]\times\mathbb{R}_+$ such that

$$\underline{V}(\bar{t},\overline{x}) \geq \overline{V}^\mu(\bar{t},\overline{x}) + 2\delta. \tag{4.7}$$

As in Lemma 5.7, we define

$$\Phi(t,x,y) := \underline{V}(t,x) - \overline{V}^\mu(t,y) - \psi(t,x,y), \quad (t,x,y) \in [0,T]\times(\mathbb{R}^2\times[0,\overline{C}])^2$$

where

$$\psi(t,x,y) = \frac{\alpha}{2}|x-y|^2 + \frac{\epsilon}{2}e^{\lambda(T-t)}(|x|^2+|y|^2).$$

Let $(t_\alpha,x_\alpha,y_\alpha)$ be the corresponding sequence of maximum points of Φ that satisfy (5.8),(5.9). Here we have dropped the dependency on ϵ for notational convenience.
Observe that

$$\Phi(t_\alpha,x_\alpha,y_\alpha) \geq \underline{V}(\bar{t},\overline{x}) - \overline{V}^\mu(\bar{t},\overline{x}) - \epsilon e^{\lambda(T-\bar{t})}|\overline{x}|^2 \geq \delta > 0, \tag{4.8}$$

for any $\alpha > 1$ when $\epsilon > 0$ is small enough. Note that this implies

$$\underline{V}(t_\alpha,x_\alpha) \geq \overline{V}^\mu(t_\alpha,y_\alpha) + \delta \tag{4.9}$$

for any $\alpha > 1$, and any ϵ sufficiently small.

Let us now look at the special case $(t_\epsilon,x_\epsilon) \in \partial\mathcal{O}_T$. Note that

$$\underline{V}(\bar{t},\overline{x}) - \overline{V}^\mu(\bar{t},\overline{x}) - \epsilon e^{\lambda(T-\bar{t})}|\overline{x}|^2 \leq \Phi(t_\alpha,x_\alpha,y_\alpha) \leq \underline{V}(t_\alpha,x_\alpha) - \overline{V}^\mu(t_\alpha,y_\alpha).$$

By the upper semicontinuity of $\underline{V}$, $-\overline{V}^\mu$ and since $\underline{V}|_{\partial\mathcal{O}_T} \leq \overline{V}^\mu_{\partial\mathcal{O}_T}$ on $[0,\infty)$, we can send $\alpha\uparrow\infty$ and then $\epsilon\downarrow 0$ in this inequality to obtain $\underline{V}(\bar{t},\overline{x}) - \overline{V}^\mu(\bar{t},\overline{x}) \leq 0$, which contradicts (4.7).

We now assume that $t_\epsilon < T$, so that $t_\alpha < T$ for any α sufficiently large. Applying the maximum principle for semicontinuous functions (see the Appendix and the references therein) to the function Φ at $(t_\alpha,x_\alpha,y_\alpha)$ yields the existence of numbers $b_\alpha, c_\alpha \in \mathbb{R}$ and matrices $P_\alpha, Q_\alpha \in \mathcal{S}^2$ such that

$$(b_\alpha, D_x\psi(t_\alpha,x_\alpha,y_\alpha), P_\alpha) \in \overline{\mathcal{P}}^{2,+}u(t_\alpha,x_\alpha)$$
$$(c_\alpha, -D_y\psi(t_\alpha,x_\alpha,y_\alpha), Q_\alpha) \in \overline{\mathcal{P}}^{2,-}u(t_\alpha,y_\alpha)$$

where $b_\alpha - c_\alpha = \partial_t\psi(t_\alpha,x_\alpha,y_\alpha)$ and

$$\begin{pmatrix} P_\alpha & 0 \\ 0 & -Q_\alpha \end{pmatrix} \leq D^2\psi(t_\alpha,x_\alpha,y_\alpha) + \frac{1}{\alpha}[D^2\psi(t_\alpha,x_\alpha,y_\alpha)]^2. \tag{4.10}$$

Next, we denote $p_\alpha = D_x\psi(t_\alpha, x_\alpha, y_\alpha)$, $q_\alpha = -D_y\psi(t_\alpha, x_\alpha, y_\alpha)$ and wish to apply Lemma 4.4 for $(b_\alpha, p_\alpha, P_\alpha) \in \overline{\mathcal{P}}^{2,+}\underline{V}(t_\alpha, x_\alpha)$, $(c_\alpha, q_\alpha, Q_\alpha) \in \overline{\mathcal{P}}^{2,-}\overline{V}(t_\alpha, y_\alpha)$ given above. First, we can directly calculate that here

$$f(t_\alpha, x_\alpha, y_\alpha, z) = \Phi(t_\alpha, x_\alpha + \gamma(t_\alpha, x_\alpha, z), y_\alpha + \gamma(t_\alpha, y_\alpha, z)) - \Phi(t_\alpha, x_\alpha, y_\alpha) \tag{4.11}$$

$$+ \frac{1}{2\epsilon}|\gamma(t_\alpha, x_\alpha, z) - \gamma(t_\alpha, y_\alpha, z)|^2 + \frac{\alpha}{2}[|\gamma(t_\alpha, x_\alpha, z)|^2 + |\gamma(t_\alpha, y_\alpha, z)|^2] \tag{4.12}$$

$$\leq \frac{\alpha}{2}|\gamma(t_\alpha, x_\alpha, z) - \gamma(t_\alpha, y_\alpha, z)|^2 + \frac{\epsilon}{2}e^{\lambda(T-t_\alpha)}[|\gamma(t_\alpha, x_\alpha, z)|^2 + |\gamma(t_\alpha, y_\alpha, z)|^2], \tag{4.13}$$

where we have used the maximality of $\Phi(t_\alpha, x_\alpha, y_\alpha)$ to deduce the last step. Noting that

$$h(t, x, y, z) := \frac{\alpha}{2}|\gamma(t, x, z) - \gamma(t, y, z)|^2 + \frac{\epsilon}{2}e^{\lambda(T-t_\alpha)}[|\gamma(t, x, z)|^2 + |\gamma(t, y, z)|^2]$$

defines an $m_t(\cdot)$-integrable function (for every (t, x, y) and α), we have that the assumptions of Lemma 4.4 hold. Thus, we have

$$\begin{aligned}
\mu \leq& \partial_t\psi(t_\alpha, x_\alpha, y_\alpha) + G(t_\alpha, x_\alpha, D_x\psi(t_\alpha, x_\alpha, y_\alpha), P_\alpha) - G(t_\alpha, y_\alpha, -D_y\psi(t_\alpha, x_\alpha, y_\alpha), Q_\alpha) \\
&+ \sup_{a\in A}\{a(\partial_{x_3}\psi(t_\alpha, x_\alpha, y_\alpha) + f(x_\alpha))\} - \sup_{a\in A}\{a(-\partial_{y_3}\psi(t_\alpha, x_\alpha, y_\alpha) + f(y_\alpha))\} \\
&+ \int_{\mathbb{R}} \{\frac{\alpha}{2}|\gamma(t_\alpha, x_\alpha, z) - \gamma(t_\alpha, y_\alpha, z)|^2 + \frac{\epsilon}{2}[|\gamma(t_\alpha, x_\alpha, z)|^2 + |\gamma(t_\alpha, y_\alpha, z)|^2]\}\eta(t_\alpha)m(dz) \\
=& -\lambda\epsilon e^{\lambda(T-t_\alpha)}(|x_\alpha|^2 + |y_\alpha|^2) + [b(t_\alpha, x_\alpha)(\alpha(x_\alpha - y_\alpha) + \epsilon e^{\lambda(T-t_\alpha)}x_\alpha) + b(t_\alpha, y_\alpha)(-\alpha(x_\alpha - y_\alpha) + \epsilon e^{\lambda(T-t_\alpha)}y_\alpha)] \\
&+ \frac{1}{2}[tr(\sigma\sigma^T(t_\alpha, x_\alpha)P_\alpha) - tr(\sigma\sigma^T(t_\alpha, y_\alpha)Q_\alpha)] \\
&+ \sup_{a\in A}\{a(\partial_{x_3}\psi(t_\alpha, x_\alpha, y_\alpha) + f(x_\alpha))\} - \sup_{a\in A}\{a(-\partial_{y_3}\psi(t_\alpha, x_\alpha, y_\alpha) + f(y_\alpha))\} \\
&+ \int_{\mathbb{R}} \frac{\alpha}{2}|\gamma(t_\alpha, x_\alpha, z) - \gamma(t_\alpha, y_\alpha, z)|^2\eta(t_\alpha)m(dz) \\
&+ \int_{\mathbb{R}} \frac{\epsilon}{2}[|\gamma(t_\alpha, x_\alpha, z)|^2 + |\gamma(t_\alpha, y_\alpha, z)|^2]\eta(t_\alpha)m(dz) \\
=:& I_1 + I_2 + I_3 + I_4 + I_5 + I_6.
\end{aligned}$$

The rest of the proof proceeds in a standard fashion, and we only review the main steps here. For the local terms we have, as usual from the viscosity solutions theory for second order differential equations that

$$\limsup_{\alpha\to\infty} I_1 \leq -\lambda\epsilon e^{\lambda(T-t_\epsilon)}|x_\epsilon|^2, \tag{4.14}$$

$$\limsup_{\alpha\to\infty} I_2 \leq K_2\frac{\epsilon}{2}e^{\lambda(T-t_\epsilon)}(1 + |x_\epsilon|^2), \tag{4.15}$$

$$\limsup_{\alpha\to\infty} I_3 \leq 0, \tag{4.16}$$

as $\alpha \to \infty$, where the matrix inequality (4.10) is used to show the last claim holds. For I_4, we also have

$$\limsup_{\alpha\to\infty} I_4 \leq K_4\frac{\epsilon}{2}e^{\lambda(T-t_\epsilon)}(1 + |x_\epsilon|^2). \tag{4.17}$$

For the nonlocal terms, global Lipschitz and growth conditions of γ together with continuity of η lead to

$$\limsup_{\alpha\to\infty} I_5 \leq 0, \tag{4.18}$$

$$\limsup_{\alpha\to\infty} I_6 \leq K_6\frac{\epsilon}{2}e^{\lambda(T-t_\epsilon)}(1 + |x_\epsilon|^2). \tag{4.19}$$

From (4.14)-(4.19), we have

$$\limsup_{\alpha\to\infty}\sum_{i=1}^{6} I_i \le \epsilon e^{\lambda(T-t_\epsilon)}[K(1+|x_\epsilon|^2)-\lambda|x_\epsilon|^2].$$

Choosing λ so large that $\lambda \ge K$, we can send $\epsilon \to 0$ to conclude

$$0<\mu \le \limsup_{\epsilon\to 0}\epsilon e^{\lambda(T-t_\epsilon)}[K(1+|x_\epsilon|^2)-\lambda|x_\epsilon|^2] \le 0,$$

a contradiction. □

5. Appendix

We have gathered here some of the more technical but fundamental results used in the article to be self-contained but not disrupt too much with the main themes. The following measurable mapping theorem is the cornerstone of our proof of the dynamic programming principle. We refer to the book by Bertzekas and Shreve [6] for this and other fundamental results and definitions in measure theory.

5.1. **Theorem** (Jankov-von Neumann theorem)**.** *Let X and Y be Borel spaces and A an analytic subset of $X \times Y$. There exists an analytically measurable function $\phi : proj_X(A) \to Y$ such that $Gr(\phi) \subset A$.*

We do not review analytic sets here, since only the fact that Borel sets are analytic is used in addition to the above theorem.

For viscosity solutions theory and results like those below, we refer to the standard reference [9]. For proving uniqueness theorems via semijets, the "Theorem of Sums" is crucial. We reproduce the statement here for reader's convenience.

5.2. **Theorem** (Theorem of Sums)**.** *Let $\mathcal{O}$ be a locally compact subset of $\mathbb{R}^n$ and $u, -v : \mathcal{O}_T \to \mathbb{R}$ be upper semicontinuous. Let ψ be defined on an open neighborhood of $(0,T)\times\mathcal{O}\times\mathcal{O}$ and such that $(t,x,y)\mapsto \psi(t,x,y)$ is once continuously differentiable in t and twice continuously differentiable in (x,y). Let $(\hat t,\hat x,\hat y)\in (0,T)\times\mathcal{O}\times\mathcal{O}$ and*

$$\Phi(t,x,y) := u(t,x)-v(t,y)-\psi(t,x,y) \le \Phi(\hat t,\hat x,\hat y)$$

for $0<t<T$ and $x,y\in\mathcal{O}$. Assume, moreover, that there is an $r>0$ such that for every $M>0$ there is a C such that

$$b_1 \le C \text{ whenever } (b_1,p,P)\in\mathcal{P}^{2,+}u(t,x), |x-\hat x|+|t-\hat t|\le r, \tag{5.3}$$

$$|u(t,x)|+|p|+|P|\le M, \tag{5.4}$$

$$b_2 \ge C \text{ whenever } (b_2,q,Q)\in\mathcal{P}^{2,-}v(t,x), |x-\hat x|+|t-\hat t|\le r, \tag{5.5}$$

$$|u(t,x)|+|q|+|Q|\le M. \tag{5.6}$$

Then for each $\kappa>0$ there are $P_\psi, Q_\psi \in \mathcal{S}_n, b_1, b_2\in\mathbb{R}$ such that

$$(b_1, D_x\psi(\hat t,\hat x,\hat y), P_\psi)\in \overline{\mathcal{P}}^{2,+}u(\hat t,\hat x),$$

$$(b_2, D_y\psi(\hat t,\hat x,\hat y), Q_\psi)\in \overline{\mathcal{P}}^{2,+}v(\hat t,\hat y),$$

$$-\Big(\frac{1}{\kappa}+\|D^2\psi(\hat t,\hat x,\hat y)\|\Big)I \le \begin{pmatrix} P_\psi & 0\\ 0 & Q_\psi\end{pmatrix} \le D^2\psi(\hat t,\hat x,\hat y)+\kappa[D^2\psi(\hat t,\hat x,\hat y)]^2$$

and

$$b_1-b_2=\partial_t\psi(\hat t,\hat x,\hat y).$$

The norm of the matrix A is $\|A\| := \sup\Big\{|Ax\cdot x|\Big||x|\le 1\Big\}$.

To use the above result, we employ a standard lemma for the doubling technique for viscosity solutions:

5.7. **Lemma.** *Let* $\underline{V} \in USC(\overline{\mathcal{O}_T})$, $\overline{V} \in LSC(\overline{\mathcal{O}_T})$ *satisfy*

$$\underline{V}(t,x), -\overline{V}(t,x) \leq C(1+x)$$

and let

$$\psi(t,x,y) := \frac{\alpha}{2}(x-y)^2 + \frac{\epsilon}{2}e^{\lambda(T-t)}(x^2+y^2).$$

Define

$$\Phi(t,x,y) := \underline{V}(t,x) - \overline{V}(t,y) - \psi(t,x,y).$$

Then

(i) *For each fixed* $0 < \epsilon < 1$, *the function* Φ *has a maximum point* $(t_\alpha, x_\alpha, y_\alpha)$,

(ii) *the maximum points satisfy*

$$x_\alpha - y_\alpha \to 0 \quad \text{as } \alpha \uparrow \infty, \tag{5.8}$$

$$\alpha|x_\alpha - y_\alpha|^2 \to 0 \quad \text{as } \alpha \uparrow \infty. \tag{5.9}$$

(iii) *There exist* t_ϵ *and* x_ϵ *such that*

$$(t_\alpha, x_\alpha, y_\alpha) \to (t_\epsilon, x_\epsilon, x_\epsilon)$$

along a subsequence as $\alpha \to \infty$.

References

[1] M. Arisawa. Corrigendum in the comparison theorems in "a new definition of viscosity solutions for a class of second-order degenerate elliptic integro-differential equations". *Ann. Inst. Henri Poincaré*, 24 (1):167–169, 2006.

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

[3] G. Barles. Convergence of numerical schemes for degenerate parabolic equations arising in finance theory. In *Numerical methods in Finance*, pages 1–21. Camb. Univ. Press, 1997.

[4] G. Barles and C. Imbert. Second-order elliptic integro-differential equations: viscosity solutions theory revisited. *preprint*, 2007.

[5] F.E. Benth, Jan Kallsen, and T. Meyer-Brandis. A non-gaussian ornstein-uhlenbeck process for electricity price modeling and derivatives pricing. *Applied Mathematical Finance*, 14 (2):153–169, 2007.

[6] D.P. Bertzekas and S. Shreve. *Stochastic Optimal Control: the Discrete Time Case*. Mathematics in Science and Engineering. Academic Press, 1978.

[7] M. Briani, C. La Chioma, and R. Natalini. Convergence of numerical schemes for viscosity solutions to integro-differential degenerate parabolic problems arising in financial theory. *Num. Math.*, 98 (4):607–646, 2004.

[8] R. Carmona and M. Ludkovski. Optimal switching with applications to energy tolling agreements. *preprint*, 2005.

[9] M.G. Crandall, H. Ishii, and P.-L. Lions. User's guide to viscosity solutions of second order partial differential equations. *Bulletin A.M.S.*, 27:1–67.

[10] A. Eydeland and K. Wolyniec. *Energy and Power Risk Management: New developments in Modeling, Pricing, and Hedging*. Wiley, 2003.

[11] W.H. Fleming and H.M. Soner. *Controlled Markov Processes and Viscosity Solutions*. Springer, New York.

[12] H. Geman. *Commodities and Commodity Derivatives*. Wiley, 2005.

[13] B. Hambly, S. Howison, and T. Kluge. Modelling spikes and swing options in electricity markets. *preprint*, 2007.

[14] J. Jacod and A.N. Shiryaev. *Limit Theorems for Stochastic Processes*. Springer, second edition, 2003.

[15] E. Jakobsen and K.H. Karlsen. A maximum principle applicable to partial integro-differential equations. *NoDEA*, 13 (2):137–165, 2006.

[16] Y. Kabanov and C. Klüppelberg. A geometric approach to portfolio optimization in models with transaction costs. *Fin. Stoch.*, 8:207–227, 2004.

[17] J. Keppo. Pricing of electricity swing options. *Journal of Derivatives*, 11:26–43, 2004.

[18] T. Kluge. *Pricing swing options and other derivatives*. PhD thesis, University of Oxford, 2006.

[19] W. Kluge. *Time-inhomogeneous Lévy processes in interest rate and credit risk models*. PhD thesis, University of Freiburg, 2005.

[20] M. Ludkovski. *Optimal switching with application to energy tolling agreements*. PhD thesis, Princeton University, 2005.

[21] A.-C. Lund and Fridthjof Ollmar. Analyzing flexible load contracts. *preprint*, 2003.

[22] K.R. Partasarathy. *Probability measures on metric spaces*. Academic Press, New York, 1967.

[23] A. Porchet, N. Touzi, and X. Warin. Valuation of a power plant under production constraints and market incompleteness. *preprint*, 2006.

[24] G. Da Prato and J. Zabczyck. *Stochastic equations in infinite dimensions*. Cambridge University Press, 1992.

[25] H.M. Soner and N. Touzi. Dynamic programming for stochastic target problems and geometric flows. *Journal of the European Mathematical Society*, 4:201–236, 2002.

[26] I. Vehviläinen. *Applying mathematical finance tools to the competitive Nordic electricity market*. PhD thesis, Helsinki University of Technology, 2006.

[27] M. Wilhelm. *Modeling, Pricing and Risk Management of Power Derivatives*. PhD thesis, ETH Zurich, 2007.

[28] J. Yong and X.Y. Zhou. *Stochastic Controls: Hamiltonian Systems and HJB Equations*. Springer, Berlin-Heidelberg-New York, 1999.

(Olli Wallin)
Centre of Mathematics for Applications (CMA)
University of Oslo
P.O. Box 1053,Blindern
N–0316 Oslo, Norway
E-mail address, Olli Wallin: `olliw@math.uio.no`