Risk Sensitive and Robust Nonlinear Filtering

Wendell H. Fleming* and William M. McEneaney†

∗Division of Applied Math., Brown Univ., Providence, RI 02912, USA, whf@cfm.brown.edu

†Dept. of Math., North Carolina State Univ., Raleigh, NC 27695, USA, wmm@eos.ncsu.edu

Abstract

A risk sensitive approach to nonlinear filtering is considered. In this case, the traditional expected mean squared error is replaced by an expected exponential–of–mean squared error. A pathwise filtering equation is used to take the risk averse limit. This leads to a robust filtering estimator. The robust estimator yields a bound on the estimate error in terms of (finite) disturbance energy.

1. Introduction

In the traditional stochastic nonlinear model, state dynamics and observation errors are modeled as white noises. The objective is to minimize an expected mean squared estimation error criterion. The theory reduces the problem to the solution of the Zakai stochastic partial differential equation (PDE) for an unnormalized conditional density, followed by integrations. Mortensen's theory of deterministic estimation is a traditional alternative to the stochastic approach. In this deterministic approach, an optimal state estimate is found as the argmax of the solution to a first–order, nonlinear Hamilton–Jacobi–Bellman (HJB) PDE. The Mortensen approach is reviewed in Section 2, and some useful properties of solutions to the HJB equation are noted. For early references on this topic, see Mortensen [Mo], Mitter [M].

We consider a stochastic risk sensitive filter model in which the traditional expected mean squared error is replaced by an expected exponential–of–mean squared error. This leads to a greater emphasis on the Brownian motion paths which lead to relatively poor estimate quality in analogy with risk-sensitive control [FMc]. In this case, the solution again depends on the unnormalized conditional density, or equivalently, on the solution to a corresponding pathwise filter equation which takes the form of a second–order, nonlinear PDE. It should be noted here that this criterion involves only the error at the current time rather than the integrated error criterion which emerges in partially–observed risk sensitive optimal control, and consequently one avoids the infinite–dimensional PDE's associated with partially–observed risk sensitive control. This filter is discussed in Section 3.

A connection between stochastic nonlinear filtering and Mortensen's deterministic estimator was made in Hijab [H]. This was done by letting state and observation noise intensities in the stochastic model depend on a small parameter, ϵ. The deterministic estimator was obtained in the $\epsilon \longrightarrow 0$ limit. Analogous results for the risk sensitive filtering problem are discussed here. In this case however, the limiting estimator is the robust filter in McEneaney's formulation [Mc1] which is a filtering analogue of robust/H_∞ control. The limit results are then obtained through a combination of a control interpretation argument, and a comparison result for viscosity solutions of the limit HJB equation. The risk averse limit is analyzed in Section 4, and the robust interpretation of the limit appears in Section 5. The proof of the convergence result is outlined in Section 6.

2. Mortensen's Deterministic Estimator

Let x_t denote the state to be estimated at time $t \geq 0$, with $x_t \in \Re^n$. The state dynamics are

$$\frac{dx}{dt} = f(x_t) + \sigma(x_t)w_t, \tag{2.1}$$

where w_t is the disturbance in the dynamics taking values in $\Re^m$. Let the observations taking values in $\Re^l$ be given by

$$y_t = h(x_t) + \rho v_t, \qquad \rho > 0 \tag{2.2}$$

where v_t is the observation disturbance taking values in $\Re^l$ also. We assume that w_t and v_t are square integrable. Let $\phi(x_0)$ be a measure of our uncertainty (to be clarified below) about the initial state x_0.

We make the following assumptions throughout. The functions f, σ, h, ϕ are C^2. Moreover, f, f_x, σ, σ^{-1}, σ_x, σ_{xx}, h, h_x, h_{xx} are bounded. We also assume that there exist $D_1, D_2 \in (0,\infty)$ such that

$$D_1|x|^2 \leq \phi(x) \leq D_2|x|^2 \qquad \forall\, x.$$

A continuous observation path, $y_\cdot$, is fixed in the following discussion.

Research partially supported by ONR grant N0014-96-1-0267, AFOSR grant F49620-95-1-0296, NSF grant DMS-9531276 and ARO grant DAAL03-92-G0115.

We seek an "optimal" estimate $\hat{x}_T$ for x_T, given observations y_t for $0 \le t \le T$. Mortensen's estimate is the following. Given a final state $x_T = x$, let

$$\begin{aligned} J(T, x, w.) &= \phi(x_0) + \frac{1}{2}\int_0^T |w_t|^2 + |v_t|^2\, dt \\ &= \phi(x_0) + \frac{1}{2}\int_0^T |w(t)|^2 + |\rho^{-1}[y_t - h(x_t)]|^2\, dt \end{aligned}$$

where we note that the state path is fixed for any particular choice of x and $w.$. Note that v also no longer appears in the second equation since it is fixed for a given x, $w.$, and observation path $y.$. Then let

$$V(T, x) = \inf_{w \in L_2} J(T, x, w.). \tag{2.3}$$

Mortensen's estimate is $\hat{x}_T \in \operatorname{argmin}_x V(T, x)$. The existence of $\hat{x}$ is guaranteed by the easily obtained estimate (see [Mc1])

$$V(T, x) \ge C_T |x|^2 \tag{2.4}$$

for some $C_T < \infty$.

The function V is a viscosity solution of the HJB equation [Mc1]

$$\begin{aligned} 0 &= -V_T - f^T(x)V_x - \frac{1}{2}V_x^T a(x) V_x \\ &\quad + \frac{1}{2}|\rho^{-1}(y_T - h(x))|^2 \\ V(0, x) &= \phi(x) \end{aligned} \tag{2.5}$$

where V_x represents the gradient and $a(x) \doteq \sigma(x)\sigma^T(x)$. In the case of constant σ, uniqueness among the class of quadratically growing solutions is obtained in [Mc2]. Recent results of Bardi–DaLio [BDL] and Ishii [I] suggest that this may be obtained without the σ–constant assumption. Here, a cone–of–dependence argument and comparison result will allow limit argument to go through.

The Solution Operator

A function $\psi(\cdot)$ is called semi-concave on $\Re^n$ if for every $R > 0$, there exists C_R such that $\phi(x) - C_R|x|^2$ is concave on the ball $\{|x| \le R\}$. By adapting standard arguments (see Fleming–Soner, Sec. 4.9 [FS]), it can be shown that $V(T, \cdot)$ is semi-concave for each $T \ge 0$.

Let S_T denote the nonlinear operator which maps the initial data, ϕ in (2.5) to the solution (2.3):

$$V(T, x) = S_T \phi(x).$$

Besides such familiar properties of S_T such as monotonicity and $S_T(\phi + c) = S_T\phi + c$ for any constant c, the following holds.

Lemma 2.1: If $\phi(x) = \min_{z \in Z} \phi_z(x)$ where Z is a finite set, then

$$S_T\phi(x) = S_T\left[\min_{z\in Z}\phi_z(x)\right] = \min_{z \in Z} S_T\phi_z(x).$$

By using convex duality, semi-concave functions can be represented as pointwise minima taken over a class of quadratic functions $\phi_z(x)$. Lemma 2.1 then provides a representation of $S_T\phi$ in terms of $S_T\phi_z$ for quadratic ϕ_z. This is useful in the design of a particular method for numerical solution to (2.5), and will be discussed in more detail in a forthcoming paper.

3. Risk Sensitive Estimates

Instead of the deterministic model in Section 2, we now consider a (standard) stochastic model in which X_t denotes the state process and Y_t denotes an accumulated observations process. They satisfy stochastic differential equations (Ito sense)

$$dX_t = f(X_t)\,dt + \sqrt{\epsilon}\sigma(X_t)\,dB_t \tag{3.1}$$
$$dY_t = h(X_t)\,dt + \sqrt{\epsilon}\rho\,d\widetilde{B}_t \qquad Y_0 = 0 \tag{3.2}$$

where $B.$, $\widetilde{B}.$ are independent Brownian motion processes. The initial state X_0 is independent of $B.$, $\widetilde{B}.$, and has density $k_\epsilon \exp(-\epsilon^{-1}\phi(x)]$ for some constant k_ϵ. In this section, the parameter, ϵ, is fixed.

Let q^ϵ denote the unnormalized conditional density of X_t, which is a solution to the Zakai stochastic PDE with initial condition $q^\epsilon(0, x) = \exp(-\epsilon^{-1}\phi(x))$. Let $V^\epsilon = -\epsilon \log q^\epsilon$. It is not difficult to show, using a comparison principle for solutions to linear, second–order, parabolic PDE's, that

$$K_1(t)|x|^2 - M_1(t) \le V^\epsilon(t, x) \le K_2(t)|x|^2 + M_2(t), \tag{3.3}$$

for suitable $K_1 > 0$, K_2, M_1, M_2, where K_1, K_2 satisfy Ricatti differential equations and depend on $Y.$.

For fixed T, let

$$\psi^\epsilon(e) = \int_{\Re^n} \exp\left[\frac{\mu}{\epsilon}|x - e|^2\right] q^\epsilon(x, T)\,dx, \tag{3.4}$$

where $0 < \mu < K_1(T)$. The risk sensitive filtering problem is to find an estimate $\hat{e}_T^\epsilon$ which minimizes $\psi^\epsilon(e)$. Routine differentiations show that $\psi^\epsilon(e)$ is strictly convex. Moreover,

$$\alpha_2|e|^2 - \beta_2 \le \epsilon \log \psi^\epsilon(e) \le \alpha_1|e|^2 + \beta_1, \tag{3.5}$$

for suitable $\alpha_i = \alpha_i(T) > 0$ and $\beta_i = \beta_i(T)$, $i = 1, 2$. Hence $\psi^\epsilon(e)$ has a minimum at a unique $\hat{e}_T^\epsilon$.

Minimizing $\psi^\epsilon(e)$ is equivalent to finding a $\mathcal{F}_T^Y$–measurable estimator e_T which minimizes the criterion $E\exp\left[\frac{\mu}{\epsilon}|X_T - e_T|^2\right]$, where $\mathcal{F}_T^Y$ is the σ–algebra generated by the accumulated observations Y_t for $0 \le t \le T$. Other authors consider instead the problem of finding $\mathcal{F}_T^Y$–measurable estimators e_t for $0 \le t \le T$ such that $E\exp\left[\frac{\mu}{\epsilon}\int_0^T |X_t - e_t|^2\,dt\right]$ is minimized. That formulation can be considered as a special case of an output–feedback, risk sensitive optimal control problem in which the estimate e_t has the role of a minimizing control. See for instance Boel–James–Petersen [BJP], and Moore–Elliott–Dey [MED], and references cited therein. In that formulation, the optimal risk sensitive estimator involves the solution to an infinite–dimensional PDE for the dynamics of the information state. By considering the exponential terminal cost criterion (3.4), the analysis remains at the level of finite–dimensional PDE's. (With the other formulation, in the deterministic limit $\epsilon \downarrow 0$ the suboptimal "certainty equivalent" filter reduces the problem back to a finite–dimensional level. See James–Baras–Elliott [JBE] and Yuliar–James [YJ]. Alternatively, one can look for specific systems which can be reduced to finite–dimensional problems [CE].)

4. Deterministic Limits

To compare the deterministic model in Section 2 and the stochastic model in Section 3, we let $\epsilon \downarrow 0$ in a manner similar to Hijab [H] and James–Baras [JB]. In the deterministic model, a function y_t is observed for $0 \le t \le T$, while in the stochastic model, a sample path of the accumulated observation process $Y_t = Y_t^\epsilon$ is observed on $[0, T]$. To compare the two models, it is convenient to use "pathwise" filtering theory.

Let

$$q^\epsilon(t,x) = p^\epsilon(t,x)\exp\left[\epsilon^{-1}Y_t \cdot h(x)\right], \qquad (4.1)$$

where $p^\epsilon(t,\cdot)$ is called a pathwise unnormalized filter density. Then p^ϵ satisfies a linear second–order parabolic PDE, with coefficients depending on Y_t. See for example, Fleming–Pardoux [FP, p. 271]. Since $Y_0 = 0$,

$$p^\epsilon(0,x) = q^\epsilon(0,x) = \exp\left[-\epsilon^{-1}\phi(x)\right].$$

According to pathwise filtering theory, we now regard $Y_\cdot$ as a *fixed* accumulated observation path. We take $Y_\cdot$ smooth (class C^1). Then $y_t = \dot{Y}_t$ is a fixed continuous observation path. Routine differentiations then show that the pathwise filter PDE for p^ϵ is equivalent to the Stratonovich form of the Zakai PDE for q^ϵ.

$$q_T^\epsilon = (L^\epsilon)^* q^\epsilon - \frac{1}{2\epsilon\rho^2}|h|^2 q^\epsilon + \frac{h \cdot y_T}{\epsilon\rho^2} q^\epsilon \qquad (4.3)$$

where L^ϵ is the generator of the Markov diffusion X_t in (3.1). Then $V^\epsilon = -\epsilon \log q^\epsilon$ satisfies a corresponding PDE which for notational simplicity we exhibit only for the special case σ = identity matrix

$$\begin{aligned} V_T^\epsilon &= \frac{\epsilon}{2}\Delta_x V^\epsilon - f \cdot V_x^\epsilon - \frac{1}{2}|V_x^\epsilon|^2 + \epsilon f_x - \frac{1}{\rho^2} y_T \cdot h \\ &\quad + \frac{1}{2\rho^2}|h|^2 \\ V^\epsilon(0,x) &= \phi(x) \end{aligned} \qquad (4.4)$$

where V_x denotes the gradient.

As $\epsilon \downarrow 0$, V^ϵ tends formally to a limit V^0 which should satisfy (in a viscosity sense) the corresponding PDE with $\epsilon = 0$

$$\begin{aligned} 0 &= -V_T^0 - f \cdot V_x^0 - \frac{1}{2}|V_x^0|^2 - \frac{1}{\rho^2} y_T \cdot h + \frac{1}{2\rho^2}|h|^2 \\ &= -V_T^0 + \inf_{w \in \Re^n}\left[(-f - w)\cdot V_x^0 + \frac{w^2}{2}\right] \\ &\quad + \frac{1}{2\rho^2}|y_T - h|^2 - \frac{1}{2\rho^2}|y_T|^2 \\ V^0(0,x) &= \phi(x) \end{aligned} \qquad (4.5)$$

Note that this PDE differs from that in (2.5) by a term $\frac{1}{2}\rho^{-2}|y_T|^2$ on the right side. This suggests the following result. The proof is quite long, and consequently, an outline of the proof is delayed to Section 6.

Theorem 4.1: As $\epsilon \downarrow 0$, $V^\epsilon \to V^0$ uniformly on compact sets, where

$$V^0(T,x) = V(T,x) - \frac{1}{2\rho^2}\int_0^T |y_t|^2\,dt \qquad (4.6)$$

with $V(T,x)$ as in (2.3).

Let $g^\epsilon(e) = \epsilon \log \psi^\epsilon(e)$. We can rewrite (3.4) as (for fixed T)

$$g^\epsilon(e) = \epsilon \log \int_{\Re^n} \exp\left[\epsilon^{-1}F^\epsilon(x,e)\right]\,dx$$

where

$$F^\epsilon(x,e) = \mu|x-e|^2 - V^\epsilon(x,T).$$

Using a rather standard Laplace–Varadhan asymptotic principle argument, one can show that $g^\epsilon(e)$ converges to a limit $g^0(e)$ as $\epsilon \downarrow 0$, uniformly on compact sets where

$$g^0(e) = \max_x \left\{\mu|x-e|^2 - V^0(x,T)\right\}. \qquad (4.7)$$

The function g^0 is strictly convex, and satisfies the same quadratic growth bounds as in (3.5). As a consequence:

Theorem 4.2: As $\epsilon \downarrow 0$, $\hat{e}_T^\epsilon \to \hat{e}_T^0$ where $\hat{e}_T^0 = \operatorname{argmin}_e g^0(e)$, provided $\mu < K_1(T)$ with $K_1(T)$ as in (3.3).

From (4.7), $\hat{e}_T^0$ minimizes $\max_x[\mu|x-e|^2 - V(T,x)]$, which is exactly the same criterion as considered in McEneaney [Mc1]. In the special case where $\mu|x|^2 - V(T,x)$ is a concave function of x, bounded above by $-c_1|x|^2 + c_2$ for some $c_1 > 0$ and c_2, the expression $\mu|x-e|^2 - V(T,x)$ has the "saddle point" property. That is,

$$\begin{aligned}\min_e \max_x [\mu|x-e|^2 - V(T,x)] \\ = \max_x \min_e [\mu|x-e|^2 - V(T,x)].\end{aligned}$$

When $\min_e$ is taken first, the minimum occurs at $e = x$. Thus, when the saddle point property holds, $\hat{e}_T^0 = \hat{x}_T$ where $\hat{x}_T$ is the deterministic Mortensen estimator in Section 2. In particular, if $f(x) = Ax$, $h(x) = Hx$, σ is constant, and $\phi(x) = x^T Dx$ is positive definite, then $V(T,x)$ is quadratic. Then $\mu|x-e|^2 - V(T,x)$ is concave provided $\mu < \widehat{\mu}$ where $\widehat{\mu}$ is a critical parameter value which makes $\mu|x-e|^2 - V(T,x) \leq 0$.

5. Robust Nonlinear Filtering

As in McEneaney [Mc1], we interpret the estimator $\hat{e}_T^0$ as a robust estimator in the following sense. (See also Krener–Duarte [KD] for a similar approach.) There exists μ^* such that for all $\mu \leq \mu^*$ (see [Mc1])

$$g^0(\hat{e}_T^0) = \min_e \max_x \left[\mu|x-e|^2 - V(T,x)\right] \leq 0. \quad (5.1)$$

Then, for all $x \in \Re^n$,

$$\mu|x - \hat{e}_T^0|^2 \leq V(T,x).$$

If we take $x = x_T$, and recall (2.1), (2.3), then equivalently, for all initial states x_0 and all square–integrable disturbances $w_\cdot$, $v_\cdot$

$$\mu|x - \hat{e}_T^0|^2 \leq \phi(x_0) + \frac{1}{2}\int_0^T \left[|w_t|^2 + |v_t|^2\right] dt.$$

Let $\mu = \gamma^{-2}$, where γ has the role of an H_∞–bound parameter. Then $\gamma^2 \geq \gamma^{*2} \doteq (\mu^*)^{-1}$ implies that $\hat{e}_T^0$ is a robust estimator in the sense that

$$|x - \hat{e}_T^0|^2 \leq \gamma^2 \left[\phi(x_0) + \frac{1}{2}\int_0^T \left[|w_t|^2 + |v_t|^2\right] dt\right] \quad (5.2)$$

for all $x_0, w_\cdot, v_\cdot$.

Thus there is a critical $\gamma^0 \leq \gamma^*$ such that (5.2) holds for all $\gamma \geq \gamma^0$. This γ^0 plays the role of the critical H_∞–bound parameter. As in H_∞ control, one would not generally search for the optimal γ^0, but would instead obtain a robust filter for some reasonable choice of $\gamma > \gamma^0$. In addition to the solution operator discussed in Section 2, one can also obtain V (and consequently $\hat{e}_T^0$) via a method of characteristics [Mc1].

6. Outline of Proof of Theorem 4.1

As the proof is quite long, and the notation can become cumbersome, we only outline the proof in the scalar case. The proof is not quite already known due to the quadratic growth of the solutions and the presence of the quadratic term in the gradient in the HJB equations. However some rather tedious but standard techniques lead to the result.

Reverse the time variable in the solutions of (4.4) and (4.5) as follows:

$$\begin{aligned}\overline{V}^\epsilon(-t,x) \doteq V^\epsilon(t,x) \\ \overline{V}^0(-t,x) \doteq V^0(t,x)\end{aligned}$$

We find that $\overline{V}^\epsilon, \overline{V}^0$ satisfy

$$\begin{aligned}
&0 = \overline{V}^\epsilon_t + \frac{\epsilon}{2}\Delta_x \overline{V}^\epsilon + \inf_{w\in\Re}\left[(-f-w)\overline{V}^\epsilon_x + \frac{w^2}{2}\right] - B_\epsilon \\
&\qquad\qquad (t,x) \in (-\infty,0)\times\Re^n \qquad (6.1) \\
&\overline{V}^\epsilon(0,x) = \phi(x) \qquad x \in \Re^n \\
&0 = \overline{V}^0_t + \inf_{w\in\Re}\left[(-f-w)\overline{V}^0_x + \frac{w^2}{2}\right] - B_0 \\
&\qquad\qquad (t,x) \in (-\infty,0)\times\Re^n \qquad (6.2) \\
&\overline{V}^0(0,x) = \phi(x) \qquad x \in \Re^n
\end{aligned}$$

where $B_\epsilon = -\epsilon f_x - \frac{1}{2\rho^2}|y_{-t} - h|^2 + \frac{1}{2\rho^2}|y_{-t}|^2$ and $B_0 = -\frac{1}{2\rho^2}|y_{-t} - h|^2 + \frac{1}{2\rho^2}|y_{-t}|^2$. By more or less standard techniques ([FS], [Mc3]), one finds that $\overline{V}^\epsilon$ and $\overline{V}^0$ are the values of the following corresponding control problems. (There is actually a slight difficulty here for the $\overline{V}^\epsilon$ problem due to the quadratic growth. This can be overcome by using processes with stopping times where the process stops when the magnitude of the state reaches some value, and then letting this value go to infinity. One needs to obtain a bound on $E\{\sup_{r\in[t,0]}[X_r^2]\}$ for each fixed t. A martingale inequality is used to eliminate the supremum.)

$$
\begin{aligned}
&dX^\epsilon_r = [-f(X^\epsilon_r) - w_r]\,dr + \sqrt{\epsilon}\,dB_r, \qquad r \in (t,0)\\
&X^\epsilon_t = x\\
&J^\epsilon(t,x,w) = E\left[\phi(X^\epsilon_0) + \int_t^0 -B_\epsilon + \frac{1}{2}w_r^2\,dr\right]\\
&\overline{V}^\epsilon(t,x) = \inf_{w\in L_2(\Omega\times(t,0))} J^\epsilon(t,x,w)\\
&\frac{dX^0}{dr} = [-f(X_r) - w_r], \qquad r \in (t,0)\\
&X^0_t = x\\
&J^0(t,x,w) = \phi(X^0_0) + \int_t^0 -B_0 + \frac{1}{2}w_r^2\,dr\\
&\overline{V}^0(t,x) = \inf_{w\in L_2} J^0(t,x,w).
\end{aligned} \tag{6.3}
$$

It is useful to define

$$
\phi^R(x) = \begin{cases} \phi(x) & \text{if } |\phi(x)| \le R\\ R & \text{if } \phi(x) > R, \end{cases}
$$

and let $\overline{V}^{\epsilon,R}$, and $\overline{V}^{0,R}$ be the solutions of (6.1), (6.2) with this replacement for ϕ. (Note that we do not claim that $\overline{V}^{\epsilon,R}$ is related to a filtering problem, but merely introduce it as a solution to the particular PDE.) By similar arguments to those alluded to just above, one may show that $\overline{V}^{\epsilon,R}$ and $\overline{V}^{0,R}$ are the values of the following control problems.

$$
\begin{aligned}
&J^{\epsilon,R}(t,x,w) = E\left[\phi^R(X^\epsilon_0) + \int_t^0 -B_\epsilon + \frac{1}{2}w_r^2\,dr\right]\\
&\overline{V}^{\epsilon,R}(t,x) = \inf_{w\in L_2(\Omega\times(t,0))} J^{\epsilon,R}(t,x,w)\\
&J^{0,R}(t,x,w) = \phi^R(X^0_0) + \int_t^0 -B_0 + \frac{1}{2}w_r^2\,dr\\
&\overline{V}^{0,R}(t,x) = \inf_{w\in L_2} J^{0,R}(t,x,w)
\end{aligned}
$$

where X^ϵ and X^0 are given in (6.3).

In the first section of the proof, one applies the Barles and Perthame technique. Specifically, let

$$
\begin{aligned}
V^R_*(t,x) &\doteq \liminf_{\epsilon\downarrow 0,(s,y)\to(t,x)} \overline{V}^{\epsilon,R}(s,y)\\
V_*(t,x) &\doteq \liminf_{\epsilon\downarrow 0,(s,y)\to(t,x)} \overline{V}^{\epsilon}(s,y).
\end{aligned}
$$

One wishes to show that

$$
V^R_* \ge \overline{V}^{0,R}. \tag{6.4}
$$

Then, by a cone of dependence argument, one shows that $\overline{V}^{0,R} = \overline{V}^0$ on compact sets where the diameter of these sets goes to infinity as $R \to \infty$. From this, (6.4) and the fact that $V_* \ge V^R_*$ one obtains

$$
V_* \ge \overline{V}^0.
$$

The proof of (6.4) proceeds in two steps. First one obtains equicontinuity estimates in $\overline{V}^{\epsilon,R}$ near $t = 0$, and uses these to obtain

$$
V^R_*(0,x) = \overline{V}^{0,R}(0,x) = \phi(x) \qquad \forall\, x \in \Re. \tag{6.5}
$$

Then one needs to prove

$$
V^R_*(t,x) \ge \overline{V}^{0,R}(t,x) \qquad \forall\, t < 0, x \in \Re.
$$

Let $\alpha > 0$, $\delta > 0$ and $\beta < 0$. Suppose

$$
\begin{aligned}
\min_{t\le 0,x\in\Re}\Big\{&(1+\alpha)V^R_*(t,x) - \overline{V}^{0,R}(t,x)\\
&+ \delta\sqrt{1+x^2} + \beta t\Big\} < 0.
\end{aligned} \tag{6.6}
$$

Note that the existence of the minimum is guaranteed by the boundedness of V^R_* and $\overline{V}^{0,R}$. Also note that since the minimum is negative, it cannot occur at $t = 0$ by (6.5). Then note that V^R_* is a lower semi–continuous viscosity supersolution of (6.2) (with ϕ^R replacing ϕ), and $\overline{V}^{0,R}$ is a continuous viscosity solution of (6.2) (with ϕ^R replacing ϕ). Using viscosity solution arguments (similar to those of [DMc], [FS]), one can show that if (6.6) is true, then one has a contradiction for sufficiently small δ, α. Taking β also close to zero yields the result.

It remains to prove that

$$
V^*(t,x) \doteq \limsup_{\epsilon\downarrow 0,(s,y)\to(t,x)} \overline{V}^\epsilon(s,y) \le \overline{V}^0(t,x)
$$

for all $t \le 0$ and $x \in \Re$. This is obtained by applying a deterministic, optimal control, w^0, for the limit problem ($\overline{V}^0$) to the prelimit problem, and then showing that

$$
J^\epsilon(t,x,w^0) \to J^0(t,x,w^0) = \overline{V}^0(t,x)
$$

which leads to $V^* \le \overline{V}^0$.

7. References

[BDL] M. Bardi and F. Da Lio, On the Bellman equation for some unbounded control problems, Preprint No. 5, Dipartimento di Matematica, Universita degli Studi di Padova (1996).

[BJP] R. K. Boel, M. R. James and I. R. Petersen, preprint.

[CE] C. D. Charalambous and R. J. Elliott, "Finite–dimensional nonlinear output feedback dynamic games and bounds for sector nonlinearities", Submitted to IEEE Trans. Auto. Control.

[DMc] P. Dupuis and W. M. McEneaney, "Risk–sensitive and robust escape criteria", SIAM J. Control and Optim., (to appear).

[FMc] W. H. Fleming and W. M. McEneaney, "Risk sensitive control on an infinite time horizon", SIAM J. Control and Optim., **33** (1995), 1881–1915.

[FP] W. H. Fleming and E. Pardoux, "Optimal control of partially–observed diffusions", SIAM J. Control and Optim., **20** (1982), 261–285.

[FS] W. H. Fleming and H. M. Soner, Controlled Markov Processes and Viscosity Solutions, Springer-Verlag, New York, 1992.

[H] O. J. Hijab, "Minimum Energy Estimation", University of California, Berkeley, PhD Thesis, 1980.

[I] H. Ishii, "Comparison results for Hamilton–Jacobi equations without a growth condition on solutions from above", preprint.

[JB] M. R. James and J. S. Baras, "Partially observed differential games, infinite dimensional HJI equations, and nonlinear H_∞ control", SIAM J. Control and Optim., **34** (1996), 1342–1364.

[JBE] M. R. James, J. S. Baras and R. J. Elliott, "Risk–sensitive control and dynamic games for partially–observed discrete–time nonlinear systems", IEEE Trans. Auto. Control, **39**, (1994), 780–792.

[KD] A. J. Krener and A. Duarte, "A Hybrid Computational Approach to Nonlinear Estimation", 35th IEEE CDC, (1996).

[M] S. K. Mitter, Lectures on Nonlinear Filtering and Stochastic Control, Springer Lecture Notes in Math., **972**.

[Mc1] W. M. McEneaney, "Robust/H_∞ filtering for nonlinear systems", Submitted to Systems and Control Letters.

[Mc2] W. M. McEneaney, "Uniqueness for viscosity solutions of nonstationary HJB equations under some a priori conditions (with applications)", SIAM J. Control and Optim., **33** (1995) 1560–1576.

[Mc3] W. M. McEneaney, "Robust control and differential games on a finite time horizon", Math. of Controls Signals and Systems, **8** (1995), 138–166.

[MED] J. B. Moore, R. J. Elliott and S. Dey, "Risk sensitive generalizations of minimum variance estimation and control", To appear in J. Math. Systems, Estimation and Control.

[Mo] R. E. Mortensen, "Maximum Likelihood Recursive Nonlinear Filtering", J. Optim. Theory Appl., **2** (1968) 386–394.

[YJ] S. Yuliar and M. R. James, "On nonlinear H_∞ filtering", preprint.