Open access • Journal Article • DOI:10.1007/S00245-009-9095-8

Viscosity Solutions for a System of Integro-PDEs and Connections to Optimal Switching and Control of Jump-Diffusion Processes — Source link

Imran H. Biswas, Espen R. Jakobsen, Kenneth H. Karlsen

Institutions: ETH Zurich, Norwegian University of Science and Technology, University of Oslo

Published on: 01 Aug 2010 - Applied Mathematics and Optimization (Springer-Verlag)

Topics: Viscosity solution, Jump diffusion, Nonlinear system, Viscosity (programming) and Differential equation

Viscosity Solutions for a System of Integro-PDEs and Connections to Optimal Switching and Control of Jump-Diffusion Processes

Journal Article

Author(s):
Biswas, Imran H.; Jakobsen, Espen R.; Karlsen, Kenneth H.

Publication date:
2010

Permanent link:
https://doi.org/10.3929/ethz-b-000020518

Originally published in:
Applied Mathematics & Optimization 62(1), https://doi.org/10.1007/s00245-009-9095-8

Appl Math Optim (2010) 62: 47–80
DOI 10.1007/s00245-009-9095-8

Viscosity Solutions for a System of Integro-PDEs and Connections to Optimal Switching and Control of Jump-Diffusion Processes

Imran H. Biswas · Espen R. Jakobsen · Kenneth H. Karlsen

Published online: 10 December 2009

Abstract We develop a viscosity solution theory for a system of nonlinear degenerate parabolic integro-partial differential equations (IPDEs) related to stochastic optimal switching and control problems or stochastic games. In the case of stochastic optimal switching and control, we prove via dynamic programming methods that the value function is a viscosity solution of the IPDEs. In our setting the value functions or the solutions of the IPDEs are not smooth, so classical verification theorems do not apply.

Keywords Integro-partial differential equations · Dynamic programming method · Viscosity solutions · Optimal stochastic control and switching · Lévy processes

Communicating Editor: Frederic Bonnans.

This research is supported by the Research Council of Norway through an Outstanding Young Investigators Award (KHK) and partially through the project "Integro-PDEs: Numerical methods, Analysis, and Applications to Finance".

I.H. Biswas (✉)
Seminar für Angewandte Mathematik, D-MATH, ETH-Zurich, HG G 56.2, Rämistrasse 101, 8092, Zurich, Switzerland
e-mail: imran.biswas@sam.math.ethz.ch

E.R. Jakobsen
Norwegian University of Science and Technology, 7491, Trondheim, Norway
e-mail: erj@math.ntnu.no

K.H. Karlsen
Centre of Mathematics for Applications, University of Oslo, P.O. Box 1053, Blindern, 0316 Oslo, Norway
e-mail: kennethk@math.uio.no
url: folk.uio.no/kennethk

1 Introduction

In this paper we analyze a system of integro-partial differential equations (IPDEs henceforth) related to stochastic optimal switching and control or stochastic games. In the case of stochastic optimal switching and control problems, we prove via the dynamic programming method that the value function is a viscosity solution of the relevant IPDE. Such results exist in the pure PDE case [19, 46], and this paper is partly motivated by a desire to extend these results to the non-local case.

The system of equations involves M equations and is of the form

$$\begin{aligned}&\mathcal{F}_i(t,x,u(t,x),\partial_t u^i(t,x),Du^i(t,x),D^2u^i(t,x),u^i(t,\cdot))=0\\&\text{in } (0,T)\times\mathbb{R}^n,\ i\in\mathcal{I},\end{aligned}\tag{1.1}$$

for $\mathcal{I}=\{1,2,\dots,M\}$. We also impose an initial condition

$$u^i(0,x)=g_i(x)\quad \text{in } \mathbb{R}^n,\ i\in\mathcal{I}.$$

Here, $g=(g_1,g_2,\dots,g_M)$ and $u=(u^1,u^2,\dots,u^M)$ are $\mathbb{R}^M$ valued functions. The nonlocal nature of the system (1.1), indicated by the term "$u^i(t,\cdot)$", is the main focus of this paper. The nonlinear and nonlocal functions $\mathcal{F}_i$ are defined as

$$\begin{aligned}\mathcal{F}_i(t,x,r,p_t,p_x,X,\varphi(\cdot))=\max\Big\{&p_t+\sup_{\alpha\in\mathcal{A}_i}\inf_{\beta\in\mathcal{B}_i}\big[\mathcal{L}_i^{\alpha,\beta}(t,x,r_i,p_x,X)-\mathcal{J}_i^{\alpha,\beta}\varphi\big];\\&r_i-\mathcal{M}^i r\Big\},\end{aligned}$$

for $(t,x,r,p_t,p_x,X)\in\mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^M\times\mathbb{R}\times\mathbb{R}^n\times\mathbb{S}^n$ ($\mathbb{S}^n$ the set of $n\times n$ symmetric matrices) and any smooth real-valued function $\varphi(t,x)$. The operators $\mathcal{L}_i^{\alpha,\beta}$, $\mathcal{J}_i^{\alpha,\beta}$, and $\mathcal{M}^i$ are defined as follows:

$$\begin{aligned}\mathcal{L}_i^{\alpha,\beta}(t,x,r_i,p_x,X)&=-\operatorname{Tr}(a_i^{\alpha,\beta}(t,x)X)-b_i^{\alpha,\beta}(t,x)p_x+c_i^{\alpha,\beta}(t,x)r_i-f_i^{\alpha,\beta}(t,x),\\\mathcal{J}_i^{\alpha,\beta}\varphi&=\int_E[\varphi(t,x+\eta_i^{\alpha,\beta}(t,x,z))-\varphi-\mathbf{1}_{|z|\le 1}\eta_i^{\alpha,\beta}(t,x,z)D_x\varphi]\nu(dz),\\\mathcal{M}^i u&=\min_{j\neq i}\{u_j+k(i,j)\},\end{aligned}$$

where $k(i,j)\ge 0$, $E=\mathbb{R}^m\backslash\{0\}$, and ν is a positive Radon measure on E (the Lévy measure) with an "at most" second order singularity at the origin and exponential decay at infinity. The sets $\mathcal{A}_i$ and $\mathcal{B}_i$ are separable metric spaces (typically subsets of some Euclidean space) and the coefficients a_i,η_i,b_i,c_i,f_i are functions taking values respectively in $\mathbb{R}^{n\times n},\mathbb{R}^n,\mathbb{R}^n,\mathbb{R},\mathbb{R}$. The specific assumptions on the coefficients will be stated later, but roughly speaking we will assume that the coefficients are Lipschitz continuous in x.

The matrices $a_i^{\alpha,\beta}$ are assumed to be merely non-negative definite and as such can vanish at some points. Similarly, the jump vectors $\eta_i^{\alpha,\beta}\ge 0$ can vanish. Consequently, there are no regularization effects in this problem coming from the second order operator ("Laplacian smoothing") or from the integral operator ("fractional Laplacian smoothing"). Because of this, the system (1.1) will in general not have classical solutions, and a suitable notion of viscosity solutions is needed.

As already mentioned, the system (1.1) is closely related to the optimal control of Jump-diffusion (Lévy) processes. It arises formally as the Bellman-Isaacs equation for zero-sum stochastic games where the state is given by controlled jump-diffusion processes involving also switching between different control regimes (indexed by i). The maximizing player (α) disposes both "continuous" and "switching controls" while the minimizing player (β) only disposes "continuous controls". If the sets $\mathcal{B}_i$ are singletons (no minimizing player), then the system (1.1) is the convex Bellman equation related to optimal control of jump-diffusion process with both continuous and switching controls.

In case of pure diffusions (i.e., $\nu = 0$), the value function of the control problem satisfies a dynamic programming principle (see [19, 22, 46]), which implies that it is a viscosity solution of a system like (1.1). However, for processes with jumps, to the best of our knowledge, there is no proved dynamic programming principle in the literature that covers the generality of (1.1). We refer to [29] for some rigorous results in two space dimensions. Generally speaking, experts expect the dynamic programming principle to hold and frequently use it without proof. In this paper, using well-known arguments, we have chosen to include a rigorous proof.

We mention that control problems involving switching have applicability to real-life problems such as production planning in a flexible manufacturing system (see [23, 42] and the references therein). In this context, the control problems are typically modeled by using diffusion processes leading to pure PDEs, but it is not far fetched to think about more general models with jumps in the state dynamics, thereby motivating the study of systems like (1.1). Another important area of application is portfolio optimization for an investor operating in multiple Lévy driven markets. It is feasible to assume that this investor has to pay a certain premium when pulling out from one market and entering into another one. In such a scenario, the investor would like to optimize the value of his portfolio by switching from one market to another and also continuously changing the portfolio while remaining in the same market. This portfolio optimization problem can be viewed as an optimal switching problem and one gets a system of nonlocal variational inequalities as the Bellman equation. In fact, while being in the same market, the agent would always look to change his holdings depending on different market modes , say the bull and bear modes, which appear randomly in an economic cycle. In such a scenario, it is possible to think that the market and the investor are engaged in a switching game; we refer to [11] for more in this direction.

In addition to the applications mentioned above, we are also motivated by the problem of deriving error estimates for numerical schemes for second order Hamilton-Jacobi-Bellman equations. This is a difficult problem that remained open for a long time before the works of Krylov [33–35] and Barles & Jakobsen [8–10]. We also mention [16, 18] as important recent contributions in this area. The methods developed in these works involve the use of carefully chosen smooth approximations of the viscosity solution of the underlying equation. In some recent developments [9, 10], Barles & Jakobsen used solutions of certain switching systems to generate suitable approximations of the viscosity solution of the Bellman equation associated with the optimal control of diffusion processes. In a separate piece of work [15], we adapt this approach to the nonlocal Bellman equation of controlled jump-diffusion processes, which is drawing a lot of interests these days due to its applications in mathematical finance (see for example [2, 3, 12, 13, 20] and the references therein). To derive error estimates like those in [9, 10] for the nonlocal Bellman equation we need to have at our disposal a viscosity solution theory for switching systems of the type (1.1).

The viscosity solution theory for second order PDEs is well developed [21] and has become an essential tool in the study of controlled diffusions [4, 24]. Expanding its availability beyond scalar equations, viscosity solution theory for systems has been advanced to understand the optimal switching of controls for both deterministic [19, 36, 43, 44, 47] and stochastic [27, 28, 48] problems; these works offer a number of results on existence, uniqueness, and qualitative properties of solutions. On the other hand, the viscosity solution approach to nonlocal equations is still under development and is currently an active research area, cf. for example [1, 2, 6, 7, 17, 30–32, 40, 41]. Contrary to its pure PDE counterpart, the available literature applying viscosity solutions to systems of integro-PDEs is very limited, but see [5] (switching systems are not covered).

The contributions of this article can be divided into two main parts. The first part includes a comprehensive study of viscosity solutions for the system (1.1), while the second one analyzes the problem of optimal switching of stochastic controls. It is not difficult to adapt techniques from stochastic analysis to prove, for example, the existence of viscosity solutions of the underlying Bellman equation. In the present context, the Bellman equation related to the optimal switching problem serves as an example of the system (1.1), but it does not cover the general form and therefore we mostly rely on PDE techniques [21] to prove our results, including existence and uniqueness of (suitably defined) viscosity solutions, continuous dependence estimates, and some regularity results.

The rest of paper is organized as follows: in Sect. 2 we list all the notations, state the full set of assumptions, and define viscosity sub- and supersolutions along with equivalent characterizations. We also state the comparison principle, uniqueness, and existence results in this section. The optimal switching problem with a jump-diffusion driven state process is introduced and analyzed in Sect. 3. The main result of this section is the proof of the dynamic programming principle. In Sect. 4 we prove the results stated in Sect. 2. Finally, Sect. 5 contains a continuous dependence estimate along with an application to the Hölder continuity of viscosity solutions.

2 Notation, Assumption, Well-posedness, and Regularity

We denote the set $\{1, \ldots, M\}$ by $\mathcal{I}$. We also use the notations Q_T and $\bar{Q}_T$ respectively for $(0, T) \times \mathbb{R}^n$ and $[0, T) \times \mathbb{R}^n$. For various constants depending on the data we mainly use N, K, C with/without subscripts.

For a bounded Lipschitz continuous function $h(x)$ defined on $\mathbb{R}^n$, its Lipschitz norm $|h|_1$ is defined as

$$|h|_1 := \sup_{x\in\mathbb{R}^n} |h(x)| + \sup_{x,y\in\mathbb{R}^n} \frac{|h(x)-h(y)|}{|x-y|},$$

and denote the space of all h so that $|h|_1 < \infty$ by $C_b^1(\mathbb{R}^n)$ or sometimes only by C_b^1. We also define

$$C_b^{\frac{1}{2},1}(\bar{Q}_T) := \left\{ h(t,x) : \sup_{(t,x)\in\bar{Q}_T} |h(t,x)| + \sup_{(t,x),(s,y)\in\bar{Q}_T} \frac{|h(t,x)-h(s,y)|}{|t-s|^{\frac{1}{2}} + |x-y|} < \infty \right\}.$$

For $|h(t,\cdot)|$ we simply mean $|\cdot|_1$ norm of $h(t,x)$ as a function of x alone and for a fixed t. Let $C^{1,2}((0,T)\times\mathbb{R}^n)$ be the space of once in time and twice in space continuously differentiable functions. Also, denote the set of all upper and lower semicontinuous functions on $\bar{Q}_T$ respectively by $USC(\bar{Q}_T)$ and $LSC(\bar{Q}_T)$. A lower index would mean polynomial growth at infinity, therefore the spaces $USC_p(\bar{Q}_T), LSC_p(\bar{Q}_T), C_p^{1,2}((0,T)\times\mathbb{R}^n)$ contain the functions h respectively from $USC(\bar{Q}_T), LSC(\bar{Q}_T), C^{1,2}((0,T)\times\mathbb{R}^n)$ satisfying the growth condition

$$|h(x)| \le C(1+|x|^p) \quad \text{for all } x\in\mathbb{R}^n \text{ (uniformly in } t \text{ if } h \text{ depends on } t).$$

We identify the spaces $USC_0(\bar{Q}_T)$ and $LSC_0(\bar{Q}_T)$ respectively with $USC_b(\bar{Q}_T)$ and $LSC_b(\bar{Q}_T)$; "b" is an index signifying boundedness. From time to time we will not explicitly mention the control parameters α,β and this will be done on occasions where the assertions are valid for all parameters.

Now we list the assumptions on the data:

(A.1) $a_i^{\alpha,\beta} = \frac{1}{2}\sigma_i^{\alpha,\beta}\sigma_i^{\alpha,\beta T}$ and $\sigma_i, b_i, c_i, f_i, \eta_i$ are continuous functions of t,x,α,β; $\mathcal{A}_i, \mathcal{B}_i$ are compact metric spaces; and the positive Radon measure ν defined on E satisfies

$$\int_{0<|z|\le 1} |z|^2\nu(dz) + \int_{|z|\ge 1} e^{\Lambda|z|}\nu(dz) \le K$$

for some $K, \Lambda > 0$.

(A.2) For any α,β and $i,j\in\mathcal{I}$; and for each $t\in[0,T]$

$$\begin{aligned} &|f_i^{\alpha,\beta}(t,\cdot)|_1 + |g_i|_1 \le K; \\ &k(i,i) = 0, \qquad k(i,j) > 0 \quad \text{for } i\neq j; \\ &g_i(x) - \mathcal{M}^i g(x) \le 0 \quad \text{for all } x\in\mathbb{R}^n. \end{aligned}$$

(A.3) $c_i^{\alpha,\beta} \ge 0$ for all i and α,β and

$$\begin{aligned} &|\sigma_i^{\alpha,\beta}(t,x) - \sigma_i^{\alpha,\beta}(t,y)| + |b_i^{\alpha,\beta}(t,x) - b_i^{\alpha,\beta}(t,y)| + |c_i^{\alpha,\beta}(t,x) - c_i^{\alpha,\beta}(t,y)| \\ &\quad \le K|x-y|. \end{aligned}$$

(A.4) For all i,t,x,α,β, $\eta_i^{\alpha,\beta}(t,x,z)$ is Borel measurable in z and

$$\begin{aligned} |\eta_i^{\alpha,\beta}(t,x,z)| &\le K\big(|z|\wedge 1\big), \\ |\eta_i^{\alpha,\beta}(t,x,z) - \eta_i^{\alpha,\beta}(t,y,z)| &\le K\big(|z|\wedge 1\big)|x-y|. \end{aligned}$$

Remark The assumptions (A.1)–(A.4) are natural and standard, except maybe for the boundedness of f, g and the decay of ν at infinity. These last assumptions can be relaxed and the results of this paper still hold in a properly modified form. The integrability assumptions on ν are natural in financial applications. Boundedness of f, g will imply bounded solutions, an assumption we make for the sake of simplicity. The requirement that $c_i \geq 0$ can be relaxed, via an exponential scaling of the solution, to the requirement that the functions c_i are bounded from below. The last assumption guarantees that the non-local part is well defined for smooth solutions with less than exponential growth at infinity.

Remark It will turn out that continuous viscosity solutions (as well as classical solutions) of (1.1) satisfy $u^i(t,x) - \mathcal{M}^i u(t,x) \leq 0$ for all i, t, x. Letting $t \to 0$ leads to $u^i(0,x) - \mathcal{M}^i u(0,x) \leq 0$. Therefore, the (compatibility) condition on g in (A.2) is necessary for viscosity solutions to be continuous in t at $t = 0$.

Next, we are going to give the definition of sub- and supersolutions to (1.1), which includes the initial condition as a part of it. Before doing so, we need to introduce the following quantities, for $\kappa \in (0,1)$:

$$\begin{aligned} \mathcal{J}_{i,\kappa}^{\alpha,\beta}(t,x,q,\phi(t,\cdot)) &= \int_{B(0,\kappa)\setminus\{0\}} \big(\varphi(t, x + \eta_i^{\alpha,\beta}(t,x,z)) \\ &\quad - \varphi - \eta_i^{\alpha,\beta}(t,x,z)q\big)\nu(dz), \\ \mathcal{J}_i^{\alpha,\beta,\kappa}(t,x,q,v(t,\cdot)) &= \int_{B(0,\kappa)^C} \big(v(t,x+\eta_i^{\alpha,\beta}(t,x,z)) - v - \mathbf{1}_{|z|\leq 1}\eta_i^{\alpha,\beta}.q\big)\nu(dz) \end{aligned}$$

and

$$\begin{aligned} &\mathcal{F}_i^\kappa\big(t,x,r,p_t,p_x,X,v(t,\cdot),\varphi(t,\cdot)\big) \\ &\quad = \max\Big[p_t + \sup_{\alpha\in\mathcal{A}_i}\inf_{\beta\in\mathcal{B}_i}\Big\{-\mathrm{Tr}(a_i^{\alpha,\beta}(t,x)X) - b_i^{\alpha,\beta}(t,x)p_x + c_i^{\alpha,\beta}(t,x)r_i \\ &\qquad - f_i^{\alpha,\beta}(t,x) - \mathcal{J}_{i,\kappa}^{\alpha,\beta}(t,x,p_x,\phi(t,\cdot)) - \mathcal{J}_i^{\alpha,\beta,\kappa}(t,x,p_x,v(t,\cdot))\Big\}; r_i - \mathcal{M}^i r\Big]. \end{aligned}$$

Definition 2.1 (i) A function $u \in USC_p([0,T]\times\mathbb{R}^n)$ is a viscosity subsolution of (1.1) if

$$u_i(0,x) \leq g_i(x), \quad x \in \mathbb{R}^n,\ 1 \leq i \leq M$$

and if for any $1 \leq i \leq M, \varphi \in C^{1,2}((0,T]\times\mathbb{R}^n)$, wherever $(t,x) \in (0,T)\times\mathbb{R}^n$ is a global maximum point of $u_i - \varphi$

$$\mathcal{F}_i^\kappa(t,x,u(t,x),\varphi_t(t,x),D\varphi(t,x),D^2\varphi(t,x),u_i(t,\cdot),\varphi(t,\cdot)) \leq 0 \quad \forall\kappa \in (0,1).$$

(ii) A function $u \in LSC_p([0,T]\times\mathbb{R}^n)$ is a viscosity supersolution of (1.1) if

$$u_i(0,x) \geq g_i(x), \quad x \in \mathbb{R}^n,\ 1 \leq i \leq M$$

and if for any $1 \le i \le M, \varphi \in C^{1,2}((0,T] \times \mathbb{R}^n)$, wherever $(t,x) \in (0,T) \times \mathbb{R}^n$ is a global minimum point of $u_i - \varphi$

$$\mathcal{F}_i^{\kappa}(t,x,u(t,x),\varphi_t(t,x),D\varphi(t,x),D^2\varphi(t,x),u_i(t,\cdot),\varphi(t,\cdot)) \ge 0 \quad \forall \kappa \in (0,1).$$

(iii) A function $u \in C_p([0,T] \times \mathbb{R}^n)$ is a viscosity solution of (1.1) if it is both a sub and a supersolution.

This definition is formulated in terms of test functions φ. Note that the test function appears in the non-local slot, which is unavoidable when ν is singular. Some growth assumptions on the sub- and supersolutions are needed for the integral term to be finite; our polynomial growth assumption is not optimal but sufficient for our needs.

As usual, any classical solution is also a viscosity solution and any smooth viscosity solution is a classical solution. Moreover, an equivalent definition is obtained by replacing "global maxima/minima" with "strict global maxima/minima" in the above definition. We may also assume that $\phi = u_i$ at the maximum/minimum point.

Next, we give an alternative definition which will be used when proving existence of solutions via Perron's method.

Lemma 2.1 (Alternative definition) *A function* $v \in USC_p([0,T] \times \mathbb{R}^n)$ (*or* $v \in LSC_p([0,T] \times \mathbb{R}^n : \mathbb{R}^M)$) *is a subsolution* (*supersolution*) *to* (1.1) *iff* $v^i(0,x) \le g(x)(v^i(0,x) \ge g(x))$ *for all* $i \in \mathcal{I}$ *and for every* $(t,x) \in Q_T$ *and* $\phi \in C_p^{1,2}(Q_T)$ *such that* (t,x) *is a global maxima* (*global minima*) *of* $v^i - \phi$ *then*

$$\mathcal{F}_i(t,x,u,\partial_t(t,x),\phi(t,x),D\phi(t,x),D^2\phi(t,x),\phi(t,\cdot)) \le 0\ (\ge 0)$$

The proof of the lemma is similar to the scalar case, see [31] or [40].

Remark The choice of (0, 1) as the domain of κ does not influence the Definition 2.1. One can equivalently replace (0, 1) by an interval of type $(0,\delta)$ for $\delta > 0$. All such choices for domain of κ could be proven to be equivalent to the alternative definition in Lemma 2.1. However, in order for our methodology to work, we need to be able to pass to the limit $\kappa \to 0$ and the Definition 2.1 is formulated with that in mind.

Remark Traditionally [21], to prove uniqueness of solutions we need to work with the sub- and superjets of a solution u. However, due to the singular non-local part of these equations, it is not straightforward to give, as in the local case, a definition in terms of sub- or superjets. In this paper these jets are introduced via a "non-local" maximum principle of semi-continuous functions [31], see Lemma 4.1. We also refer to [7] for slightly different but (in this setting) equivalent way of doing this.

Next, we state the comparison, existence, uniqueness, and regularity results for bounded viscosity solutions of (1.1). The proofs will be given in Sects. 4 and 5.

Theorem 2.2 (Comparison) *Assume* (A.1)–(A.4). *Let* $u, -v \in USC(\bar{\mathbb{Q}}_T; \mathbb{R}^M)$ *be respectively sub- and supersolutions of* (1.1) *such that* $u_i(x), -v_i(x) \le C(1+|x|^2)$ *for*

$i = 1, \dots, M$, *then*

$$u_i \le v_i \quad \text{for } i = 1, \dots, M.$$

Theorem 2.3 (Existence) *Assume* (A.1)–(A.4), *and the existence of two functions* $\bar{u} \in USC_b([0, T) \times \mathbb{R}^n : \mathbb{R}^M)$ *and* $\bar{v} \in LSC_b([0, T) \times \mathbb{R}^n : \mathbb{R}^M)$ *which are respectively sub- and supersolutions of* (1.1). *Then there exists a unique viscosity solution* $u \in C_b([0, T) \times \mathbb{R}^n : \mathbb{R}^M)$ *to the system* (1.1) *satisfying* $\bar{u} \le u \le \bar{v}$.

Since $\pm(Kt + |g|_0)$ are sub- and supersolutions of (1.1), the two previous theorems immediately give existence and uniqueness of a bounded viscosity solution of (1.1):

Corollary 2.4 *Assume* (A.1)–(A.4). *There exists a unique viscosity solution* $u \in C_b(\bar{Q}_T)$ *of the system* (1.1) *satisfying*

$$|u^i(t, x)| \le Kt + |g|_0$$

for all $(t, x) \in \bar{Q}_T$, *where* K *comes from* (*A*.2).

The comparison principle is stated for sub- and supersolutions of quadratic growth. This is more than what is needed for uniqueness and existence of bounded solutions, but we will need it later when we prove time regularity of the solution.

The unique viscosity solution of (1.1) enjoys the following regularity:

Theorem 2.5 (Regularity) *Assume* (A.1)–(A.4), *and let* u *be the viscosity solution of* (1.1). *Then there is a constant* C, *depending on the data*, *such that*

$$|u^i(t, x) - u^i(s, y)| \le C\big[|x - y| + (1 + |x| + |y|)|t - s|^{\frac{1}{2}}\big],$$

for all $(t, x), (s, y) \in \bar{Q}_T$ *and* $i \in \mathcal{I}$.

3 Optimal Switching of Stochastic Controls

We want to prove a connection between optimal switching problems for Lévy processes and systems of nonlocal equations of the form (1.1). If $\mathcal{A}_i = \mathcal{U}$ for all i and the sets B_i are singletons (no β dependence), we prove that the value function of the switching control problem is a viscosity solution of a system like (1.1).

For $(t, x) \in [0, T) \times \mathbb{R}^n$, consider the following stochastic differential equation on a filtered probability space $\big(\Omega, \mathcal{F}_t, P, \mathcal{F}_{t,\cdot}\big)$ [where $\mathcal{F}_t$ is a σ algebra and $\mathcal{F}_{t,\cdot}$ is the shorthand for a filtration $(\mathcal{F}_{t,s})_{s \ge t}$.]:

$$\begin{aligned} dY(s) &= b(s, Y(s); a(s), \zeta(s))ds + \sigma(s, Y(s); a(s), \zeta(s))dW(s) \\ &\quad + \int_{\mathbb{R}^m \setminus \{0\}} \eta(s, Y(s^-), z; a(s^-), \zeta(s^-))d\tilde{N}(ds, dz) \end{aligned} \tag{3.1}$$

with

$$Y(t) := x \in \mathbb{R}^n \quad \text{and} \quad s \in (t, T]$$

for some positive constant $T > 0$. In the above SDE, the b, η's are $\mathbb{R}^n$ valued functions, N is Poisson random measure on $\mathbb{R}^m \times (\Omega, \mathcal{F}, P)$ and $W(s)$ is a k-dimensional Brownian motion defined on $(\Omega, \mathcal{F}, P)$. The diffusion coefficients σ are $n \times k$ matrices. The control processes $\zeta(s)$ and $a(s)$ take values respectively in a metric space $\mathbb{U}$ and in a finite set $\mathcal{A} = \{1, 2, \ldots, M\}$.

Definition 3.1 (Admissible control)

(i) An admissible (*continuous*) control $\zeta(s)$, for $s \in [t, T]$, is a $\mathcal{U}$-valued càdlàg process adapted to the filtration $\mathcal{F}_{t,\cdot}$.
(ii) An admissible (*switching*) control a is a sequence of switching times τ_i and switching decisions d_i, i.e.,

$$a := \{\tau_i, d_i\}_{i \geq 0}$$

such that each τ_i is a $\mathcal{F}_{t,\cdot}$ stopping time with

$$t = \tau_0 \leq \tau_1 \leq \tau_2 \leq \cdots \leq \tau_i \leq \tau_{i+1} \leq \cdots \leq T$$

and d_i is $\mathcal{F}_{t,\tau_i}$ measurable with values in $\mathcal{A}$.

For each $d = 1, 2, \ldots, M$ we denote the set of all admissible (*switching*) controls starting with d as $\mathcal{A}^d(t)$ and set of all (*continuous*) admissible controls by $\mathcal{U}(t)$, i.e.,

$$\mathcal{A}^d(t) = \big\{a = \{\tau_i, d_i\}_{i\geq 0} : a \text{ is an admissible (\textit{switching}) control and } d_0 = d\big\},$$
$$\mathcal{U}(t) = \big\{\zeta(\cdot) : \zeta(s) \text{ is an admissible (\textit{continuous}) control on } [t, T]\big\}.$$

Any admissible switching control $a = \{d_i, \tau_i\}_{i\geq 0}$ could be thought of as a control process $a(s)$ as follows:

$$a(s) = \sum_{i\geq 1} d_{i-1}\chi_{[\tau_{i-1},\tau_i)}(s),$$

which is obviously cádlág .

For given control processes $a(\cdot) \in \mathcal{A}^d(t)$ and $\zeta(\cdot) \in \mathcal{U}(t)$, the cost functional associated with the control problem is given by the expectation value

$$J^d_{t,x}(a(\cdot), \zeta(\cdot)) := \mathbb{E}_{t,x}\left[\int_t^T f(s, Y(s), a(s), \zeta(s))ds + g(Y(T)) + \sum_{j\geq 1} k(d_{j-1}, d_j)\right],$$

where $Y(\cdot)$ is the solution to (3.1) with controls $a(s), \zeta(s)$ and $k(i, j)$ is the cost of switching from decision i to decision j for all $i, j \in \mathcal{A}$. We also note that due to the nontrivial switching cost (i.e., $k(i, j) > 0$ for $i \neq j$), one is likely to get different values of the cost functional for different initial values of a. Next, we formally state the control problem.

Optimal Switching Problem: For any $(t,x,d) \in [0,T) \times \mathbb{R}^d \times \mathcal{A}$, determine $(a^*(\cdot), \zeta^*(\cdot)) \in \mathcal{A}^d(t) \times \mathcal{U}(t)$ such that

$$V^d(t,x) \equiv J^d_{t,x}(a^*(\cdot), \zeta^*(\cdot)) = \inf_{(a(\cdot),\zeta(\cdot)) \in \mathcal{A}^d(t) \times \mathcal{U}(t)} J^d_{t,x}(a(\cdot), \zeta(\cdot)).$$

The vector valued function $V(t,x) := (V^1(t,x), V^2(t,x), \dots, V^M(t,x))$ is called the value function of the control problem.

Remark In our definition of admissible control we allow an infinite number of switching times. Since such controls incur an infinite switching cost, they will never be minimizing costs for the control problem, and hence we could restrict $\mathcal{A}^d(t)$ to controls having a finite number of switchings. We also add that $\sigma, b, f, \eta, k(\cdot,\cdot), \nu$ satisfy assumptions (A.1), (A.2), (A.3), and (A.4) with the convention that $g_i(x) = g(x)$ for all i (g is scalar now).

Optimal switching control problems have been studied by many authors over the last few decades, we refer to [19, 36, 45–47] and references therein. These references mainly consider processes without jumps (continuous sample paths) and the corresponding Bellman-Isaacs equations are pure PDEs. An exception is a series of papers by Lenhart and co-workers on piece-wise deterministic processes (with finite Lévy measures), see, e.g., [36]. To the best of our knowledge the optimal switching problem has not been studied before in a general Lévy setting. In this section we provide results for the general Lévy case. The analysis mainly follow [46] but we have to overcome additional non-trivial technical difficulties due to the fact that the state evolution has discontinuous sample paths. The classical approach is to prove that the value function is the unique viscosity solution of the underlying Bellman equation is via dynamic programming principle. In [46] the authors use this approach but with a canonical choice of the underlying probability space, the Wiener space $C_0([t,T]:\mathbb{R}^n)$.

In the case of stochastic evolutions driven by Lévy processes, the canonical sample space consists of all $\mathbb{R}^{k+m}$-valued cádlág functions on $[t,T]$ starting at 0. This space equipped with a complete separable metric, the so called Skorohod metric, is called the Skorohod space and is denoted by $D[t,T]$. The Skorohod space, its defining topology and analysis of probability measures on this space is way more complicated than the Wiener space and one is required to be careful while drawing conclusions on technical grounds. For more information about the Skorohod space we refer to [14].

3.1 The Canonical Sample Space

For the problem (3.1), the canonical sample space $\Omega_{t,s}$, $0 \le t < s \le T$, is defined as

$$\Omega_{t,s} = D[t,s] = \Big\{ w \in \text{cadlag}\big([t,s]; \mathbb{R}^{k+m}\big)\Big\}.$$

Let $\mathcal{F}_{t,s}$ denote the Borel σ-algebra on $\Omega_{t,s}$ (with the Skorohod topology). We will use the convention that $\Omega_{t,T} = \Omega_t$ and $\mathcal{F}_{t,T} \equiv \mathcal{F}_t$.

The next issue is to ensure the existence of a probability measure and a compatible Lévy process which will be our candidate for driving the dynamics. We would like to recall that a Lévy process is characterized by its distribution, which is infinitely divisible in nature and equivalently characterized by its characteristic triplet (γ, A, ν), where γ is the drift of the process, A being the co-variance matrix, and ν is the so-called Lévy measure. To this end, we define a positive Lévy measure ν' on $\mathbb{R}^{k+m}$ and a $(k+m)\times(k+m)$ covariance matrix I' as follows

$$\nu'(G) = \nu(\mathbb{P}_m(G)); \qquad I' = \begin{pmatrix} I_{k\times k} & 0 \\ 0 & 0 \end{pmatrix}$$

where $G \subset \mathbb{R}^{k+m}$ is a Borel set and $\mathbb{P}_m : \mathbb{R}^k \times \mathbb{R}^m \to \mathbb{R}^m$ is the usual projection. For the characteristic triplet $(0, I', \nu')$, there exists a probability measure P_t and a compatible Lévy process $X_t(s)$ taking values in $\mathbb{R}^{k+m}$ with the same characteristic triplet. In view of the Lévy-Ito decomposition, $X_t = (W_t(\cdot), N(dw, ds))$ where $W_t(\cdot)$ is a k-dimensional Brownian motion and $N(dw, ds)$ could be considered as a Poisson random measure on $\mathbb{R}^m\backslash\{0\}$. The probability measure P_t is the one induced by the random variable $X_t(T)$. For more on existence and related topics on Lévy processes we refer to [14, 39].

Once we have made the choice for $(P_t, X_t(\cdot))$, we choose the driving Lévy process $X_s(\cdot)$, for $s > t$, in the following manner:

$$X_s(r) := X_t(s) - X_t(r)$$

which is also a Lévy process starting at s, thanks to the generic properties of Lévy processes and that the probability measure on Ω_s has been chosen as the one induced by $X_s(T)$.

Next, we address some technical issues and verify some assertions so that we can argue along the lines of [46]. For any $\tau \in (t, T)$ (deterministic) and $\omega \in \Omega_t$ define

$$\begin{aligned} \omega_1 &:= \omega|_{[t,\tau^-]} \quad (\tau^- \text{ signifies the left limits}), \\ \omega_2 &:= (\omega - \omega_\tau)|_{[\tau,T]}, \\ \pi(\omega) &:= (\omega_1, \omega_2). \end{aligned}$$

The map $\pi : \Omega_t \to \Omega_{t,\tau} \times \Omega_{\tau,T}$ is well defined. Next, we prove the following lemma:

Lemma 3.1 *For any $\tau \in (t, T)$ (deterministic),*

$$P_t\left(\{\omega \in \Omega_t : \omega \text{ is discontinuous at } \tau\}\right) = 0.$$

Proof Let Ω_1 be the sample paths of the Lévy process $X_t(\cdot)$, i.e.,

$$\Omega_1 = \left\{X_t(\cdot, \omega) : \omega \in \Omega_t\right\}.$$

Then $P_t(\Omega_1) = P_t(X_t(T)(\Omega_t)) = 1$, and therefore

$$P_t(\{\omega \in \Omega_t : \omega \text{ is discontinuous at } \tau\}) = P_t(\{\omega \text{ is discontinuous at } \tau\} \cap \Omega_1) = 0,$$

where the last equality follows by stochastic continuity of the Lévy process X_t. $\square$

Now, $\pi^{-1} : \Omega_{t,\tau} \times \Omega_{\tau,T} \to \Omega_t$ is the following map

$$\pi^{-1}(\omega_1, \omega_2) = \begin{cases} \omega_1(s) & \text{if } s \in [t, \tau), \\ \omega_2(s) + \omega_1(\tau) & \text{if } s \in [\tau, T]. \end{cases}$$

The map π^{-1} generates the paths in Ω_t which are continuous at τ. This fact, along with the independence of increments of Lévy processes, implies that

$$P_t = P_{t,\tau} \otimes P_\tau.$$

With all the technical preparations being completed, we can now finally claim that there exists a unique solution $Y_{t,x}(\cdot)$ to the SDE (3.1) for any 4-tuple $(t, x, a(\cdot), \zeta(\cdot)) \in [0, T) \times \mathbb{R}^n \times \mathcal{A}^d(t) \times \mathcal{U}(t)$, i.e.,

$$\begin{aligned} Y_{t,x}(s) = x &+ \int_t^s b(r, Y_{t,x}(r), a(r), \zeta(r))dr + \int_t^s \sigma(r, Y_{t,x}(r), a(r), \zeta(r))dW_r \\ &+ \int_t^s \int_{\mathbb{R}^m\setminus\{0\}} \eta(r, Y_{t,x}(r^-), a(r^-), \zeta(r^-); z)d\tilde{N}(dr, dz) \end{aligned}$$

which also mean that, by the canonical choice of the driving process on Ω_τ for any $\tau \in (t, s]$, we have

$$\begin{aligned} Y_{t,x}(s) = Y_{t,x}(\tau) &+ \int_\tau^s b(r, Y_{t,x}(r), a(r), \zeta(r))dr + \int_\tau^s \sigma(r, Y_{t,x}(r), a(r), \zeta(r))dW'_r \\ &+ \int_\tau^s \int_{\mathbb{R}^m\setminus\{0\}} \eta(r, Y_{t,x}(r^-), a(r^-), \zeta(r^-); z)d\tilde{N}'(dr, dz), \end{aligned}$$

where $(W', \tilde{N}')$ is the canonical driving process on Ω_τ. Thereby arguing along the lines as in [25], we have the following Markov property:

Lemma 3.2 *For any bounded continuous function* φ, *any* $\xi(\cdot) = (a(\cdot), \zeta(\cdot)) \in \mathcal{A}^d(t) \times \mathcal{U}(t)$, *and any* $\tau \in [s, T]$ (*deterministic*),

$$E_{t,x}[\varphi(Y_{t,x}(\tau), \xi(\tau))|\mathcal{F}_{t,s}] = E_{s,Y_{t,x}(s)}[\varphi(Y_{t,x}(\tau), \xi \circ \pi^{-1}(\tau))], \quad P_{t,s} \text{ a.s.} \tag{3.2}$$

3.2 The Dynamic Programming Principle

To derive the dynamic programming principle we need the following continuity properties of the value functions:

Lemma 3.3 *Assumption* (A.1)–(A.4) *hold. Then there exists a constant* $C > 0$, *such that for all* $t \in [0, T]$, $x_1, x_2 \in \mathbb{R}^n$ *and* $d \in \mathcal{A}$,

$$|V^d(x_1)| \le C \quad \text{and} \quad |V^d(t, x_1) - V^d(t, x_2)| \le C|x_1 - x_2|.$$

The proof uses Lipschitz continuity and boundedness of the data, moment estimates for the stochastic processes, and Gronwall's inequality. We do not give the proof here; the proof can be pieced together combining arguments from [38] (controlled jump-diffusions) and [46] (optimal switching for pure diffusions).

Next, we prove the dynamic programming principle. The proof is similar to the one in [46] except that we are working in the Skorohod space (not in C_0).

Theorem 3.4 (Dynamic Programming Principle) *Suppose that assumptions* (A.1)–(A.4) *hold. Let* $V(t,x) = (V^d(t,x))_{d\in\mathcal{A}}$ *be the value function of the optimal switching problem. Then for* $(t,x,d) \in [0,T) \times \mathbb{R}^n \times \mathcal{A}$ *and* $s \in (t,T]$,

$$V^d(t,x) = \inf_{\xi(\cdot)=(a(\cdot),\zeta(\cdot))\in\mathcal{A}^d(t)\times\mathcal{U}(t)} E\left[\int_t^s f(r, Y_{t,x}^{\xi(\cdot)}(r), a(r), \zeta(r))dr + V^{a(s)}(s, Y_{t,x}^{\xi(\cdot)}(s)) + \sum_{\tau_i<s} k(d_i, d_{i-1})\right], \tag{3.3}$$

where $\{\tau_i, d_i\}$ *is the elaborated form of* $a(\cdot)$.

Proof Let the right hand side of (3.3) be $W(t,x)$. Then for every $\epsilon > 0$, there exists $\hat{\xi}(\cdot) = (\hat{a}(\cdot), \hat{\zeta}(\cdot)) \in \mathcal{A}^d(t) \times \mathcal{U}(t)$ such that

$$W(t,x) + \epsilon \geq E\left[V^{\hat{a}(s)}(s, Y_{t,x}^{\hat{\xi}(\cdot)}(s)) + \int_t^s f(r, Y_{t,x}^{\hat{\xi}(\cdot)}(r), \hat{a}(r))dr + \sum_{\hat{\tau}_i<s} k(\hat{d}_i, \hat{d}_{i-1})\right] \tag{3.4}$$

where $\hat{a}(\cdot) = \{\hat{\tau}_i, \hat{d}_i\}$. Moreover, from the definition $V(t,x)$ we have for every $(s,z,b) \in [0,T] \times \mathbb{R}^n \times A$ that there exists $\xi_{s,z,b}(\cdot) \in \mathcal{A}^b[s,T] \times \mathcal{U}[s,T]$ such that

$$V^b(s,z) \geq J_{s,z}^b\big(\xi_{s,z,b}(\cdot)\big) - \epsilon.$$

Next, by uniform continuity in x of V^b uniformly in b (Lemma 3.3) we can choose a partition $\{B_i, i \geq 1\}$ of $\mathbb{R}^n$ such that each of B_i is a Borel set satisfying

$$|V^b(s,x_1) - V^b(s,x_2)| \leq \epsilon, \quad \text{for all } b \in A,\ x_1, x_2 \in B_i.$$

Furthermore, by x-uniform continuity of $J_{t,x}^d(\xi(\cdot))$ uniformly in $\xi(\cdot) \in \mathcal{A}(t) \times \mathcal{U}(t)$ (essentially Lemma 3.3), we may also assume that

$$|J_{t,x_1}^d(\xi(\cdot)) - J_{t,x_2}^d(\xi(\cdot))| \leq \epsilon, \quad \text{for all } d \in A,\ \xi(\cdot) \in \mathcal{A}^d \times \mathcal{U}(t),\ x_1, x_2 \in B_i. \tag{3.5}$$

Now we fix an $\beta_i \in B_i$ for each $i \geq 1$ and define a control $\tilde{\xi}(r) \in \mathcal{A}^d \times \mathcal{U}(t)$ as follows:

$$\tilde{\xi}(r) = \begin{cases} \hat{\xi}(r), & r \in [t,s), \\ \sum_{i\geq 1}\sum_j^m \xi_{s,\beta_i,j}(r)\chi_{B_i}(Y_{t,x}^{\hat{\xi}(\cdot)}(s))\chi_{\{\tilde{a}(s)=j\}}(s), & r \in [s,T]. \end{cases}$$

From the definition we immediately conclude that $\tilde{\xi}(\cdot) \in \mathcal{A}^d \times \mathcal{U}(t)$, and from (3.4)–(3.5) we get

$$\begin{aligned}
J^d_{t,x}(\tilde{\xi}(\cdot)) = E\Bigg[& \int_t^s f(r, Y^{\hat{\xi}(\cdot)}_{t,x}(r), \hat{a}(r))dr + \sum_{\hat{\tau}_j < s} k(\hat{d}_j, \hat{d}_{j-1}) \\
& + \underbrace{\int_s^T f(r, Y^{\tilde{\xi}(\cdot)}_{t,x}(r), \tilde{\xi}(r))dr + \sum_{T \ge \tilde{\tau}_j \ge s} k(\tilde{d}_j, \tilde{d}_{j-1}) + g(Y^{\tilde{\xi}(r)}_{t,x}(T))\Bigg]}_{\sum_i \sum_j J^j_{s, Y^{\hat{\xi}(\cdot)}_{t,x}(s)}(\xi_{s,\beta_i,j})\chi_{B_i}(Y^{\hat{\xi}(\cdot)}_{t,x}(s))\chi_{\hat{a}(s)=j}(s)} \\
\le & E\Bigg[\int_t^s f(r, Y^{\hat{\xi}(\cdot)}_{t,x}, \hat{\xi}(r))dr + \sum_{\hat{\tau}_j < s} k(\hat{d}_j, \hat{d}_{j-1}) + V^{\hat{a}(s)}_{t,x}(Y^{\hat{\xi}(\cdot)}_{t,x}(s)) + 2\varepsilon \Bigg] \\
\le & W(t,x) + 3\epsilon.
\end{aligned}$$

This implies

$$V^d(t,x) \le W(t,x).$$

To get the opposite inequality we argue using the Markov property (3.2). From the definition of $V^d(t,x)$ there exists $\bar{\xi}(\cdot) \in \mathcal{A}^d(t) \times \mathcal{U}(t)$ such that

$$V^d(t,x) + \epsilon \ge J^d_{t,x}(\bar{\xi}(\cdot)). \tag{3.6}$$

We split $J^d_{t,x}(\bar{\xi}(\cdot))$ into two parts, one on $[t,s]$ and one on $(s,T]$. By Lemma 3.2 the second part can be estimated as follows:

$$\begin{aligned}
& E\Bigg\{ \int_s^T f(r, Y^{\bar{\xi}(\cdot)}_{t,x}(r), \bar{\xi}(r))dr + \sum_{T \ge \bar{\tau}_j \ge s} k(\bar{d}_j, \bar{d}_{j-1}) + g(Y^{\bar{\xi}(\cdot)}_{t,x}(T)) \Bigg\} \\
& = E\Bigg\{ E\Bigg[\int_s^T f(r, Y^{\bar{\xi}(\cdot)}_{t,x}(r), \bar{\xi}(r))dr + \sum_{T \ge \bar{\tau}_j \ge s} k(\bar{d}_j, \bar{d}_{j-1}) + g(Y^{\bar{\xi}(\cdot)}_{t,x}(T)) \Bigg| \mathcal{F}_{t,s} \Bigg] \Bigg\} \\
& = E\Big(J^{\bar{\xi}(s)}_{s, Y^{\bar{\xi}(\cdot)}_{t,x}(s)} \Big(\bar{\xi}\big(\cdot; \pi^{-1}(\omega_1, \omega_2)\big) \Big) \Big) \ge E V^{\bar{a}(s)}(s, Y^{\bar{\xi}(\cdot)}_{t,x}(s)).
\end{aligned}$$

By this inequality, the definition of W, and (3.6) we get

$$V^d(t,x) + \epsilon \ge W(t,x),$$

and this completes the proof of the theorem. □

Recall that, for $W : [0,T] \times \mathbb{R}^n \mapsto \mathbb{R}^M$ and $(t,x) \in [0,T] \times \mathbb{R}^n$, the switching operators $\mathcal{M}^d$ are defined as follows

$$\mathcal{M}^d W(t,x) = \min_{d \ne \tilde{d}} \{ W^d(t,x) + k(d, \tilde{d}) \}.$$

As a consequence of the dynamic programming principle we have

Theorem 3.5 *Suppose that assumptions* (A.1)–(A.4) *hold. Then the value function* $V(t,x)$ *satisfies the following properties*,

(i) *For any* $(t,x,d) \in [0,T] \times \mathbb{R}^n \times \mathcal{A}$

$$V^d(t,x) \le \mathcal{M}^d V(t,x). \tag{3.7}$$

(ii) *If for some* $(t,x,d) \in [0,T] \times \mathbb{R}^n \times \mathcal{A}$, (3.7) *fails to hold with an equality, then there exists* $s_0 \in (t,T]$, *such that*

$$V^d(t,x) = \inf_{\zeta(\cdot)\in\mathcal{U}(t)} E\left[\int_t^s f(r, Y_{t,x}(r), d, \zeta(r))dr + V^d(s, Y_{t,x}(s))\right], \tag{3.8}$$

for all $s \in [t, s_0]$ *where* $Y_{t,x}(\cdot)$ *is the solution of* (3.1) *with the control pair* $(d, \zeta(\cdot))$.

The proof of this theorem is very similar to the ones in [19] and [46]. We choose to outline the proof here mainly because of its importance to derive the underlying system of IPDEs.

Proof We prove (i). For every $d, \tilde{d} \in \mathcal{A}$, $d \ne \tilde{d}$ and $\tilde{a} \in \mathcal{A}^{\tilde{d}}$ we define $a(\cdot) \in \mathcal{A}^d$ by

$$\tilde{d}_{i-1} = d_i, \qquad d_0 = d, \qquad \tilde{\tau}_{i-1} = \tau_i, \qquad \tau_0 = t.$$

Note that $\tau_0 = \tau_1 = t$. Let $\tilde{a}(\cdot) = \{\tilde{d}_i, \tilde{\tau}_i\}$ and $a(\cdot) = \{d_i, \tau_i\}$, then

$$V^d(t,x) \le J^d_{t,x}(a(\cdot), \zeta(\cdot)) = J^{\tilde{d}}_{t,x}(\tilde{a}(\cdot), \zeta(\cdot)) + k(d, \tilde{d})$$

for all $\tilde{a}(\cdot) \in \mathcal{A}^{\tilde{d}}$ and $\zeta(\cdot) \in \mathcal{U}(t)$. Hence $V^d(t,x) \le V^{\tilde{d}}(t,x) + k(d,\tilde{d})$ and (i) follows.

To prove (ii) we first observe that if $\{d_i; 1 \le i \le i_0\} \subset \mathcal{A}$ with $i_0 \ge 2$ and $d_i \ne d_{i+1}$ for some $1 \le i \le i_0$ then by (3.7) and the definition of $\mathcal{M}$,

$$\begin{aligned} V^{d_{i_0}}(t,x) + \sum_{i=1}^{i_0-1} k(d_i, d_{i-1}) &\ge \mathcal{M}^{d_{i_0-1}} V(t,x) + \sum_{i=1}^{i_0-2} k(d_i, d_{i-1}) \\ &\ge V^{d_{i_0-1}}(t,x) + \sum_{i=1}^{i_0-2} k(d_i, d_{i-1}) \ge \cdots \\ &\ge \mathcal{M}^{d_1} V(t,x). \end{aligned} \tag{3.9}$$

Next we observe that the inequality "≤" follows from Theorem 3.4 if the controller chooses not to switch, and therefore we only have to prove the "≥" inequality. We use contrapositive argument starting by assuming the contrary: There exists a $\delta > 0$ and sequences $s_p \to t$, $\epsilon_p > 0$ such that

$$V^d(t,x)+\epsilon_p < \inf_{\zeta(\cdot)\in\mathcal{U}(t)} E\bigg[\int_t^{s_p} f(r, Y_{t,x}(r), d, \zeta(\cdot))dr + V^d(s_p, Y_{t,x}(s_p))\bigg], \tag{3.10}$$

$$\mathcal{M}^d V(t,x) - V^d(t,x) = \delta. \tag{3.11}$$

On the other hand, by definition there exists $\xi_p(\cdot) = (a_p(\cdot), \zeta_p(\cdot)) \in \mathcal{A}^d(t) \times \mathcal{U}(t)$ such that,

$$\begin{aligned} V^d(t,x)+\epsilon_p \geq E\bigg[\int_t^{s_p} f(r, Y_{t,x}^{\xi_p(\cdot)}(r), \xi_p(r))dr + V^{a_p(s_p)}(s_p, Y_{t,x}^{\xi_p(\cdot)}(s_p)) \\ + \sum_{t\leq \tau_{p,j}<s_p} k(d_{p,j}, d_{p,j-1})\bigg]. \end{aligned} \tag{3.12}$$

Define $B = \{t \leq \tau_{p,1} < s_p\}$, and note that by (3.10) and (3.12) we must have

$$E\big[\chi_B\big] > 0, \quad \text{and} \quad 0 > I_1 + I_2 + I_3, \tag{3.13}$$

where

$$\begin{aligned} I_1 &:= E\bigg[\int_t^{s_p} \big[f(r, Y_{t,x}^{a_p(\cdot),\zeta_p(\cdot)}(r), a_p(\cdot), \zeta_p(\cdot)) - f(r, Y_{t,x}^{d,\zeta_p(\cdot)}(r), d, \zeta_p(\cdot))\big]dr\bigg] \\ &= o(1)E[\chi_B], \end{aligned} \tag{3.14}$$

$$\begin{aligned} I_2 &:= E\Big[V^{a_p(s_p)}(s_p, Y_{t,x}^{a_p(\cdot),\zeta_p(\cdot)}(s_p)) - V^d(s_p, Y_{t,x}^{d,\zeta_p(\cdot)}(s_p))\Big] \\ &= E\big(V^{a_p(s)}(s_p,x)\chi_B\big) - V^d(s_p,x)E\big(\chi_B\big) + o(1)E\big(\chi_B\big), \end{aligned} \tag{3.15}$$

$$I_3 := E\bigg(\sum_{t\leq \tau_{p,j}<s_p} k(d_{p,j}, d_{p,j-1})\bigg). \tag{3.16}$$

The derivation of (3.14) and (3.15) can be made rigorous using the regularity of V and f, moment estimates on Y and Gronwall's inequality. Now we use (3.13)–(3.16), (3.9), and (3.11) to conclude that as $s_p \to t$,

$$0 > I_1 + I_2 + I_3 \geq \Big[o(1) + \mathcal{M}^d V(s_p,x) - V^d(s_p,x)\Big]E\big(\chi_B\big) \geq [\delta + o(1)]E\big(\chi_B\big),$$

which is a contradiction. □

In view of Theorem 3.5, it is now quite easy to prove that the value function V of the optimal switching problem solves a system of nonlocal quasi-variational inequalities. For every $d \in \mathcal{A}$, $(t,x,r,p_x,X,z) \in [0,T]\times\mathbb{R}^n\times\mathbb{R}^M\times\mathbb{R}^n\times\mathbb{S}^n\times\mathbb{R}^m$ and smooth function φ we define

$$\begin{aligned} &F^d(t,x,r,p,X,\varphi(t,\cdot)) \\ &\quad := \inf_{\alpha\in\mathcal{U}} \bigg\{\frac{1}{2}\mathrm{Tr}\big[\sigma(t,x;d,\alpha)^T X\sigma(t,x;d,\alpha)\big] + b(t,x;d,\alpha).p + f(t,x;d,\alpha) \end{aligned}$$

$$+ \int_{\mathbb{R}^m \setminus \{0\}} \Big\{ \varphi(t, x + \eta(t, x, z; d, \alpha)) - \varphi(t, x) - \eta(t, x, z; d, \alpha).D_x \varphi(t, x) \Big\} \times \nu(dz) \Big\}.$$

We have the following result.

Theorem 3.6 *Suppose that assumptions* (A.1)–(A.4) *hold. Then the value function* $V(t, x)$ *of the optimal switching problem is the unique viscosity solution of the following system of non-local variational inequalities*:

$$\max \Big\{ -\partial_t V^d(t, x) - F^d\big(t, x, V^d(t, x), DV^d(t, x), D^2 V^d(t, x), V^d(t, \cdot)\big); \\ V^d(t, x) - \mathcal{M}^d V(t, x) \Big\} = 0 \quad \textit{in } [0, T) \times \mathbb{R}^n \tag{3.17}$$

with

$$V^d(T, x) := g(x). \tag{3.18}$$

Proof The proof is an immediate consequence of Theorem 3.5 and Dynkin's lemma for Jump-Diffusion processes (see, e.g., [37]). We only prove that $V(t, x)$ is a supersolution, the proof for V being a subsolution is similar. Let $(t, x, d) \in (0, T) \times \mathbb{R}^n \times \mathcal{A}$ and note that if

$$V^d(t, x) = \mathcal{M}^d V(t, x),$$

then (3.17) holds. Otherwise, there exists $s_0 > t$ such that (3.8) holds for all $s \in (t, s_0]$. Let us introduce the following notation:

$$F^d(\alpha; t, x, r, p, X, \varphi(t, \cdot)) \\ := \frac{1}{2} \mathrm{Tr}\big[\sigma(t, x; d, \alpha)^T X \sigma(t, x; d, \alpha)\big] + b(t, x; d, \alpha).p + f(t, x; d, \alpha) \\ + \int_{\mathbb{R}^m \setminus \{0\}} \Big\{ \varphi(t, x + \eta(t, x, z; d, \alpha)) - \varphi(t, x) - \eta(t, x, z; d, \alpha).D_x \varphi(t, x) \Big\} \\ \times \nu(dz).$$

If $\varphi \in C^{1,2}[(0, T) \times \mathbb{R}^n]$, $V^d - \phi$ has a global minima at (t, x), and $V^d(t, x) = \varphi(t, x)$, then by (3.8) and Dynkin's lemma we have

$$V^d(t, x) = \inf_{\zeta(\cdot) \in \mathcal{U}(t)} E\bigg[\int_t^s f(r, Y_{t,x}(r), d, \zeta(r)) dr + V^d(s, Y_{t,x}(s))\bigg], \\ \geq \inf_{\zeta(\cdot) \in \mathcal{U}(t)} E\bigg[\int_t^s f(r, Y_{t,x}(r), d, \zeta(r)) dr + \varphi(s, Y_{t,x}(s))\bigg]$$

$$= \phi(t,x) + \inf_{\zeta(\cdot)\in\mathcal{U}(t)} E\bigg[\int_t^s \varphi_r(r, Y_{t,x}(r)) + F^d\big(\zeta(r); r, Y_{t,x}(r),$$

$$\varphi(r, Y_{t,x}(r)), D\varphi(r, Y_{t,x}(r), D^2\varphi(r, Y_{t,x}(r)), \varphi(r,\cdot))\big)dr\bigg].$$

We may rewrite this inequality as

$$\inf_{\zeta(\cdot)\in\mathcal{U}(t)} E\bigg[\frac{1}{s-t}\int_t^s \varphi_r(r, Y_{t,x}(r)) + F^d\big(\zeta(r); r, Y_{t,x}(r), \varphi(r, Y_{t,x}(r)),$$

$$D\varphi(r, Y_{t,x}(r), D^2\varphi(r, Y_{t,x}(r)), \varphi(r,\cdot))\big)dr\bigg] \le 0,$$

so by letting $s \downarrow t$ and using the moment estimates for solutions of the SDE (3.1), we get

$$\partial_t\varphi(t,x) + \inf_{\zeta(\cdot)\in\mathcal{U}(t)} F^d(\zeta(t+); t, x, \varphi, D\varphi, D^2\varphi, \varphi(t,\cdot)) \le 0,$$

which is equivalent to

$$-\partial_t\varphi(t,x) - F^d(t, x, \varphi(t,x), D\varphi(t,x), D^2\varphi(t,x), \varphi(t,\cdot)) \ge 0.$$

Hence V is a supersolution of (3.17). □

Remark The system of variational inequalities (3.17)–(3.18) is a terminal value problem, which easily can be converted to an initial value problem. Once we do that, any result derived for (1.1) applies to the above system as well. Therefore, the system (3.17)–(3.18) has a unique solution which is the value function $V(t,x)$ and satisfies the regularity estimate in Theorem 2.5.

4 Comparison Principle, Perron's Method, and Existence of Solutions

We start this section with the proof of the comparison principle, cf. Theorem 2.2. This result is the backbone of any viscosity solution theory. The basic idea of the proof is same as the pure PDE case, i.e., to reduce the problem to the scalar case using a no-loop argument and then follow the usual approach to get the final result. In the proof we will need the so-called maximum principle for semicontinuous functions, suitably adapted to the nonlocal system. This result along with the no-loop argument is summarized in the following lemma:

Lemma 4.1 *Let $u \in USC([0,T]\times\mathbb{R}^n;\mathbb{R}^M)$ be a subsolution of (1.1) and $\hat{u} \in LSC([0,T]\times\mathbb{R}^n;\mathbb{R}^M)$ be a supersolution of another variant of (1.1) (for example, the system (5.1)) where the operators $\mathcal{L}_i^{\alpha,\beta}$, $\mathcal{J}_i^{\alpha,\beta}$ are respectively replaced by $\hat{\mathcal{L}}_i^{\alpha,\beta}$, $\hat{\mathcal{J}}_i^{\alpha,\beta}$ satisfying the same assumptions. Let $\phi(t,x,y) \in C_p^{1,2}[0,T]\times\mathbb{R}^{2n}$ be bounded from below and denote*

$$\Psi_i(t,x,y) := u^i(t,x) - \hat{u}^i(t,y) - \phi(t,x,y).$$

If $\mathcal{D} := \sup_{i,t,x,y} \Psi_i(t,x,y)$ *exists finitely and there is a (maximum) point* $(i', t_0, x_0, y_0) \in \mathcal{I} \times (0,T) \times \mathbb{R}^{2n}$ *such that* $\psi_{i'}(t_0,x_0,y_0) = \mathcal{D}$, *then there exists* $i_0 \in \mathcal{I}$ *such that* $\mathcal{D} := \Psi_{i_0}(t_0,x_0,y_0)$ *and* $\hat{u}^{i_0}(t_0,x_0) < \mathcal{M}^{i_0}\hat{u}(t_0,x_0)$.

Furthermore, if in neighborhood of (t_0,x_0,y_0), *there are continuous functions* $h_0 : [0,T] \times \mathbb{R}^{2n} \to \mathbb{R}$, $h, \hat{h} : \bar{Q}_T \to \mathbb{S}^n$ *such that* $h_0(t_0,x_0,y_0) > 0$ *and*

$$D^2\phi \le h_0(t,x,y)\begin{pmatrix} I & -I \\ -I & I \end{pmatrix} + \begin{pmatrix} h(t,x) & 0 \\ 0 & \hat{h}(t,y) \end{pmatrix},$$

then for each $\kappa \in (0,1)$ *there are* $a, b \in \mathbb{R}$ *and* $X, Y \in \mathbb{S}^n$ *satisfying*

$$a - b = \phi_t(t_0,x_0,y_0)$$

and

$$\begin{pmatrix} X & 0 \\ 0 & -Y \end{pmatrix} \le 2h_0(t_0,x_0,y_0)\begin{pmatrix} I & -I \\ -I & I \end{pmatrix} + \begin{pmatrix} h(t_0,x_0) & 0 \\ 0 & \hat{h}(t_0,y_0) \end{pmatrix}$$

such that

$$\begin{aligned} a + \sup_{\alpha\in\mathcal{A}_{i_0}} \inf_{\beta\in\mathcal{B}_{i_0}} \big[&\mathcal{L}_{i_0}^{\alpha,\beta}(t_0,x_0,u^{i_0},D_x\phi(t_0,x_0,y_0),X) \\ &- \mathcal{J}_{i_0,\kappa}^{\alpha,\beta}(t_0,x_0,D_x\phi(t_0,x_0,y_0),\phi(t_0,\cdot,y_0)) \\ &- \mathcal{J}_{i_0}^{\alpha,\beta,\kappa}(t_0,x_0,D_x\phi(t_0,x_0,y_0),u^{i_0}(t_0,\cdot))\big] \le 0, \\ b + \sup_{\alpha\in\mathcal{A}_{i_0}} \inf_{\beta\in\mathcal{B}_{i_0}} \big[&\hat{\mathcal{L}}_{i_0}^{\alpha,\beta}(t_0,y_0,\hat{u}^{i_0},-D_y\phi(t_0,x_0,y_0),Y) \\ &- \hat{\mathcal{J}}_{i_0,\kappa}^{\alpha,\beta}(t_0,y_0,-D_y\phi(t_0,x_0,y_0),-\phi(t_0,x_0,\cdot)) \\ &- \hat{\mathcal{J}}_{i_0}^{\alpha,\beta,\kappa}(t_0,y_0,-D_y\phi(t_0,x_0,y_0),\hat{u}^{i_0}(t_0,\cdot))\big] \ge 0. \end{aligned}$$

The first part of the above Lemma follows exactly in the same way as Lemma A.2 in [9]. Once we have the first part, then for the supersolution it says that at the point (t_0,x_0,y_0) we can ignore the term $\hat{u}^{i_0} - \mathcal{M}^{i_0}\hat{u}$, and then the second part follows as a consequence of Theorem 2.2 of [31].

Proof of Theorem 2.2 For constants $\lambda, \theta, \gamma, \epsilon > 0$ we define the following (test) function:

$$\phi(t,x,y) = e^{\lambda t}\frac{\theta}{2}|x-y|^2 + e^{\lambda t}\frac{\epsilon}{2+\gamma}\big(|x|^{2+\gamma} + |y|^{2+\gamma}\big) \tag{4.1}$$

on $[0,T] \times \mathbb{R}^n \times \mathbb{R}^n$. We double the variables defining for $i \in \mathcal{I}$,

$$\Psi_i(t,x,y) = u^i(t,x) - v^i(t,y) - \phi(t,x,y) - \frac{\delta\sigma t}{T} - \frac{\bar{\epsilon}}{T-t},$$

where $0<\delta<1$, $\bar{\epsilon}>0$, and

$$\sigma_0 = \sup_{i,x,y}\left\{u^i(0,x) - v^i(0,y) - \phi(0,x,y) - \frac{\bar{\epsilon}}{T}\right\}^+,$$

$$\sigma = \sup_{i,t,x,y}\left\{u^i(t,x) - v^i(t,y) - \phi(t,x,y) - \frac{\bar{\epsilon}}{T-t}\right\} - \sigma_0.$$

The main step of this proof is to derive an upper bound on $\sigma+\sigma_0$ by deriving a positive upper bound on σ. Note that if $\sigma\le 0$ then we can take 0 as the upper bound and we are done; therefore we will assume in the following that $\sigma>0$. By the upper semicontinuity of $u^i - v^i$, the growth assumptions, and the penalization term, there exists $(i_0,t_0,x_0,y_0)\in\mathcal{I}\times[0,T)\times\mathbb{R}^n\times\mathbb{R}^n$ such that

$$\Psi_{i_0}(t_0,x_0,y_0) = \sup_{i,t,x,y}\Psi_i(t,x,y).$$

The assumption that $\sigma>0$ forces $t_0\neq 0$, so that $0<t_0<T$. Since,

$$\begin{aligned}\Psi_{i_0}(t_0,x_0,y_0) &\ge \sup_{i,t,x,y}\left\{u^i(t,x) - v^i(t,y) - \phi(t,x,y) - \frac{\bar{\epsilon}}{T-t}\right\} - \delta\sigma\\ &= \sigma_0 + (1-\delta)\sigma > \sigma_0,\end{aligned}$$

while on the other hand $t_0=0$ would imply $\Psi_{i_0}(t_0,x_0,y_0)\le\sigma_0$.

Now we are in a position apply the maximum principle for semicontinuous functions adapted to the present non-local system, i.e. Lemma 4.1. By this lemma, for each $0<\kappa\le 1$ there are numbers p_t and q_t, symmetric matrices X and Y, and an index i_0 such that

$$p_t - q_t = \frac{\delta\sigma}{T} + \frac{\bar{\epsilon}}{(T-t_0)^2} + \phi_t(t_0,x_0,y_0)$$

and

$$\begin{aligned}p_t - q_t \le \sup_{\alpha\in\mathcal{A}_{i_0}}\inf_{\beta\in\mathcal{B}_{i_0}}&\big[\mathcal{L}^{\alpha,\beta}_{i_0}(t_0,y_0,v^{i_0},-D_y\phi(t_0,x_0,y_0),Y)\\ &-\mathcal{J}^{\alpha,\beta}_{i_0,\kappa}(t_0,y_0,-D_y\phi(t_0,x_0,y_0),-\phi(t_0,x_0,\cdot))\\ &-\mathcal{J}^{\alpha,\beta,\kappa}_{i_0}(t_0,y_0,-D_y\phi(t_0,x_0,y_0),v^{i_0}(t_0,\cdot))\big]\\ &-\sup_{\alpha\in\mathcal{A}_{i_0}}\inf_{\beta\in\mathcal{B}_{i_0}}\big[\mathcal{L}^{\alpha,\beta}_{i_0}(t_0,x_0,u^{i_0},D_x\phi(t_0,x_0,y_0),X)\\ &-\mathcal{J}^{\alpha,\beta}_{i_0,\kappa}(t_0,x_0,D_x\phi(t_0,x_0,y_0),\phi(t_0,x_0,\cdot))\\ &-\mathcal{J}^{\alpha,\beta,\kappa}_{i_0}(t_0,x_0,-D_x\phi(t_0,x_0,y_0),v^{i_0}(t_0,\cdot))\big]\end{aligned} \tag{4.2}$$

with

$$\begin{pmatrix}X & 0\\ 0 & -Y\end{pmatrix} \le 2e^{\lambda t_0}\theta\begin{pmatrix}I & -I\\ -I & I\end{pmatrix} + \epsilon e^{\lambda t_0}(1+\gamma)\begin{pmatrix}|x_0|^\gamma I & 0\\ 0 & |y_0|^\gamma I\end{pmatrix}. \tag{4.3}$$

The upper bound on σ will be obtained from (4.2). We start by estimating the right hand side of (4.2). First note that

$$\begin{aligned}
&\mathcal{L}_{i_0}(t_0, x_0, u^{i_0}, D_x\phi(t_0, x_0, y_0), X) - \mathcal{L}_{i_0}(t_0, y_0, v^{i_0}, -D_y\phi(t_0, x_0, y_0), Y)\\
&\quad = \big[\mathrm{Tr}(a_{i_0}(x_0, t_0)X - a_{i_0}(x_0, t_0)Y)\big]\\
&\qquad + \big[b_{i_0} D_x\phi(t_0, x_0, y_0) + b_{i_0}(t_0, y_0) D_y\phi(t_0, x_0, y_0)\big]\\
&\qquad + \big[c_{i_0}(t_0, y_0)v^{i_0}(t_0, y_0) - c_{i_0}(t_0, x_0)u^{i_0}(t_0, x_0)\big] + \big[f_{i_0}(t_0, y_0) - f_{i_0}(t_0, x_0)\big],
\end{aligned}$$

and

$$\begin{aligned}
D_y\phi(t_0, x_0, y_0) &= -\theta e^{\lambda t_0}(x_0 - y_0) + \epsilon e^{\lambda t_0} y_0|y_0|^\gamma,\\
D_y\phi(t_0, x_0, y_0) &= \theta e^{\lambda t_0}(x_0 - y_0) + \epsilon e^{\lambda t_0} x_0|x_0|^\gamma.
\end{aligned}$$

By the definition of ϕ, inequality (4.3), assumptions (A.1), (A.3), and standard computations,

$$\mathrm{Tr}[a_{i_0}(x_0, t_0)X - a_{i_0}(x_0, t_0)Y] \le K_1 e^{\lambda t_0}\big\{\theta|x_0 - y_0|^2 + \epsilon(1 + |x_0|^{2+\gamma} + |y_0|^{2+\gamma})\big\}. \tag{4.4}$$

Note that $\sigma > 0$ implies that $u^{i_0}(t_0, x_0) - v^{i_0}(t_0, y_0) > 0$, so by (A.2), (A.3), and the growth assumptions on u, v we easily see that

$$\begin{aligned}
&b_{i_0}(t_0, x_0)D_x\phi(t_0, x_0, y_0) + b_{i_0}(t_0, y_0)D_y\phi(t_0, x_0, y_0)\\
&\quad \le K_2 e^{\lambda t_0}(\theta|x_0 - y_0|^2 + \epsilon(1 + |x_0|^{2+\gamma} + |y_0|^{2+\gamma})), && (4.5)\\
&c_{i_0}(t_0, y_0)v^{i_0}(t_0, y_0) - c_{i_0}(t_0, x_0)u^{i_0}(t_0, x_0) \le K_3(1 + |x_0|^2 + |y_0|^2)|x_0 - y_0|, && (4.6)\\
&f_{i_0}(t_0, y_0) - f_{i_0}(t_0, x_0) \le K_4|x_0 - y_0|. && (4.7)
\end{aligned}$$

By (A.1) and (A.4) it follows that

$$\begin{aligned}
&\mathcal{J}_{i_0,k}(t_0, x_0, D_x\phi(t_0, x_0, y_0); \phi(t_0, \cdot, y_0))\\
&\quad - \mathcal{J}_{i_0,k}(t_0, y_0, -D_y\phi(t_0, x_0, y_0); -\phi(t_0, x_0, \cdot)) = \mathcal{O}(\kappa).
\end{aligned} \tag{4.8}$$

Using the fact that (t_0, x_0, y_0) is a maximum point of Ψ_{i_0} we have,

$$\begin{aligned}
&\int_{\kappa\le|z|\le 1}\Big[u^{i_0}(t_0, x_0 + \eta(t_0, x_0, z)) - u^{i_0}(t_0, x_0) - \eta(t_0, x_0, z)D_x\phi(t_0, x_0, y_0)\\
&\qquad - v^{i_0}(t_0, y_0 + \eta(t_0, y_0, z)) + v^{i_0}(t_0, y_0) - \eta(t_0, y_0, z)D_y\phi(t_0, x_0, y_0)\Big]\nu(dz)\\
&\quad \le \int_{\kappa\le|z|\le 1}\Big[\phi(t_0, x_0 + \eta(t_0, x_0, z), y_0 + \eta(t_0, y_0, z)) - \phi(t_0, x_0, y_0)\\
&\qquad - (x_0 + \eta(t_0, x_0, z), y_0 + \eta(t_0, y_0, z))D_{x,y}\phi(t_0, x_0, y_0)\Big]\nu(dz)\\
&\quad \le K_6 e^{\lambda t_0}\big\{\epsilon(1 + |x_0|^{2+\gamma} + |y_0|^{2+\gamma}) + \theta|x_0 - y_0|^2\big\}.
\end{aligned} \tag{4.9}$$

In a similar manner we also have

$$\int_{1\le|z|}\Big[u^{i_0}(t_0,x_0+\eta(t_0,x_0,z))-u^{i_0}(t_0,x_0)$$
$$-v^{i_0}(t_0,y_0+\eta(t_0,y_0,z))+v^{i_0}(t_0,y_0)\Big]\nu(dz)$$
$$\le K_7e^{\lambda t_0}\Big[\theta|x_0-y_0|^2+\epsilon(1+|x_0|^{2+\gamma}+|y_0|^{2+\gamma})\Big]. \tag{4.10}$$

Combining the different above estimates (4.4)–(4.10) and using (4.2) we have,

$$\frac{\delta\sigma}{T}+\lambda e^{\lambda t_0}\Big(\theta|x_0-y_0|^2+\frac{\epsilon}{2+\gamma}(|x_0|^{2+\gamma}+|y_0|^{2+\gamma})\Big)$$
$$\le C_1e^{\lambda t_0}\Big(\theta|x_0-y_0|^2+\epsilon(1+|x_0|^{2+\gamma}+|y_0|^{2+\gamma})\Big)$$
$$+C_2(1+|x_0|^2+|y_0|^2)|x_0-y_0|+\mathcal{O}(\kappa).$$

Notice that the point (t_0,x_0,y_0) does not depend on κ, so after letting $\kappa\to 0$ and rearranging the terms with the choice $\lambda=\max\{(2+\gamma)C_1,2C_2\}+2$ we get

$$\frac{\delta\sigma}{T}\le C_3(1+|x_0|+|y_0|)^2|x_0-y_0|-e^{\lambda t_0}\theta|x_0-y_0|^2$$
$$-C_4\epsilon e^{\lambda t_0}(1+|x_0|+|y_0|)^{2+\gamma}+\mathcal{O}(\epsilon).$$

After a maximization on the right-hand side of the above inequality with respect to $|x_0-y_0|$ we obtain

$$\frac{\delta\sigma}{T}\le C_5\frac{(1+|x_0|+|y_0|)^4}{e^{\lambda t_0}\theta}-C_4\epsilon e^{\lambda t_0}(1+|x_0|+|y_0|)^{2+\gamma}+\mathcal{O}(\epsilon).$$

Now choose $\gamma=6$ and maximize the right-hand side of the above inequality with respect to $(1+|x_0|+|y_0|)$ and let $\delta\to 1$; the result is

$$\sigma\le TC\frac{1}{\theta^2\epsilon}+\mathcal{O}(\epsilon).$$

We estimate σ_0 using the Lipschitz continuity of $u(0,x)$ and $v(0,x)$,

$$\sigma_0\le\max_i|(u^i(0,\cdot)-v^i(0,\cdot))^+|_0+\sup_{x,y}\Big(K|x-y|-\frac{\theta}{2}|x-y|^2\Big)$$
$$=\frac{K^2}{2\theta}, \tag{4.11}$$

where we have also used the fact that $u^i(0,\cdot)\le v^i(0,\cdot)$ and maximization of with respect to $|x-y|$ in the first line. Therefore, for all $t\in[0,T)$, $x\in\mathbb{R}^n$, and $i\in\mathcal{I}$,

$$u^i(t,x)-v^i(t,x)-\frac{\bar\epsilon}{T-t}-\frac{1}{4}\epsilon|x|^8\le\sigma+\sigma_0\le TC\frac{1}{\theta^2\epsilon}+\mathcal{O}(\epsilon)+\frac{K^2}{2\theta},$$

and letting $\bar{\epsilon} \to 0$, $\theta \to \infty$ for a fixed ϵ gives

$$u^i(t,x) - v^i(t,x) - \frac{1}{4}\epsilon|x|^8 \le \mathcal{O}(\epsilon).$$

Finally, letting $\epsilon \to 0$ concludes the theorem. □

Now we turn to the existence of viscosity solutions of the system of IPDEs (1.1); we will use Perron's method as developed by Ishii [26] and its adaptation to the scalar nonlocal equations by Alvarez & Tourin [1]. Different from [1], we face a system of equations and an unbounded Lévy measure ν.

Proof of Theorem 2.3 We only prove existence since uniqueness follows from the comparison principle. Define $v(t,x) = (v^1, v^2, \dots, v^M)$ as

$$v^i(t,x) = \sup\big\{u^i(t,x) : u = (u^1, u^2, \dots, u^i, \dots, u^M) \text{ is a subsolution of (1.1)}\big\}$$

for each $i \in \mathcal{I}$. Next, let v^* and v_* denote the upper and lower semi-continuous envelopes of $v(t,x)$:

$$v^{*,i}(t,x) = \lim_{r\downarrow 0} \sup\{v^i(s,y) : (s,y) \in B_r(t,x) \cap [0,T) \times \mathbb{R}^n\},$$

and $v^i_*(t,x) = -(-v^i(t,x))^*$. From the definition it is clear that

$$\bar{u} \le v^*, \qquad v_* \le \bar{v}, \quad \text{and} \quad v_* \le v^*.$$

We want to show that v^* and v_* are respectively sub- and supersolutions of (1.1). Then we are done, since by the comparison principle

$$v^* \le v_*,$$

and hence $v^* = v_* = v$ is the sought after (continuous) viscosity solution of (1.1).

We now prove that v^* is subsolution of (1.1). First, we check that the initial condition is satisfied using a barrier argument. For every $z \in \mathbb{R}^n$ and $\epsilon > 0$, define

$$\Psi^i_{z,\epsilon}(x) = g^i(z) + L^i\big(|x-z|^2 + \epsilon\big)^{\frac{1}{2}},$$

where L^i is the Lipschitz constant of $g_i(x)$. It follows that

$$\Psi^i_{z,\epsilon}(x) \ge g_i(x) \quad \text{for all } x, z \in \mathbb{R}^n,\ i \in \mathcal{I},\ \epsilon > 0.$$

A simple computation now shows that there is a constant $A_\epsilon \ge 0$ such that

$$U^i_{z,\epsilon}(t,x) := A_\epsilon t + \Psi^i_{z,\epsilon}(x)$$

is a continuous supersolution to (1.1). Therefore, by the comparison principle,

$$v^i(t,x) \le U^i_{z,\epsilon}(t,x) \quad \text{for all } x, z \in \mathbb{R}^n,\ i \in \mathcal{I},\ \epsilon > 0,$$

and hence $v^{*,i}(t,x) \le (U^i_{z,\epsilon})^*(t,x) = U^i_{z,\epsilon}(t,x)$. So the initial condition follows after setting $t = 0$ and minimizing w.r.t. z, ε:

$$v^{*,i}(0,x) \le \inf_{\epsilon,z} U^i_{z,\epsilon}(0,x) = \inf_{\epsilon,z} \Psi^i_{z,\epsilon}(x) = g_i(x).$$

Next, we want to show that subsolution condition for the system of equations holds. For each $i \in \mathcal{I}$ and $(t,x) \in (0,T) \times \mathbb{R}^n$ there exists a sequence $(t_p, x_p, u_p(t_p, x_p))$ such that

$$\lim_{p\to\infty} \big(t_p, x_p, u^i_p(t_p, x_p)\big) = (t, x, v^{*,i}(t,x)),$$

and u_p is a subsolution for each $p \in \mathbb{N}$. Now if $\phi \in C^{1,2}$ and $v^{*,i} - \phi$ has a strict global maximum at $(t,x) \in [0,T) \times \mathbb{R}^n$, then there will be a sequence (s_p, y_p) of global maxima of $u^i_p - \phi$ (for p large enough) such that

$$\lim_{p\to\infty} \big(s_p, y_p, u^i_p(s_p, y_p)\big) = \big(t, x, v^{*,i}(t,x)\big).$$

Again if p is large enough, $s_p > 0$ and the definition of subsolution gives

$$\max\Big(\phi_t(s_p, y_p) + \sup_{\alpha\in\mathcal{A}_i} \inf_{\beta\in\mathcal{B}_i} \big\{\mathcal{L}^{\alpha,\beta}_i(s_p, y_p, u^i_p(s_p, y_p), D\phi(s_p, y_p), D^2\phi(s_p, y_p)) \\ - \mathcal{J}^{\alpha,\beta}_i \phi(s_p, \cdot)\big\}; u^i_p(s_p, y_p) - \mathcal{M}^i u_p(s_p, y_p)\Big) \le 0.$$

Passing to the limit $p \to \infty$ and using the regularity of ϕ, v^* and the continuity of the equation, we get

$$\max\Big(\phi_t(t,x) + \sup_{\alpha\in\mathcal{A}_i} \inf_{\beta\in\mathcal{B}_i} \big\{\mathcal{L}^{\alpha,\beta}_i(t, x, v^{*,i}(t,x), D\phi(t,x), D^2\phi(t,x)) \\ - \mathcal{J}^{\alpha,\beta}_i \phi(t, \cdot)\big\}, v^{*,i}(t,x) - \mathcal{M}^i v^*(t,x)\Big) \le 0.$$

This completes the proof that v^* is a subsolution.

Next we prove that v_* is a supersolution of (1.1). We start by checking the initial condition. For $z \in \mathbb{R}^n$ and $\epsilon > 0$, let

$$\Phi^i_{z,\epsilon}(x) = g^i(z) - L\big(|x - z|^2 + \epsilon\big)^{\frac{1}{2}} \quad \text{and} \quad V^i_{\epsilon,z}(t,x) = -A^\epsilon t + \Phi^i_{z,\epsilon}(x),$$

where $L = \max_i\{L_i\}$ and A^ϵ is a constant to be determined later. Note that

$$\Phi^i_{z,\epsilon}(x) \le g_i(x) \quad \text{for all } x, z, \epsilon,$$

and since $g_i - \mathcal{M}^i g \le 0$ by assumption (A.2), we see that

$$V^i_{\epsilon,z}(t,x) - \mathcal{M}^i V_{\epsilon,z}(t,x) \le 0.$$

Now it is straightforward to see that there is a constant A^ϵ such that $V_{\epsilon,z}$ is a subsolution to (1.1). Therefore, by the definition of $v(t,x)$,

$$V^i_{\epsilon,z}(t,x) \le v^i(t,x) \quad \text{for all } t, x, z, \varepsilon.$$

It follows that $V^i_{\epsilon,z}(t,x) \le v^i_*(t,x)$ and hence the initial condition holds because

$$v^i_*(0,x) \ge \sup_{\epsilon,z} V^i_{\epsilon,z}(0,x) = \sup_{\epsilon,z} \Phi^i_{z,\epsilon}(x) = g_i(x).$$

We continue with proving that the system of equations is satisfied. Assume by contradiction that v_* is not a supersolution. Then there are $(i,t,x) \in \mathcal{I} \times (0,T) \times \mathbb{R}^n$ and $\phi \in C^{1,2}_p$ satisfying $v^i_* = \phi$ at (t,x), $v^i_* - \phi$ has a global strict minimum at (t,x), and

$$\begin{aligned} \max\big\{\phi_t + \sup_{\alpha\in\mathcal{A}_i} \inf_{\beta\in\mathcal{B}_i} \big\{\mathcal{L}^{\alpha,\beta}_i(t,x,v^i_*(t,x), D\phi(t,x), D^2\phi(t,x)) \\ - \mathcal{J}^{\alpha,\beta}_i \phi(t,\cdot)\big\}, v^i_*(t,x) - \mathcal{M}^i v_*(t,x)\big\} < 0. \end{aligned} \tag{4.12}$$

Let us prove that $v^i_*(t,x) < \bar{v}^i(t,x)$. By the definition of v_*, $v^i_*(t,x) \le \bar{v}^i(t,x)$, so if by contradiction this equality is not strict, then $\phi(t,x) = v^i_*(t,x) = \bar{v}^i(t,x)$. But then $\bar{v}^i - \phi$ has a global minimum at (t,x), and since $\bar{v}$ is a supersolution,

$$\begin{aligned} \max\big\{\phi_t(t,x) + \sup_{\alpha\in\mathcal{A}_i} \inf_{\beta\in\mathcal{B}_i} \big\{\mathcal{L}^{\alpha,\beta}_i(t,x,\bar{v}^i(t,x), D\phi(t,x), D^2\phi(t,x)) \\ - \mathcal{J}^{\alpha,\beta}_i \phi(t,\cdot)\big\}, \bar{v}^i(t,x) - \mathcal{M}^i \bar{v}(t,x)\big\} \ge 0. \end{aligned}$$

Now $\bar{v}^i(t,x) = v^i_*(t,x)$ and $-\mathcal{M}^i \bar{v}(t,x) \le -\mathcal{M}^i v_*(t,x)$, so this is a contradiction to (4.12) and the inequality is strict. By continuity of $\phi, \bar{v}$ it immediately follows that there are constants $\epsilon_1, \delta_1 \ge 0$ such that

$$\phi + \epsilon_1 \le \bar{v} \quad \text{in } B_{\delta_1}(t,x) \subset Q_T.$$

Therefore, by (4.12), continuity of the equation, regularity of ϕ and lower semicontinuity of v_*, there exist two constants $\epsilon_2, \delta_2 \ge 0$ such that

$$\begin{aligned} \max\big\{&(\phi+\epsilon)_t \\ &+ \sup_{\alpha\in\mathcal{A}_i} \inf_{\beta\in\mathcal{B}_i} \big\{\mathcal{L}^{\alpha,\beta}_i(s,y,(\phi+\epsilon)(s,y), D(\phi+\epsilon)(s,y), D^2(\phi+\epsilon)(s,y)) \\ &- \mathcal{J}^{\alpha,\beta}_i(\phi+\epsilon)(t,\cdot)\big\}, (\phi(s,y)+\epsilon) - \mathcal{M}^i v_*(s,y)\big\} \le 0 \end{aligned} \tag{4.13}$$

for all $(s,y) \in B_{\delta_2}(t,x) \subset Q_T$ and $0 \le \epsilon \le \epsilon_2$.

Since (t,x) is a strict minimum point of $v^i_* - \phi$, there are constants $\epsilon_3 \ge 0$ and $\delta_0 \le \min(\delta_1,\delta_2)$ such that $v^i_* - \phi > \epsilon_3$ on $\partial B_{\delta_0}(t,x)$. Now set $\epsilon_0 = \min(\epsilon_1,\epsilon_2,\epsilon_3)$ and define

$$w^j = v^{*,j} \quad \text{if } j \ne i, \qquad w^i = \begin{cases} \max(\phi+\epsilon_0, v^{*,i}) & \text{on } B_{\delta_0}(t,x) \cap \bar{Q}_T, \\ v^{*,i} & \text{elsewhere.} \end{cases}$$

Note that w is upper semicontinuous. We will prove that w is a subsolution of (1.1). For $j \neq i$,

$$w_j - \mathcal{M}^j w = v^{*,j} - \mathcal{M}^j w \leq v^{*,j} - \mathcal{M}^j v^* \leq 0,$$

and the subsolution inequalities hold as in the first part of the proof. For $j = i$ and $(s, y) \in B_{\delta_0}(t, x) \cap \bar{Q}_T$,

$$\begin{aligned} w^i(s, y) - \mathcal{M}^i w(s, y) &\leq \max\big\{\phi(s, y) + \epsilon_0 - \mathcal{M}^i v^*, v^{i,*}(s, y) - \mathcal{M}^i v^*(s, y)\big\} \\ &\leq \max\big\{\phi(s, y) + \epsilon_0 - \mathcal{M}^i v_*(s, y), v^{i,*}(s, y) - \mathcal{M}^i v^*(s, y)\big\} \\ &\leq 0, \end{aligned}$$

where we have used (4.13) and the fact that v^* is a subsolution. Outside the region $B_{\delta_0}(t, x) \cap \bar{Q}_T$, it trivially holds that $w^i(s, y) - \mathcal{M}^i w(s, y) \leq 0$. Take $(s, y) \in [0, T) \times \mathbb{R}^N$ and $\psi \in C_p^{1,2}$ such that $w^i(s, y) = \psi(s, y)$ and $w^i - \psi$ has a strict a global maximum at (s, y). Depending on whether $w^i = v^{*,i}$ or $\phi + \epsilon_0 = w^i$ at (s, y), either $v_*^i - \psi$ or $\phi + \epsilon_0 - \psi$ has a global maximum here. In the first case, the subsolution inequality involving the test function ψ is a consequence of v^* being a subsolution. In the other case,

$$\begin{aligned} &\partial_t \phi(s, y) \geq \partial_t \psi(s, y), \qquad D\phi(s, y) = D\psi(s, y), \qquad D^2\phi(s, y) \leq D^2\psi(s, y), \\ &\phi(s, y + \eta^{\alpha,\beta}(s, y, z)) - \phi(s, y) \leq \psi(s, y + \eta^{\alpha,\beta}(s, y, z)) - \psi(s, y), \end{aligned}$$

and hence (4.13) implies that

$$\psi_t(s, y) + \sup_{\alpha \in \mathcal{A}_i} \inf_{\beta \in \mathcal{B}_i} \big\{\mathcal{L}_i^{\alpha,\beta}(s, y, \psi(s, y), D\psi(s, y), D^2\psi(s, y)) - \mathcal{J}_i^{\alpha,\beta}\psi(s, \cdot)\big\} \leq 0.$$

This completes the proof that w is a viscosity subsolution of (1.1).

We can now conclude the proof since w is a subsolution satisfying

$$w_*^i(t, x) \geq \sup\{\phi(t, x) + \epsilon_0, v_*^i(t, x)\} = \phi(t, x) + \epsilon_0 \geq v_*^i(t, x) + \epsilon_0,$$

i.e., $w^i(s, y) > v^i(s, y)$ for some (s, y) thereby contradicting the definition of v. □

5 Continuous Dependence Estimate and Regularity Properties

In this section

$$\hat{\sigma}_i^{\alpha,\beta}(t, x), \quad \hat{b}_i^{\alpha,\beta}(t, x), \quad \hat{c}_i^{\alpha,\beta}(t, x), \quad \hat{f}_i^{\alpha,\beta}(t, x), \quad \hat{\eta}_i^{\alpha,\beta}(t, x, z), \quad \hat{\nu}(dz),$$

will denote another set of coefficients/Lévy measure satisfying assumptions (A.1)–(A.4). We define the operators $\hat{\mathcal{L}}_i^{\alpha,\beta}$ and $\hat{\mathcal{J}}_i^{\alpha,\beta}$ in the obvious way, and consider the new initial value problem

$$\begin{aligned}&\max\Big[\partial_t u^i(t,x)+\sup_{\alpha\in\mathcal{A}_i}\inf_{\beta\in\mathcal{B}_i}\big\{\hat{\mathcal{L}}_i^{\alpha,\beta}(t,x,u^i(t,x),Du^i(t,x),D^2u^i(t,x))\\&\quad-\hat{\mathcal{J}}_i^{\alpha,\beta}(u^i)(t,\cdot)\big\};u^i-\mathcal{M}^i u\Big]=0\quad\text{in } Q_T,\\&u^i(0,x)=\hat{g}_i(x)\quad\text{in }\mathbb{R}^N,\end{aligned}\tag{5.1}$$

and $\hat{g}=(\hat{g}_i(x))_i$ satisfies (A.2).

The objective is to estimate the difference between the viscosity solutions of (1.1) and (5.1) in terms of the difference between the "nonlinearities" and the initial conditions. Such continuous dependence estimates are important in themselves, as they quantify the stability properties of viscosity solutions, and have many important consequences and uses. One immediate consequence is Lipschitz continuity in the spatial variable of a viscosity solution, and with some additional reasoning also Hölder continuity in time. Another (recent) application concerns their relevance in Krylov's method of shaking the coefficients, which is used in numerical analysis of convex fully-nonlinear PDEs, see for example [10, 35].

Let us now state the continuous dependence estimate.

Theorem 5.1 *Suppose that* (A.1)–(A.4) *hold for both sets of coefficients. Let* $u,-\hat{u}\in USC_b(\bar{Q}_T;\mathbb{R}^M)$ *be respectively sub- and supersolutions of* (1.1) *and* (5.1) *satisfying*

$$|Du^i(0,x)|\le K,\quad |D\hat{u}^i(0,x)|\le K\quad \textit{for all } i\in\mathcal{I}.$$

Then there exists a constant C, *depending on the data*, *such that for all* $j\in\mathcal{I}$,

$$\begin{aligned}u^j-\hat{u}^j\le \max_i\Bigg[&|(u^i-\hat{u}^i)^+(0,\cdot)|_0+T\sup_{\alpha,\beta}\big(|f_i-\hat{f}_i|_0+|u|_0\vee|\hat{u}|_0|c_i-\hat{c}_i|_0\big)\\&+CT^{\frac12}\sup_{\alpha,\beta}\Bigg\{|\sigma_i-\hat{\sigma}_i|_0+|b_i-\hat{b}_i|_0+\bigg|\int|\bar{\eta}_i|^2|\nu-\hat{\nu}|(dz)\bigg|_0^{\frac12}\\&+\bigg|\int|\eta_i-\hat{\eta}_i|^2\bar{\nu}(dz)\bigg|_0^{\frac12}\Bigg\}\Bigg],\end{aligned}\tag{5.2}$$

where $\bar{\nu}=\max(\nu,\hat{\nu})$ *and* $|\bar{\eta}_i|^2=\max(|\eta_i|^2,|\hat{\eta}_i|^2)$.

Proof The proof is essentially a refined version of the proof of the comparison principle. We begin by introducing the quantities:

$$\Psi_i(t,x,y)=u^i(t,x)-\hat{u}^i(t,y)-\phi(t,x,y)-\frac{\delta\sigma}{T}t-\frac{\bar{\epsilon}}{T-t},$$

where $\delta,\bar{\epsilon}\in(0,1)$, $\phi(t,x,y)$ is defined at (4.1) and γ is chosen to be 0, and

$$\sigma_0=\sup_{i,x,y}\bigg\{u^i(0,x)-\hat{u}^i(0,y)-\phi(0,x,y)-\frac{\bar{\epsilon}}{T}\bigg\}^+,$$

$$\sigma = \sup_{i,t,x,y} \left\{ u^i(t,x) - \hat{u}^i(t,y) - \phi(t,x,y) - \frac{\bar{\epsilon}}{T-t} \right\} - \sigma_0.$$

From the semicontinuity of $u, \hat{u}$ and the growth properties of ϕ along with the penalization term $\frac{\bar{\epsilon}}{T-t}$, there exists $(i_0, t_0, x_0, y_0) \in \mathcal{I} \times [0,T) \times \mathbb{R}^{2n}$ such that

$$\Psi_{i_0}(t_0, x_0, y_0) = \sup_{i,t,x,y} \Psi(t,x,y).$$

We are interested in deriving a positive upper bound on σ; therefore, without loss of generality, we may assume that $\sigma > 0$. This implies that $t_0 > 0$, and we may apply Lemma 4.1. Hence we can choose i_0 so that, $\hat{u}^{i_0}(t_0, y_0) < \mathcal{M}^{i_0}\hat{u}(t_0, y_0)$ and for each $\kappa \in (0,1)$ there exist two symmetric matrices X and Y satisfying

$$\begin{aligned}
&\phi_t(t_0, x_0, y_0) + \frac{\delta\sigma}{T} + \frac{\bar{\epsilon}}{(T-t)^2} \\
&\quad \le \sup_{\alpha \in \mathcal{A}_{i_0}} \inf_{\beta \in \mathcal{B}_{i_0}} \big[\hat{\mathcal{L}}_{i_0}^{\alpha,\beta}(t_0, y_0, \hat{u}^{i_0}(t_0, x_0, y_0), -D_y\phi(t_0, x_0, y_0), Y) \\
&\qquad - \hat{\mathcal{J}}_{i_0,\kappa}^{\alpha,\beta}(t_0, y_0, -D_y\phi(t_0, x_0, y_0), -\phi(t_0, x_0, \cdot)) \\
&\qquad - \hat{\mathcal{J}}_{i_0}^{\alpha,\beta,\kappa}(t_0, y_0, -D_y\phi(t_0, x_0, y_0), \hat{u}^{i_0}(t_0, \cdot)) \big] \\
&\qquad - \sup_{\alpha \in \mathcal{A}_{i_0}} \inf_{\beta \in \mathcal{B}_{i_0}} \big[\mathcal{L}_{i_0}^{\alpha,\beta}(t_0, x_0, u^{i_0}(t_0, x_0, y_0), D_x\phi(t_0, x_0, y_0), X) \\
&\qquad - \mathcal{J}_{i_0,\kappa}^{\alpha,\beta}(t_0, x_0, D_x\phi(t_0, x_0, y_0), \phi(t_0, \cdot, y_0)) \\
&\qquad - \mathcal{J}_{i_0}^{\alpha,\beta,\kappa}(t_0, x_0, D_x\phi(t_0, x_0, y_0), u^{i_0}(t_0, \cdot)) \big],
\end{aligned} \tag{5.3}$$

where the symmetric matrices X and Y will satisfy

$$\begin{pmatrix} X & 0 \\ 0 & -Y \end{pmatrix} \le 2e^{\lambda t_0}\theta \begin{pmatrix} I & -I \\ -I & I \end{pmatrix} + 2\epsilon e^{\lambda t_0} \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}. \tag{5.4}$$

Relation (5.4) along with (A.1)–(A.4) and standard computations yield

$$\begin{aligned}
&\mathrm{Tr}(a_i(t_0, x_0)X) - \mathrm{Tr}(\hat{a}_i(t_0, y_0)Y) \\
&\quad \le K_1 e^{\lambda t_0} \big(\theta|x_0 - y_0|^2 + \theta|\sigma - \hat{\sigma}|_0^2 + \epsilon(1 + |x_0|^2 + |y_0|^2) \big), \\
&\hat{b}_i(t_0, y_0) D_y\phi(t_0, x_0, y_0) + b_i(t_0, x_0) D_x\phi(t_0, x_0, y_0) \\
&\quad = \hat{b}_i(t_0, y_0)(\theta e^{\lambda t_0}(y_0 - x_0) + \epsilon e^{\lambda t_0} y_0) + b_i(t_0, x_0)(\theta e^{\lambda t_0}(x_0 - y_0) + \epsilon e^{\lambda t_0} x_0) \\
&\quad \le K_2 e^{\lambda t_0} \big\{ \theta|b_i - \hat{b}_i|_0^2 + \theta|x_0 - y_0|^2 + \epsilon(1 + |x_0|^2 + |y_0|^2) \big\}, \\
&|\hat{c}_i(t_0, y_0)\hat{u}^i(t_0, y_0) - c_{i_0}(t_0, x_0)u^i(t_0, x_0)| + |\hat{f}_i(t_0, y_0) - f_i(t_0, x_0)| \\
&\quad \le |u^i|_0 \vee |\hat{u}^i| \big(|\hat{c}_i - c_i|_0 + K|x_0 - y_0| \big) + (|f_i - \hat{f}_i|_0 + K|x_0 - y_0|).
\end{aligned}$$

We now turn to the non-local terms. First, observe that

$$\mathcal{J}_{i,\kappa}^{\alpha,\beta}(t_0,x_0,D_x\phi(t_0,x_0,y_0),\phi(t_0,\cdot,y_0)) - \mathcal{J}_{i,\kappa}^{\alpha,\beta}(t_0,y_0,-D_y\phi(t_0,x_0,y_0),-\phi(t_0,x_0,\cdot)) \le (\theta+\epsilon)e^{\lambda t_0}\mathcal{O}(\kappa).$$

Exploiting the fact that (t_0,x_0,y_0) is a point of maximum of $\Psi_{i_0}(t,x,y)$ we get

$$\begin{aligned}
&\int_{\kappa\le|z|\le1}\Big\{u^{i_0}(t_0,x_0+\eta_{i_0}(t_0,x_0,z)) - u^{i_0}(t_0,x_0) - \eta_{i_0}.D_x\phi(t_0,x_0,y_0)\Big\}\nu(dz)\\
&\quad - \int_{\kappa\le|z|\le1}\Big\{\hat{u}^{i_0}(t_0,y_0+\hat{\eta}_{i_0}(t_0,y_0,z)) - \hat{u}^{i_0}(t_0,y_0) - \hat{\eta}_{i_0}.D_y\phi(t_0,x_0,y_0)\Big\}\hat{\nu}(dz)\\
&\le \int_{\kappa\le|z|\le1} e^{\lambda t_0}|\eta_{i_0}(t_0,x_0,z)-\hat{\eta}_{i_0}(t_0,y_0,z)|^2\bar{\nu}(dz)\\
&\quad + \int_{\kappa\le|z|\le1}\epsilon e^{\lambda t_0}\Big(|\eta_{i_0}|^2+|\hat{\eta}_{i_0}|^2\Big)\bar{\nu}(dz)\\
&\quad + (\theta+\epsilon)\int_{\kappa\le|z|\le1}\max\big(|\eta_{i_0}|^2,|\hat{\eta_{i_0}}|^2\big)|\nu-\hat{\nu}|(dz)\\
&\le K_4\theta e^{\lambda t_0}\Bigg\{|x_0-y_0|^2 + \Bigg|\int_{B(0,1)}|\eta_{i_0}-\hat{\eta}_{i_0}|^2\bar{\nu}(dz)\Bigg|_0\\
&\quad + \int_{B(0,1)}\max\big(|\eta_{i_0}|^2,|\hat{\eta}_{i_0}|^2\big)|\nu-\hat{\nu}|(dz)\Bigg\} + K_5\epsilon e^{\lambda t_0}(1+|x_0|^2+|y_0|^2),
\end{aligned}$$

where $\bar{\nu}=\max(\nu,\hat{\nu})$. Once again using that (t_0,x_0,y_0) is a point of maximum for Ψ_{i_0}, along with standing assumptions, we obtain

$$\begin{aligned}
&\int_{|z|\ge1}\Big\{u^{i_0}(t_0,x_0+\eta_{i_0}(t_0,x_0,z)) - u^{i_0}(t_0,x_0)\Big\}\nu(dz)\\
&\quad - \int_{|z|\ge1}\Big\{\hat{u}^{i_0}(t_0,y_0+\hat{\eta}_{i_0}(t_0,y_0,z)) - \hat{u}^{i_0}(t_0,y_0)\Big\}\hat{\nu}(dz)\\
&\le \theta e^{\lambda t_0}K_6\Bigg(|x_0-y_0|^2 + \int_{|z|\ge1}|\eta_{i_0}-\hat{\eta}_{i_0}|^2\bar{\nu}(dz)\\
&\quad + \int_{|z|\ge1}\max\big(|\eta_{i_0}|^2,|\hat{\eta}_{i_0}|^2\big)|\nu-\hat{\nu}|(dz)\Bigg) + \epsilon e^{\lambda t_0}K_6(1+|x|_0^2+|y_0|^2).
\end{aligned}$$

Now by (5.3), the above estimates, and the form of $\phi_t(t_0,x_0,y_0)$, it follows that

$$\begin{aligned}
&\lambda\Big[e^{\lambda t_0}\theta|x_0-y_0|^2 + e^{\lambda t_0}\epsilon(|x_0|^2+|y_0|^2)\Big] + \frac{\delta\sigma}{T} + \frac{\bar{\epsilon}}{(T-t)^2}\\
&\quad \le C_1 e^{\lambda t_0}\theta\max_{i\in\mathcal{I}}\sup_{\alpha\mathcal{A}_i,\beta\in\mathcal{B}_i}\Bigg\{|\sigma_i-\hat{\sigma}_i|_0^2 + |b_i-\hat{b}_i|_0^2 + \Bigg|\int|\eta_i-\hat{\eta}_i|^2\bar{\nu}(dz)\Bigg|_0
\end{aligned}$$

$$+ \left| \int \max\big(|\eta_i|^2, |\hat{\eta}_i|^2\big)|\nu - \hat{\nu}|(dz)\right|_0 \Bigg\} + |u|_0 \vee |\hat{u}|_0 \sup_{i,\alpha,\beta} |c_i - \hat{c}_i|_0$$

$$+ \sup_{i,\alpha,\beta} |f_i - \hat{f}_i|_0 + C_2 e^{\lambda t_0}\theta |x_0 - y_0|^2 + C_3|x_0 - y_0|$$

$$+ C_4 e^{\lambda t_0}\epsilon(1 + |x_0|^2 + |y_0|^2) + \mathcal{O}(\kappa),$$

where the constants only depend on the data. In the above relation the point (t_0, x_0, y_0) is independent of κ, so we can let $\kappa \to 0$ and ignore the term $\mathcal{O}(\kappa)$. Next, we choose $\lambda = 2\max(C_1, C_2, C_3, C_4) + 1$, which gives us

$$\frac{\delta\sigma}{T} \leq C_1 e^{\lambda t_0}\theta \max_{i\in\mathcal{I}} \sup_{\alpha\mathcal{A}_i, \beta\in\mathcal{B}_i} \Bigg\{ |\sigma_i - \hat{\sigma}_i|_0^2 + |b_i - \hat{b}_i|^2 + \left| \int |\eta_i - \hat{\eta}_i|^2 \bar{\nu}(dz)\right|_0$$

$$+ \left| \int \max\big(|\eta_i|^2, |\hat{\eta}_i|^2\big)|\nu - \hat{\nu}|(dz)\right|_0 \Bigg\} + |u|_0 \vee |\hat{u}|_0 \sup_{i,\alpha,\beta} |c_i - \hat{c}_i|_0$$

$$+ \sup_{i,\alpha,\beta} |f_i - \hat{f}_i|_0 + C_3|x_0 - y_0| - \mathrm{e}^{\lambda t_0}\theta|x_0 - y_0|^2 + \mathcal{O}(\epsilon).$$

After a maximization with respect to in $|x_0 - y_0|$ and sending $\delta \to 1$, we obtain

$$\frac{\sigma}{T} \leq C_1 e^{\lambda t_0}\theta \max_{i\in\mathcal{I}} \sup_{\alpha\mathcal{A}_i, \beta\in\mathcal{B}_i} \Bigg\{ |\sigma_i - \hat{\sigma}_i|_0^2 + |b_i - \hat{b}_i|^2$$

$$+ \left| \int |\eta_i - \hat{\eta}_i|^2 \bar{\nu}(dz)\right|_0 + \left| \int \max\big(|\eta_i|^2, |\hat{\eta}_i|^2\big)|\nu - \hat{\nu}|(dz)\right|_0 \Bigg\}$$

$$+ |u|_0 \vee |\hat{u}|_0 \sup_{i,\alpha,\beta} |c_i - \hat{c}_i|_0 + \sup_{i,\alpha,\beta} |f_i - \hat{f}_i|_0 + \frac{C_2}{4\theta e^{\lambda t_0}} + \mathcal{O}(\epsilon). \tag{5.5}$$

Next we estimate σ_0 using the Lipschitz continuity of $u(0, x)$ and $\hat{u}(0, x)$,

$$\sigma_0 \leq \max_i |(u^i(0,\cdot) - \hat{u}^i(0,\cdot))^+|_0 + \sup_{x,y}\left(K|x-y| - \frac{\theta}{2}|x-y|^2\right)$$

$$= \max_i |(u^i(0,\cdot) - \hat{u}^i(0,\cdot))^+|_0 + \frac{K^2}{2\theta}. \tag{5.6}$$

Therefore adding (5.5) and (5.6) and minimizing w.r.t. θ leads to

$$\sigma + \sigma_0 \leq C T^{\frac{1}{2}} \max_{i\in\mathcal{I}} \sup_{\alpha\mathcal{A}_i, \beta\in\mathcal{B}_i} \Bigg\{ |\sigma_i - \hat{\sigma}_i|_0 + |b_i - \hat{b}_i|$$

$$+ \left| \int |\eta_i - \hat{\eta}_i|^2 \bar{\nu}(dz)\right|_0^{\frac{1}{2}} + \left| \int \max\big(|\eta_i|^2, |\hat{\eta}_i|^2\big)|\nu - \hat{\nu}|(dz)\right|_0^{\frac{1}{2}} \Bigg\}$$

$$+ T|u|_0 \vee |\hat{u}|_0 \sup_{i,\alpha,\beta} |c_i - \hat{c}_i|_0 + T \sup_{i,\alpha,\beta} |f_i - \hat{f}_i|_0$$

$$+ \max_i |(u^i(0,\cdot) - \hat{u}^i(0,\cdot))^+|_0 + \mathcal{O}(\epsilon).$$

For every $(t,x) \in [0,T) \times \mathbb{R}^n$,

$$u^i(t,x) - \hat{u}^i(t,x) - \epsilon|x|^2 - \frac{\bar{\epsilon}}{T-t} \le \sigma + \sigma_0,$$

so the proof is complete since (5.2) follows from the last two estimates if we send $\epsilon, \bar{\epsilon} \to 0$. □

As a simple consequence of the continuous dependence estimate, we have the Lipschitz continuity in the x-variable of the viscosity solution of (1.1).

Lemma 5.2 *Assume that* (A.1)–(A.4) *hold, and let* $u \in C_b(Q_T; \mathbb{R}^M)$ *be the unique viscosity solution of* (1.1). *Then there is a constant* L, *depending only the data* (*and* T), *such that*

$$|u^i(t,x+h) - u^i(t,x)| \le L|h|$$

for all $h \in \mathbb{R}^n$, $(t,x) \in [0,T) \times \mathbb{R}^n$, *and* $i \in \mathcal{I}$.

Proof For all $i \in \mathcal{I}$ and $\alpha \in \mathcal{A}_i$, $\beta \in \mathcal{B}_i$, define

$$\Big(\hat{\sigma}_i, \hat{b}_i, \hat{c}_i, \hat{f}_i, \hat{\eta}_i\Big)(t,x) := \Big(\sigma_i, b_i, c_i, f_i, \eta_i\Big)(t, x+h)$$

with $\hat{\nu} = \nu$ and $\hat{u}(0,x) = u(0,x+h)$. By uniqueness, $\hat{u} = u(t,x+h)$ is the unique viscosity solution of (5.1) and then the rest of the proof is just a consequence of Theorem 5.1, once we observe that the right-hand side of (5.2) can be estimated by $L|h|$ with the constant L depending only the data. □

Next, we prove a Hölder continuity result in the time variable. Remember that the data are only continuous in time and thus, as in the scalar case, the equation induces some extra regularity in time on the solution.

Lemma 5.3 *Assume that* (A.1)–(A.4) *hold, and let* $u(t,x)$ *be the unique viscosity solution of* (1.1). *Then there is a constant* C, *depending only on the data and* T, *such that*

$$|u^i(t,y) - u^i(t',y)| \le C(1+|y|)|t-t'|^{\frac{1}{2}},$$

for all $y \in \mathbb{R}^n$ *and* $t, t' \in [0,T)$.

Proof Without loss of generality we may assume that $t' = 0$ and $|t| \le 1$ (since solutions are bounded). For $y \in \mathbb{R}^n$, define

$$\psi_i(s,x) = \lambda L\Big[e^{Ds}|x-y|^2 + \gamma s(1+|y|^2)\Big] + Ks + \lambda^{-1}L + u^i(0,y),$$

for all $(s,x) \in Q_T$ and $i \in \mathcal{I}$, with L being the Lipschitz constant defined in Lemma 5.2 and D, γ are constants to be chosen later. Observe that

$$\partial_s \psi_i(s,x) = \lambda L\big[De^{Ds}|x-y|^2 + \gamma(1+|y^2|)\big] + K,$$

$$\mathcal{L}_i^{\alpha,\beta}(s,x,D\psi_i,D^2\psi_i) - \mathcal{J}_i^{\alpha,\beta}(\psi_i(s,\cdot)) \\ \geq -\lambda L N_0 e^{Ds}\big((1+|x|^2) + |x-y|(1+|x|)\big) - K,$$

for all α, β. Therefore,

$$\partial_s \psi_i(s,x) + \sup_\alpha \inf_\beta \Big[\mathcal{L}_i^{\alpha,\beta}(s,x,D\psi_i,D^2\psi_i) - \mathcal{J}_i^{\alpha,\beta}(\psi_i(s,\cdot))\Big] \geq 0$$

for all s, x whenever D and γ are chosen large enough. Furthermore,

$$\psi_i(0,x) = \lambda L|x-y|^2 + \lambda^{-1}L + u^i(0,y) \\ \geq L|x-y| + u^i(0,y) \geq u^i(0,x), \quad \forall x \in \mathbb{R}^n.$$

We conclude that $\psi = (\psi_1, \psi_2, \dots, \psi_M)$ is a supersolution of (1.1), and hence the comparison principle yields

$$u^i(t,y) \leq \lambda L\gamma(1+|y|^2)t + Kt + \lambda^{-1}L + u^i(0,y).$$

Upon minimizing the right-hand side with respect to λ along with $|t| \leq 1$, we obtain

$$u^i(t,y) - u^i(0,y) \leq N(1+|y|)t^{\frac{1}{2}}.$$

The other inequality follows in a similar manner. □

In view of Lemmas 5.2 and 5.3, the proof of Theorem 2.5 is now concluded.

References

1. Alvarez, O., Tourin, A.: Viscosity solutions of nonlinear integro-differential equations. Ann. Inst. Henri Poincaré Anal. Non Linéaire **13**(3), 293–317 (1996)
2. Amadori, A.L.: The obstacle problem for nonlinear integro-differential operators arising in option pricing. Quaderno IAC Q21-000 (2000)
3. Amadori, A.L.: Nonlinear integro-differential evolution problems arising in option pricing: a viscosity solutions approach. Differ. Integr. Equ. **16**(7), 787–811 (2003)
4. Bardi, M., Capuzzo-Dolcetta, I.: Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations. In: Systems & Control: Foundations & Applications. Birkhäuser Boston, Cambridge (1997)
5. Barles, G., Buckdahn, R., Pardoux, E.: Backward stochastic differential equations and integral-partial differential equations. Stoch. Stoch. Rep. **60**(1–2), 57–83 (1997)
6. Barles, G., Chasseigne, E., Imbert, C.: On the Dirichlet problem for second-order elliptic integro-differential equations. Indiana Univ. Math. J. **57**(1), 213–246 (2008)
7. Barles, G., Imbert, C.: Second-order elliptic integro-differential equations: viscosity solutions' theory revisited. Ann. Inst. Henri Poincaré Anal. Non Linéaire **25**(3), 567–585 (2008)
8. Barles, G., Jakobsen, E.R.: On the convergence rate of approximation schemes for Hamilton-Jacobi-Bellman equations. M2AN Math. Model. Numer. Anal. **36**(1), 33–54 (2002)
9. Barles, G., Jakobsen, E.R.: Error bounds for monotone approximation schemes for Hamilton-Jacobi-Bellman equations. SIAM J. Numer. Anal. **43**(2), 540–558 (2005) (electronic)
10. Barles, G., Jakobsen, E.R.: Error bounds for monotone approximation schemes for parabolic Hamilton-Jacobi-Bellman equations. Math. Comput. **76**(240), 1861–1893 (2007)
11. Basu, A., Ghosh, M.K.: Stochastic differential games with multiple modes and applications to portfolio optimization. Stoch. Anal. Appl. **25**, 845–867 (2007)

12. Benth, F.E., Karlsen, K.H., Reikvam, K.: Optimal portfolio selection with consumption and nonlinear integrodifferential equations with gradient constraint: a viscosity solution approach. Finance Stoch. **5**, 275–303 (2001)
13. Benth, F.E., Karlsen, K.H., Reikvam, K.: Portfolio optimization in a Lévy market with intertemporal substitution and transaction costs. Stoch. Stoch. Rep. **74**(3–4), 517–569 (2002)
14. Billingsley, P.: Convergence of Probability Measures, 2nd edn. Wiley Series in Probability and Statistics. Wiley, New York (1999)
15. Biswas, I.H., Jakobsen, E.R., Karlsen, K.H.: Difference-quadrature schemes for nonlinear degenerate parabolic integro-PDE (2009, submitted)
16. Bonnans, J.F., Baroso, S., Zidani, H.: Error estimates for a stochastic impulse control problem. Appl. Math. Optim. **55**(3), 327–357 (2007)
17. Caffarelli, L., Silvestre, L.: Regularity theory for fully nonlinear integro-differential equations. Commun. Pure Appl. Math. **62**(5), 597–638 (2009)
18. Caffarelli, L., Souganidis, P.E.: A rate of convergence for monotone finite difference approximations to fully nonlinear, uniformly elliptic PDEs. Commun. Pure Appl. Math. **61**(1), 1–17 (2008)
19. Capuzzo Dolcetta, I., Evans, L.C.: Optimal switching for ordinary differential equations. SIAM J. Control. Optim. **22**, 365–389 (1984)
20. Cont, R., Tankov, P.: Financial Modeling with Jump Processes. Chapman & Hall/CRC, Boca Raton (2004)
21. Crandall, M.G., Ishii, H., Lions, P.-L.: User's guide to viscosity solutions of second order partial differential equations. Bull. Am. Math. Soc. (N.S.) **27**(1), 1–67 (1992)
22. Evans, L.C., Friedman, A.: Optimal stochastic switching and the Dirichlet problem for the Bellman equation. Trans. Am. Math. Soc. **253**, 365–389 (1979)
23. Fleming, W.H., Sethi, S.P., Soner, H.M.: An optimal stochastic production planning with randomly fluctuating demand. SIAM J. Control. Optim. **25**, 1494–1502 (1987)
24. Fleming, W.H., Soner, H.M.: Controlled Markov Processes and Viscosity Solutions. Springer, New York (1993)
25. Fleming, W.H., Souganidis, P.E.: On the existence of value functions of two player, zero-sum stochastic differential Games. Indiana Univ. Math. J. **38**(2), 293–314 (1989)
26. Ishii, H.: Perron's method for Hamilton-Jacobi equations. Duke Math. J. **55**, 369–384 (1987)
27. Ishii, H., Koike, S.: Viscosity solutions of a system of nonlinear second-order elliptic PDEs arising in switching games. Funkc. Ekvacioj. **34**(1), 143–155 (1991)
28. Ishii, H., Koike, S.: Viscosity solutions for monotone systems of second-order elliptic PDEs. Commun. Partial Differ. Equ. **16**, 1095–1128 (1991)
29. Ishikawa, Y.: Optimal control problem associated with jump processes. Appl. Math. Optim. **50**(1), 21–65 (2004)
30. Jakobsen, E.R., Karlsen, K.H.: Continuous dependence estimates for viscosity solutions of fully nonlinear degenerate elliptic equations. Electron. J. Differ. Equ. **39**, 1–10 (2002)
31. Jakobsen, E.R., Karlsen, K.H.: Continuous dependence estimates for viscosity solutions of integro-PDEs. J. Differ. Equ. **212**(2), 278–318 (2005)
32. Jakobsen, E.R., Karlsen, K.H.: A "maximum principle for semicontinuous functions" applicable to integro-partial differential equations. NoDEA Nonlinear Differ. Equ. Appl. **13**(2), 137–165 (2006)
33. Krylov, N.V.: On the rate of convergence of finite-difference approximations for Bellman's equations. Algebra Anal. **9**(3), 245–256 (1997)
34. Krylov, N.V.: On the rate of convergence of finite-difference approximations for Bellman's equations with variable coefficients. Probab. Theory Relat. Fields **117**(1), 1–16 (2000)
35. Krylov, N.V.: The rate of convergence of finite-difference approximations for Bellman equations with Lipschitz coefficients. Appl. Math. Optim. **52**(3), 365–399 (2005)
36. Lenhart, S.M., Liao, Y.C.: Switching control of piecewise-deterministic processes. J. Optim. Theory Appl. **59**(1), 99–115 (1988)
37. Øksendal, B., Sulem, A.: Applied Stochastic Control of Jump Diffusions. Springer, Berlin (2005)
38. Pham, H.: Optimal stopping of controlled jump diffusion processes: a viscosity solution approach. J. Math. Syst. Estim. Control **8**(1), 27 (1998) (electronic)
39. Sato, K.I.: Levy Processes and Infinite Divisibility. Cambridge University Press, Cambridge (2000)
40. Sayah, A.: Équations d'Hamilton-Jacobi du premier ordre avec termes intégro-différentiels. I. Unicité des solutions de viscosité. Commun. Partial Differ. Equ. **16**(6–7), 1057–1074 (1991)
41. Sayah, A.: Équations d'Hamilton-Jacobi du premier ordre avec termes intégro-différentiels. II. Existence de solutions de viscosité. Commun. Partial Differ. Equ. **16**(6–7), 1075–1093 (1991)

42. Sharifnia, A.: Production control of a manufacturing systems with multiple machine states. IEEE Trans. Automat. Control **AC-33**, 620–625 (1988)
43. Stojanovic, S., Yong, J.M.: Optimal switching for partial differential equations. I. J. Math. Anal. Appl. **138**(2), 418–438 (1989)
44. Stojanovic, S., Yong, J.M.: Optimal switching for partial differential equations. II. J. Math. Anal. Appl. **138**(2), 439–460 (1989)
45. Tang, S., Hou, S.: Switching games of stochastic differential systems. SIAM J. Control Optim. **46**(3), 900–929 (2007)
46. Tang, S., Yong, J.M.: Finite horizon stochastic optimal switching and impulse controls with viscosity solution approach. Stoch. Stoch. Rep. **45**, 145–176 (1993)
47. Yong, J.M.: Systems governed by ordinary differential equations with continuous, switching and impulse controls. Appl. Math. Optim. **20**(3), 223–235 (1989)
48. Yamada, N.: Viscosity solutions for a system of elliptic inequalities with bilateral obstacles. Funkc. Ekvacioj. **30**, 417–425 (1987)