Proceedings of the 26th Conference
on Decision and Control
Los Angeles, CA • December 1987

NONLINEAR FILTERING AND LARGE DEVIATIONS *

M. R. James J. S. Baras

Systems Research Center
University of Maryland
College Park MD 20742

Abstract

We consider the nonlinear filtering problem $dx = f(x)dt + \sqrt{\epsilon}dw$, $dy = h(x)dt + \sqrt{\epsilon}dv$, and obtain $\lim_{\epsilon\to 0} \epsilon \log q^\epsilon(x,t) = -W(x,t)$ for unnormalised conditional densities $q^\epsilon(x,t)$ using PDE methods. Here, $W(x,t)$ is the value function for a deterministic optimal control problem arising in Mortensen's deterministic estimation, and is the unique viscosity solution of a Hamilton–Jacobi–Bellman equation.

Introduction

An important problem in system theory is the construction of observers for nonlinear control systems. Baras, Bensoussan and James [1] have studied a method for constructing an observer as a limit of nonlinear filters for a family of associated filtering problems (2), parameterised by $\epsilon > 0$. It is of interest then to study the asymptotic behaviour of the corresponding unnormalised conditional densities $q^\epsilon(x,t)$ as $\epsilon \to 0$, via the Zakai equation (3). We obtain the asymptotic formula

$$q^\epsilon(x,t) = e^{-\frac{1}{\epsilon}(W(x,t)+o(1))}, \tag{1}$$

as $\epsilon \to 0$, where $W(x,t)$ is the value function corresponding to a deterministic optimal control problem, namely that arising in deterministic estimation.

Our method is inspired by the work of Fleming and Mitter [4], and Evans and Ishii [3]. A logarithmic transformation is applied to the robust form of the Zakai equation, yielding a Hamilton–Jacobi equation in the limit. A related Hamilton–Jacobi equation is interpreted as the Bellman equation for the optimal control problem arising in deterministic estimation, of which $W(x,t)$ is the unique viscosity solution. In particular, $W(x,t)$ is not assumed to be smooth.

This problem has been studied by Hijab [5] using different methods. Hijab also obtained a large deviation principle for conditional measures on $C([0,T];\mathbb{R}^n)$. An extension of his result is presented in James and Baras [6], which includes complete proofs of the results discussed in the present paper.

Problem Formulation

We consider a family of diffusion processes in $\mathbb{R}^n$ with real valued observations:

$$\begin{aligned} dx^\epsilon(t) &= f(x^\epsilon(t))dt + \sqrt{\epsilon}dw(t), \quad x^\epsilon(0) = x_0^\epsilon, \\ dy^\epsilon(t) &= h(x^\epsilon(t))dt + \sqrt{\epsilon}dv(t), \quad y^\epsilon(0) = 0. \end{aligned} \tag{2}$$

Here w, v are independent Wiener processes independent of the initial conditions x_0^ϵ, which have (unnormalised) densities $q_0^\epsilon(x) = C_\epsilon e^{-\frac{1}{\epsilon}S_0(x)}$ where $\lim_{\epsilon\to 0}\epsilon \log C_\epsilon = 0$ and $S_0 \geq 0$ is smooth and bounded. As $\epsilon \to 0$ the trajectories of (2) converge in probability to the trajectory of a corresponding deterministic system. We assume throughout the following: f, h are bounded C^∞ functions with bounded derivatives of orders 1 and 2.

The *Zakai equation* for an unnormalised conditional density $q^\epsilon(x,t)$ is

$$\begin{aligned} dq^\epsilon(x,t) &= A_\epsilon^* q^\epsilon(x,t) + \frac{1}{\epsilon}h(x)q^\epsilon(x,t)dy^\epsilon(t), \\ q^\epsilon(x,0) &= q_0^\epsilon(x), \end{aligned} \tag{3}$$

where A_ϵ^* is the formal adjoint of the diffusion operator. Defining

$$p^\epsilon(x,t) = \exp\left(-\frac{1}{\epsilon}y^\epsilon(t)h(x)\right) q^\epsilon(x,t), \tag{4}$$

the *robust* form of the Zakai equation is

$$\begin{aligned} \tfrac{\partial}{\partial t}p^\epsilon(x,t) - \frac{\epsilon}{2}\Delta p^\epsilon(x,t) + Dp^\epsilon(x,t)g^\epsilon(x,t) + \frac{1}{\epsilon}V^\epsilon(x,t)p^\epsilon(x,0) &= 0, \\ p^\epsilon(x,t) &= q_0^\epsilon(x). \end{aligned} \tag{5}$$

Note that (5) is a linear parabolic PDE and the coefficient V^ϵ depends on the observation path $t \mapsto y(t)$. We shall omit the ϵ–dependence of y, and view (5) as a functional of the observation path $y \in \Omega_0 = C([0,T], \mathbb{R}^n;\ y(0) = 0)$. This transformation provides a convenient choice of a version of the conditional density, and under our assumptions we can recover the unnormalised density $q^\epsilon(x,t)$ from the solution of (5).

Following Fleming and Mitter [4], who considered filtering problems with $\epsilon = 1$, we apply the logarithmic transformation

$$S^\epsilon(x,t) = -\epsilon \log p^\epsilon(x,t). \tag{6}$$

Then $S^\epsilon(x,t)$ satisfies

$$\begin{aligned} \tfrac{\partial}{\partial t}S^\epsilon(x,t) - \frac{\epsilon}{2}\Delta S^\epsilon(x,t) + H^\epsilon(x,t,DS^\epsilon(x,t)) &= 0, \\ S^\epsilon(x,0) &= S_0(x), \end{aligned} \tag{7}$$

where

$$H^\epsilon(x,t,\lambda) = \lambda g^\epsilon(x,t) + \frac{1}{2}\mid\lambda\mid^2 - V^\epsilon(x,t). \tag{8}$$

Equation (7) is a nonlinear parabolic PDE. Formally letting $\epsilon \to 0$ we obtain a Hamilton–Jacobi equation

$$\begin{aligned} \tfrac{\partial}{\partial t}S(x,t) + H(x,t,DS(x,t)) &= 0, \\ S(x,0) &= S_0(x), \end{aligned} \tag{9}$$

where

$$H(x,t,\lambda) = \lambda g_0(x,t) + \frac{1}{2} \mid \lambda \mid^2 - V(x,t), \qquad (10)$$

Note that $g^\epsilon \to g_0$, $V^\epsilon \to V$, and $H^\epsilon \to H$ uniformly on compact subsets. We shall interpret solutions of (9) in the viscosity sense. If we define

$$W(x,t) = S(x,t) - y(t)h(x), \quad y \in \Omega_0, \qquad (11)$$

then, for $y \in \Omega_0 \cap C^1$, $W(x,t)$ satisfies a Hamilton–Jacobi equation, which is presented as the Bellman equation for the deterministic estimation control problem below.

Deterministic Estimation

We begin by reviewing Mortensen's method [5] of deterministic minimum energy estimation. Given an observation record $\mathcal{Y}_t = \{y(s),\ 0 \le s \le t\}$, $0 \le t \le T$, of the deterministic system

$$\begin{aligned} \dot{x} &= f(x) + u, \quad x(0) = x_0, \\ \dot{y} &= h(x) + v, \quad y(0) = 0, \end{aligned} \qquad (12)$$

we wish to estimate the state at time t, the initial condition x_0 being unknown. Define

$$J_t(x_0,u) = S_0(x_0) + \int_0^t L(x(s),u(s),s)ds, \qquad (13)$$

where

$$L(x,u,s) = \frac{1}{2} \mid u \mid^2 + \frac{1}{2}h(x)^2 - \dot{y}(s)h(x). \qquad (14)$$

We now minimise J_t over pairs (x_0,u). The *deterministic* or minimum energy *estimate* $\hat{x}(t)$ given $\mathcal{Y}_t$ is defined to be the endpoint of the optimal trajectory $s \mapsto x^*(s)$, $0 \le s \le t$, corresponding to a minimum energy pair (x_0^*,u^*) : $\hat{x}(t) = x^*(t)$.

We use dynamic programming to study this problem. Define a *value function*

$$W(x,t) = \inf_{(x_0,u)} \{J_t(x_0,u) : x(0) = x_0,\ x(t) = x\}. \qquad (15)$$

By using standard methods, we see that $W(x,t)$ is continuous and formally satisfies the *Bellman equation*

$$\begin{aligned} \tfrac{\partial}{\partial t}W(x,t) + \tilde{H}(x,t,DW(x,t)) &= 0, \\ W(x,0) &= S_0(x), \end{aligned} \qquad (16)$$

where

$$\tilde{H}(x,t,\lambda) = \max_{u \in U} \{\lambda(f(x)+u) - L(x,u,t)\}. \qquad (17)$$

To obtain $\hat{x}(t)$, one minimises $W(x,t)$ over x. In fact, using the definition of viscosity solutions in Crandall, Evans and Lions [2], we can prove:

Theorem *The value function $W(x,t)$ defined by* (15) *is the unique viscosity solution of the Hamilton-Jacobi-Bellman equation* (16). *In addition, the function $S(x,t)$ defined by* (6) *is the unique viscosity solution of the Hamilton-Jacobi equation* (9).

Some Estimates

Let $S^\epsilon(x,t)$ be the solution of (7). The following estimates are used to prove that $S^\epsilon \to S$.

Theorem *For every compact subset $Q \subset I\!R^n \times [0,T]$, there exists $\epsilon_0 > 0$ and $K > 0$ such that for $0 < \epsilon < \epsilon_0$ we have*

$$\begin{aligned} \mid S^\epsilon(x,t) \mid &\le K, \quad \text{for all } (x,t) \in Q, \qquad (18) \\ \mid DS^\epsilon(x,t) \mid &\le K, \quad \text{for all } (x,t) \in Q. \qquad (19) \end{aligned}$$

To prove (18), we use a comparison theorem which depends on the maximum principle for linear parabolic PDE. The gradient estimate (19) uses a variant of the techniques presented in Evans and Ishii [3], as suggested to us by L. C. Evans.

Main Result

We are now in a position to state and prove our main result.

Theorem *Under the above assumptions, we have*

$$\lim_{\epsilon \to 0} \epsilon \log q^\epsilon(x,t) = -W(x,t) \qquad (20)$$

uniformly on compact subsets of $I\!R^n \times [0,T]$, where $W(x,t)$ is defined by (11).

Proof: From the above estimates and the Arzela–Ascoli theorem, there is a subsequence $\epsilon_k \to 0$ such that S^{ϵ_k} converges uniformly on compact subsets to a continuous function $\tilde{S}$. By the "vanishing viscosity" theorem [3], $\tilde{S}$ is a viscosity solution of (9). By uniqueness, $\tilde{S} = S$. In fact, $S^\epsilon \to S$ as $\epsilon \to 0$.

From this we have

$$\lim_{\epsilon \to 0} \epsilon \log q^\epsilon(x,t) = -(S(x,t) - y(t)h(x))$$

uniformly on compact subsets, for $y \in \Omega_0$. Using the definition (11) of $W(x,t)$ completes the proof.

References

[1] J. S. Baras, A. Bensoussan and M. R. James, *Dynamic Observers as Asymptotic Limits of Recursive Filters: Special Cases*, SRC–TR–86–79, Systems Research Center, Univ. of MD, Dec. 1986. To appear, SIAM J. Applied Math.

[2] M. G. Crandall, L. C. Evans and P. L. Lions, *Some Properties of Viscosity Solutions of Hamilton–Jacobi Equations*, Trans. AMS, 282, No.2, April 1984, pp487–502.

[3] L. C. Evans and H. Ishii, *A PDE approach to some asymptotic problems concerning random differential equations with small noise intensities*, Ann. Inst. Henri Poincaré–Analyse non linéire, 2, No.1, 1985, pp1–20.

[4] W. H. Fleming and S. K. Mitter, *Optimal Control and Nonlinear Filtering for Nondegenerate Diffusion Processes*, Stochastics, 8, 1982, pp63–77.

[5] O. Hijab, *Minimum Energy Estimation*, PhD Dissertation, University of California, Berkely, December, 1980.

[6] M. R. James and J. S. Baras, *Nonlinear Filtering and Large Deviations: A PDE-Control Theoretic Approach*, SRC–TR–87–27, Systems Research Center, Univ. of MD, Feb. 1987.

* Research partially supported by NSF grant CDR-85-00108, and AFOSR contract AFOSR-870073.