

New Finite-Dimensional Risk-Sensitive Filters: Small Noise Limits

Charalambos D. Charalambous, Subhrakanti Dey,
and Robert J. Elliott

Abstract— This paper is concerned with continuous-time nonlinear risk-sensitive filters. It is shown that for large classes of nonlinearities entering both the dynamics and measurements, these filters are finite-dimensional generalizations of the Benes filters. Specific examples are discussed. The small noise limiting analog is discussed using change of probability measures.

Index Terms—Filtering, finite-dimensional, risk-sensitive.

I. INTRODUCTION

Risk-sensitive estimation optimizes the expectation of an exponential of quadratic (or a more general convex) cost criterion. Compared with minimum variance estimation, which achieves the minimization of a quadratic error criteria, risk-sensitive estimation robustifies the estimator against noise uncertainties by penalizing all the higher order moments of the estimation error energy. For example, it has been shown in [2], that discrete-time risk-sensitive filters for hidden Markov models with finite-discrete states perform better than standard filters in situations involving uncertainties in the noise statistics. Also, in the small noise limit, risk-sensitive problems are shown to be related to estimation/control problems, formulated as deterministic minimax dynamic games (\mathcal{H}^∞ estimation/control problems for linear/nonlinear systems) [3]–[9].

It is well known that the conditional density of continuous-time partially observed stochastic systems satisfies the Duncan–Mortensen–Zakai (DMZ) equation. Aside from the case of linear systems, solutions of this equation are rare. In fact, the only success is that noted by Benes [1] and the subsequent generalizations again with linear measurements.

Motivated by the finite-dimensional solutions of the DMZ equation and its Feynman–Kac version, it was recently noted independently by Charalambous–Naidu–Moore [10], Charalambous [11], and Bensoussan–Elliott [12] that the existence of finite-dimensional filters for nonlinear systems is closely related to the existence of finite-dimensional controllers for nonlinear partially observed stochastic optimal control problems. The most general extensions are given in [13], where the nonlinearities are also allowed to enter the measurements.

The main result of this paper is the derivation of finite-dimensional risk-sensitive filters when nonlinearities enter the dynamics and measurement equations. In all the existing literature on nonlinear filtering, finite-dimensional filter estimates are computed based on the model of Benes (see [1]), which assumes nonlinear dynamics

Manuscript received December 30, 1996. The work of C. D. Charalambous was supported by the Natural Sciences and Engineering Research Council of Canada under Grant OGP0183720.

C. D. Charalambous is with the Department of Electrical Engineering, McGill University, Montreal, PQ, H3A 2A7 Canada (e-mail: chadcha@cim.mcgill.ca).

S. Dey is with the Department of Systems Engineering, Research School of Information Sciences and Engineering, Australian National University, Canberra ACT 0200, Australia.

R. J. Elliott is with the Department of Mathematical Sciences, University of Alberta, Edmonton, Alberta, Canada T6G 2G1.

Publisher Item Identifier S 0018-9286(98)06601-X.

and linear measurements. Here we show that when the measurements include the increments of the state process, and are thus nonlinear, the risk-sensitive filter is finite-dimensional. In principle, this is expected from the finite-dimensionality of the analogous control problems treated in [13]. However, because the formulation of the risk-sensitive cost criterion is different from the formulation of the information state of [13], and finite-dimensional filters are essentially nonexistent, these filters should be treated separately. Another reason for treating these filters separately is their connection to analogous finite-dimensional minimax filters, which we explore in this paper.

In Section II, we introduce the class of signal models we are dealing with and formulate the risk-sensitive filtering problem. Using change of probability measure techniques, an unnormalized measure (also known as the information state), is defined that satisfies a Feynman–Kac version of the DMZ equation. The optimal risk-sensitive estimate is given as the minimizing argument of an integral involving this information state. This minimization can be recast as a pointwise minimization by employing dynamic programming to derive a Hamilton–Jacobi equation with infinite-dimensional state space. Section III presents finite-dimensional risk-sensitive filtering results. Section IV includes some discussions on small noise limits which enable us to obtain finite-dimensional deterministic minimax filters.

II. PROBLEM STATEMENT AND METHODOLOGY

Let $[0, T]$, $T \in \mathbb{R}$ denote the time interval, and let $(\Omega, \mathcal{A}, \tilde{P})$ be a complete probability space that carries the standard Wiener processes $\{w_t; t \in [0, T]\}$, $\{b_t; t \in [0, T]\}$ which take values in \mathbb{R}^m , \mathbb{R}^d , respectively, and the random variable $x(0)$. Throughout this paper it is assumed that these Wiener processes and $x(0)$ are mutually independent under measure \tilde{P} .

On the probability space $(\Omega, \mathcal{A}, \tilde{P})$, the processes $\{x_t; t \in [0, T]\}$, $\{y_t; t \in [0, T]\}$ are solutions of the stochastic differential equations

$$dx_t = f(t, x_t) dt + G_t dw_t, \quad x(0) \in \mathbb{R}^n \quad (1)$$

$$dy_t = h(t, x_t) dt + \alpha_t dx_t + N_t^{1/2} db_t \\ y(0) = 0 \in \mathbb{R}^d. \quad (2)$$

Here $t \rightarrow G_t$, α_t , $N_t^{1/2}$ are known $n \times m$, $d \times n$, $d \times d$, real-valued matrices, respectively, continuous in t , and $G_t G_t' \geq 0$; $N_t \doteq N_t^{1/2} N_t^{1/2}{}' > 0$. Also, $|f(t, x)| \leq K(1 + |x|)$, $|h(t, x)| \leq K(1 + |x|)$, and $f(\cdot, x)$, $h(\cdot, x)$ are continuous in t . $\{x_t; t \in [0, T]\}$ is a Markov diffusion process, and so for any C_x^2 function $\Phi: \mathbb{R}^n \rightarrow \mathbb{R}$, with compact support, its generator is defined by

$$A(t)\Phi(x) \doteq \frac{\text{Tr}}{2} (G_t G_t' D_x^2 \Phi(x)) + f(t, x) \cdot D_x \Phi(x) \quad (3)$$

where $D_x \doteq (\partial/\partial x_1 \partial/\partial x_2 \cdots \partial/\partial x_n)'$ and $\alpha \cdot \beta \doteq \alpha' \beta$ denotes the scalar product in \mathbb{R}^n .

x_t is the state of the system at time t , which is not directly measurable. Rather, the only information available for x_t is through the observations process $\{y_s; 0 \leq s \leq t < T\}$. Unlike the minimum variance estimation problem which minimizes the expected value of the square of the error at time t given $\{y_s; 0 \leq s \leq t\}$, the problem to be discussed here is the so-called risk-sensitive estimation problem. This aims to minimize the expected value of an exponential-of-integral of the error, given the past and present measurements

$\{y_s; 0 \leq s \leq t < T\}$. More precisely, letting $\{\mathcal{F}_{0,t}^y; t \in [0, T]\}$ denote the complete filtration generated by the observation σ -algebras $\sigma\{y_s; 0 \leq s \leq t\}$, the risk-sensitive estimation problem is to find a random process which is a nonanticipative functional of the observations process namely, $\{\hat{x}_t^* = \mu^*(t, \{y_s; 0 \leq s \leq t\})\}$ $t \in [0, T]$ such that

$$\hat{x}^*(\cdot) \doteq \arg \min_{\hat{x}(\cdot)} J(\hat{x}(\cdot))$$

where

$$J(\hat{x}(\cdot)) = \tilde{E} \left\{ \exp \frac{\theta}{2} \left(\int_0^t |Q_s^{1/2}(x_s - \hat{x}_s)|^2 ds \right) \right\} \\ Q_t = Q_t' \doteq \tilde{Q}_t^{1/2} Q_t^{1/2'} > 0. \quad (4)$$

Here Q is Borel measurable and bounded, $\theta > 0$, and \tilde{E} denotes expectation with respect to measure \tilde{P} . That is, for each $t > 0$, an admissible risk-sensitive estimator is $\mathcal{F}_{0,t}^y$ measurable. Thus, if there exists an $\mathcal{F}_{0,t}^y$ measurable density function $\tilde{q}_t^\theta \equiv \tilde{q}^\theta(x, t)$ such that $\tilde{E} \left\{ \Phi(x_t) \exp(\theta/2) \left(\int_0^t |Q_s^{1/2}(x_s - \hat{x}_s)|^2 ds \right) | \mathcal{F}_{0,t}^y \right\} = \int_{\mathbb{R}^n} \Phi(x) \tilde{q}^\theta(x, t) dx$, then (4) is expressed as

$$J(\hat{x}(\cdot)) = \tilde{E} \left\{ \tilde{E} \left\{ \exp \frac{\theta}{2} \left(\int_0^t |Q_s^{1/2}(x_s - \hat{x}_s)|^2 ds \right) | \mathcal{F}_{0,t}^y \right\} \right\} \\ = \tilde{E} \left\{ \int_{\mathbb{R}^n} \tilde{q}^\theta(x, t) dx \right\}. \quad (5)$$

In our case such density functions exist and they are unique provided θ is sufficiently small. Note that \tilde{q}_t^θ is by construction an infinite-dimensional information state and so formally we may write $\hat{x}_t = \mu(t, \{y_s; 0 \leq s \leq t\}) = \tilde{\mu}(t, \tilde{q}_t^\theta)$, $\forall t \in [0, T]$. Moreover, if \tilde{q}_t^θ is expressed in terms of a finite number of quantities such that the dependence of \tilde{q}_t^θ on $\{\hat{x}_s, s \leq t\}$ is only through them, then these finite-number quantities are sufficient statistics for the risk-sensitive estimation problem; that is, $\{\hat{x}_t^*; t \in [0, T]\}$ is a finite-dimensional risk-sensitive estimator, which is a functional of the sufficient statistics.

Toward introducing an equation governing \tilde{q}_t^θ , or some version of it, recall that (2) is also equivalent to $dy_t = (h(t, x_t) + \alpha_t f(t, x_t)) dt + \alpha_t G_t dw_t + N_t^{1/2} db_t$, so we may conveniently rewrite

$$dy_t = \tilde{h}(t, x_t) dt + \tilde{\alpha}_t dw_t + N_t^{1/2} db_t \\ \tilde{h}(t, x) = h(t, x) + \alpha_t f(t, x), \quad \tilde{\alpha}_t = \alpha_t G_t. \quad (6)$$

Certainly, we shall need results for correlated risk-sensitive problems; these are drawn from [11] and [15]. To this end introduce

$$\tilde{C}_t = \tilde{\alpha}_t \tilde{\alpha}_t' + N_t, \quad \tilde{D}_t = I_m - \tilde{\alpha}_t' \tilde{C}_t^{-1} \tilde{\alpha}_t. \quad (7)$$

From the linear growth of $f(t, x)$ and $h(t, x)$, (assuming $x(0)$ has finite second moment), the measure \tilde{P} (under which the system (1) and (2) is defined) is absolutely continuous with the measure P defined by the system

$$dx_t = (f(t, x_t) - G_t \tilde{\alpha}_t' \tilde{C}_t^{-1} \tilde{h}(t, x_t)) dt + G_t \tilde{D}_t^{1/2} d\tilde{w}_t \\ + G_t \tilde{\alpha}_t' \tilde{C}_t^{-1/2} d\tilde{y}_t, \quad x(0) \in \mathbb{R}^n \quad (8)$$

$$dy_t = \alpha_t G_t dw_t + N_t^{1/2} db_t, \quad y(0) = 0 \in \mathbb{R}^d \quad (9)$$

where

$$d\tilde{y}_t \doteq \tilde{C}_t^{-1/2} dy_t \\ d\tilde{w}_t \doteq \tilde{D}_t^{-1/2} \{dw_t - \tilde{\alpha}_t' \tilde{C}_t^{-1} dy_t\}. \quad (10)$$

In fact the Radon Nikodym derivative on $\mathcal{F}_{0,t}^y \times \mathcal{F}_{0,t}^w \times \mathcal{F}^{x(0)}$ ($\mathcal{F}_{0,t}^w \doteq \sigma\{w(s); 0 \leq s \leq t\}$, $\mathcal{F}^{x(0)} \doteq \sigma\{x(0)\}$) is given by

$$\frac{d\tilde{P}}{dP} \doteq \Lambda_{0,t} \\ = \exp \left\{ \int_0^t \tilde{h}(s, x_s) \cdot \tilde{C}_s^{-1/2} d\tilde{y}_s \right. \\ \left. - \frac{1}{2} \int_0^t |\tilde{C}_s^{-1/2} \tilde{h}(s, x_s)|^2 ds \right\}. \quad (11)$$

Note that under measure P , $\{\tilde{y}_t, t \in [0, T]\}$ and $\{\tilde{w}_t, t \in [0, T]\}$ are independent standard Wiener processes. Consequently, denoting E the expectation with respect to measure P we have

$$J(\hat{x}(\cdot)) = \tilde{E} \left\{ \exp \frac{\theta}{2} \left(\int_0^t |Q_s^{1/2}(x_s - \hat{x}_s)|^2 ds \right) \right\} \\ = E \left\{ \Lambda_{0,t} \exp \frac{\theta}{2} \left(\int_0^t |Q_s^{1/2}(x_s - \hat{x}_s)|^2 ds \right) \right\} \\ = E \left\{ E \left\{ \Lambda_{0,t} \exp \frac{\theta}{2} \right. \right. \\ \left. \left. \cdot \left(\int_0^t |Q_s^{1/2}(x_s - \hat{x}_s)|^2 ds \right) | \mathcal{F}_{0,t}^y \right\} \right\}. \quad (12)$$

Define

$$q^\theta(x, t) dx \doteq E \left[I_{x_t \in dx} \Lambda_{0,t} \exp \frac{\theta}{2} \right. \\ \left. \cdot \left(\int_0^t |Q_s^{1/2}(x_s - \hat{x}_s)|^2 ds \right) | \mathcal{F}_{0,t}^y \right]. \quad (13)$$

Then by (5), (12), and (13) we have the alternative representation

$$J(\hat{x}(\cdot)) = \tilde{E} \left\{ \int_{\mathbb{R}^n} \tilde{q}^\theta(x, t) dx \right\} \\ = E \left\{ \int_{\mathbb{R}^n} q^\theta(x, t) dx \right\}. \quad (14)$$

Here $q_t^\theta \equiv q^\theta(x, t)$ is the unnormalized version of the conditional density $\tilde{q}^\theta(x, t)$, which is a solution of the Feynman-Kac version of the DMZ equation (see [11] and [15]), namely

$$dq^\theta(x, t) = A(t)^* q^\theta(x, t) dt + \frac{\theta}{2} |Q_t^{1/2}(x - \hat{x}_t)|^2 q^\theta(x, t) dt \\ + \tilde{h}(t, x) q^\theta(x, t) \cdot \tilde{C}_t^{-1} dy_t \\ - D_x \cdot (q^\theta(x, t) G_t \tilde{\alpha}_t') \tilde{C}_t^{-1} dy_t, \\ (x, t) \in \mathbb{R}^n \times (0, T], \quad (15)$$

$$q^\theta(x, 0) = q_0^\theta(x), \quad x \in \mathbb{R}^n. \quad (16)$$

Here $A(t)^*$ is the formal adjoint of $A(t)$ defined in (3).

III. FINITE-DIMENSIONAL RISK-SENSITIVE FILTERS

In this section, we present the finite-dimensional risk-sensitive filtering results. Note that under uniqueness of solutions of (15) and (16), the finite-dimensionality of the information state implies finite-dimensionality of the risk-sensitive estimate $\hat{x}(\cdot)$. Define

$$C_t \doteq \alpha_t \alpha_t' + N_t, \quad Q_t^\theta \doteq Q_t - \frac{1}{\theta} \tilde{Q}_t \\ \tilde{Q}_t = \tilde{Q}_t', \quad m_t^\theta \doteq -\frac{1}{\theta} \tilde{m}_t, \quad \delta^\theta \doteq -\frac{1}{\theta} \tilde{\delta}_t$$

where $\tilde{Q} = \tilde{Q}'$, \tilde{m} are $n \times n$, $1 \times n$, Borel measurable, bounded, real-valued matrices, respectively, and $\tilde{\delta}$ is a Borel measurable, bounded, real-valued function.

Theorem 3.1: Consider the risk-sensitive estimation problem

$$dx_t = (F_t x_t + g(t, x)) dt + G_t dw_t, \quad x(0) \in \mathfrak{R}^n \quad (17)$$

$$dy_t = H_t x_t dt + h(t, x) dt + \alpha_t dw_t + N_t^{1/2} db_t$$

$$y(0) = 0 \in \mathfrak{R}^d \quad (18)$$

where we have the following.

A1) The nonlinear drift and observation functions are gradients of some potential function

$$g(t, x) = G_t G_t' D_x \phi(x, t)$$

$$h(t, x) = \alpha_t G_t' D_x \phi(x, t)$$

$$\phi \in C_{x,t}^{2,1}(\mathfrak{R}^n \times [0, T]) \quad (19)$$

and $\phi_t \equiv \phi(x, t)$ is a solution of the second-order partial differential equation (PDE)

$$\frac{\partial}{\partial t} \phi_t + \frac{1}{2} \text{Tr}(G_t G_t' D_x^2 \phi_t) + \frac{1}{2} |G_t' D_x \phi_t|^2 + F_t x \cdot D_x \phi_t$$

$$= \frac{1}{2} (\tilde{Q}_t x \cdot x + 2 \tilde{m}_t x + \tilde{\delta}_t). \quad (20)$$

A2) The random variable $x(0)$ has density

$$q_0^\theta(x) = \exp(\phi(x, 0)) \times \hat{q}_0^\theta(x)$$

$$\hat{q}_0^\theta(x) = \frac{\exp(-P_0^{-1}(x - \xi) \cdot (x - \xi))}{(2\pi)^{n/2} |P_0|^{1/2}}, \quad P_0 = P_0' \geq 0. \quad (21)$$

Suppose there exists some $0 < \theta \leq \theta^*$ such that $H_t' C_t^{-1} H_t - \theta Q_t^\theta \geq 0$, $\forall t \in [0, T]$. Then a solution of the Feynman-Kac information state equation is given by

$$q_t^\theta = \exp(\phi(x, t)) \times \frac{\exp\left(-\frac{1}{2} |P_t^{-1/2}(x - r_t(\hat{x}))|^2\right)}{(2\pi)^{n/2} |P_t|^{1/2}}$$

$$\times \exp(c_{0,t}(\hat{x}) + \lambda_{0,t}(\hat{x}))$$

$$= \exp(\phi(x, t)) \times \hat{q}_t^\theta \quad (22)$$

where $r: [0, T] \times \Omega \rightarrow \mathfrak{R}^n$, $P = P': [0, T] \rightarrow \mathfrak{R}^{n \times n}$, $P_t \geq 0$, $\forall t \in [0, T]$, $c: [0, T] \times \Omega \rightarrow \mathfrak{R}$, $\lambda: [0, T] \times \Omega \rightarrow \mathfrak{R}$ are the sufficient statistics given by

$$dr_t(\hat{x}) = (F_t + \theta P_t Q_t^\theta) r_t(\hat{x}) dt - \theta P_t Q_t \hat{x}_t dt + \theta P_t m_t^{\theta, \prime} dt$$

$$+ (P_t H_t' + G_t \alpha_t') C_t^{-1} (dy_t - H_t r_t(\hat{x}) dt)$$

$$r(0) = \xi \quad (23)$$

$$\dot{P}_t = F_t P_t + P_t F_t' + \theta P_t Q_t^\theta P_t$$

$$- (P_t H_t' + G_t \alpha_t') C_t^{-1} (H_t P_t + \alpha_t G_t')$$

$$+ G_t G_t', \quad P(0) = P_0 \quad (24)$$

$$c_{0,t}(\hat{x}) = \int_0^t H_s r_s(\hat{x}) \cdot C_s^{-1} dy_s$$

$$- \frac{1}{2} \int_0^t |C_s^{-1/2} H_s r_s(\hat{x})|^2 ds \quad (25)$$

$$\lambda_{0,t}(\hat{x}) = \frac{\theta}{2} \int_0^t [Q_s^\theta r_s(\hat{x}) \cdot r_s(\hat{x}) + Q_s \hat{x}_s \cdot \hat{x}_s - 2Q_s \hat{x}_s \cdot r_s(\hat{x})$$

$$+ 2m_s^\theta \cdot r_s(\hat{x}) + \delta_s^\theta + \text{Tr}(P_s Q_s^\theta)] ds. \quad (26)$$

Moreover, letting

$$\hat{\varphi}(r, t) = \frac{2}{\theta} \log \int_{\mathfrak{R}^n} \exp(\phi(x, t))$$

$$\times \frac{\exp\left(-\frac{1}{2} |P_t^{-1/2}(x - r)|^2\right)}{(2\pi)^{n/2} |P_t|^{1/2}} dx \quad (27)$$

the risk-sensitive estimation problem is equivalent to finding an $\hat{x}^*(\cdot)$ minimizing the expression

$$J(\hat{x}(\cdot)) = \tilde{E} \left\{ \int_{\mathfrak{R}^n} \tilde{q}^\theta(x, t) dx \right\}$$

$$= E \left\{ \int_{\mathfrak{R}^n} \exp(\phi(x, t)) \times \hat{q}^\theta(x, t) dx \right\}$$

$$= \hat{E} \left\{ \exp \frac{\theta}{2} (\hat{\varphi}(r_t(\hat{x}), t)) \times \exp(\lambda_{0,t}(\hat{x})) \right\} \quad (28)$$

where under measure \hat{P} , $\hat{b}_t \doteq y_t - \int_0^t H r_s(\hat{x}) ds$ is the innovations process.

If, instead of (18), we consider observations $dy_t = H_t x_t dt + \alpha_t dx_t + N_t^{1/2} db_t$, then the following modifications need to be introduced in the above equations:

$$\alpha_t \rightarrow \alpha_t G_t, \quad H_t \rightarrow H_t + \alpha_t F_t. \quad (29)$$

Remark 3.2: Let $g(t, x) = 0$, $h(t, x) = 0$ in Theorem 3.1 so (17) and (18) specialize to the linear case

$$dx_t = F_t x_t dt + G_t dw_t, \quad x(0) \in \mathfrak{R}^n$$

$$dy_t = H_t x_t dt + \alpha_t dx_t + N_t^{1/2} db_t, \quad y(0) = 0 \in \mathfrak{R}^d.$$

Then, q_t^θ is given by (22) with $\phi(x, t) = 0$, $\tilde{Q}_t = 0$, $\tilde{m}_t = 0$, and $\tilde{\delta}_t = 0$, and, therefore by (28), the risk-sensitive estimator $\hat{x}^*(\cdot)$ is obtained by minimizing the expression

$$J(\hat{x}(\cdot)) = \hat{E} \left\{ \exp \frac{\theta}{2} \left(\int_0^t [|Q_s^{1/2}(r_s(\hat{x}) - \hat{x}_s)|^2 \right. \right.$$

$$\left. \left. + \text{Tr}(P_s Q_s) \right] ds \right\}. \quad (30)$$

Consequently, $\hat{x}_t^* = r_t$, $\forall t \in [0, T]$, where

$$d\hat{x}_t^* = F_t \hat{x}_t^* dt + (P_t H_t' + G_t \alpha_t') C_t^{-1} \hat{b}_t$$

$$d\hat{b}_t = dy_t - H_t \hat{x}_t^* dt, \quad \hat{x}^*(0) = \xi \quad (31)$$

$$\dot{P}_t = F_t P_t + P_t F_t' + \theta P_t Q_t P_t$$

$$- (P_t H_t' + G_t \alpha_t') C_t^{-1} (H_t P_t + \alpha_t G_t')$$

$$+ G_t G_t', \quad P(0) = P_0. \quad (32)$$

These are precisely the filtering equations of deterministic linear-quadratic minimax filtering problems, with L^2 additive disturbances.

Remark 3.3: Consider the scalar risk-sensitive estimation problem

$$dx_t = f(x_t) dt + dw_t, \quad x(0) \in \mathfrak{R}$$

$$dy_t = x_t dt + dx_t + db_t, \quad y(0) = 0 \in \mathfrak{R}$$

when \hat{x}^* is a law minimizing (4), and $f(x)$ is a solution of the Riccati equation $D_x f(x) + f(x)^2 = \text{quadratic in } x$. Following Benes [1], let $f(x) = D_x \phi(x)$, $\phi(x) = \log w(x)$. By (20)

$$\frac{1}{2} D_x^2 \phi(x) + \frac{1}{2} D_x \phi(x)^2 = \frac{1}{2} (\tilde{Q} x^2 + 2\tilde{m} x + \tilde{\delta})$$

$$D_x^2 w(x) = w(x) (\tilde{Q} x^2 + 2\tilde{m} x + \tilde{\delta}).$$

Hence, if $w(x) > 0$ then $f(x) = D_x w(x)/w(x)$, and Theorem 3.1 applies. The function

$$f(x) = \frac{A \exp(x) - B \exp(-x)}{A \exp(x) + B \exp(-x)}, \quad A \in \mathfrak{R}, B \in \mathfrak{R} \quad (33)$$

satisfies $D_x f(x) + f(x)^2 = 1$, with $\tilde{\delta} = 1$, $\tilde{Q} = 0$, $\tilde{m} = 0$, so Theorem 3.1 holds with $F_t = \tilde{Q}_t = \tilde{m}_t = 0$, $H_t = 1$, $\alpha_t = G_t = N_t = 1$, $\tilde{\delta}_t = 1$. For $A = B$, $f(x) = \tanh(x)$. Thus, Theorem 3.1 generalizes the class of finite-dimensional Benes filters (see [1]), from observations $dy_t = x_t dt + db_t$, to the more general case $dy_t = x_t dt + dx_t + db_t$.

Proof of Theorem 3.1: The equation governing q_t^θ is

$$\begin{aligned} dq_t^\theta &= \frac{1}{2} \text{Tr}(G_t G_t' D_x^2 q_t^\theta) dt - D_x \cdot (q_t^\theta (F_t x + g(t, x))) dt \\ &+ \frac{\theta}{2} |Q_t^{1/2} (x - \hat{x}_t)|^2 q_t^\theta dt + (H_t x + h(t, x)) q_t^\theta \cdot C_t^{-1} dy_t \\ &- D_x \cdot (q_t^\theta G_t \alpha_t') C_t^{-1} dy_t, \quad q^\theta(x, 0) = q_0^\theta(x). \end{aligned} \quad (34)$$

Introduce the gauge transformation

$$\hat{q}_t^\theta = \exp(-\phi(x, t)) \times q_t^\theta. \quad (35)$$

Substituting (35) into (34) we easily deduce the following equation governing $\hat{q}^\theta(\cdot)$:

$$\begin{aligned} d\hat{q}_t^\theta &= \frac{1}{2} \text{Tr}(G_t G_t' D_x^2 \hat{q}_t^\theta) dt - D_x \cdot (F_t x \hat{q}_t^\theta) dt \\ &+ \frac{\theta}{2} |Q_t^{1/2} (x - \hat{x}_t)|^2 \hat{q}_t^\theta dt + H_t x \hat{q}_t^\theta \cdot C_t^{-1} dy_t \\ &- D_x \cdot (\hat{q}_t^\theta G_t \alpha_t') C_t^{-1} dy_t \\ &+ [G_t G_t' D_x \phi_t - g(t, x)] \cdot D_x \hat{q}_t^\theta dt \\ &+ [h(t, x) - \alpha_t G_t' D_x \phi_t] \hat{q}_t^\theta \cdot C_t^{-1} dy_t \\ &+ \left[-\frac{\partial}{\partial t} \phi_t + \frac{1}{2} \text{Tr}(G_t G_t' D_x^2 \phi_t) \right. \\ &\quad \left. + \frac{1}{2} D_x \phi_t \cdot G_t G_t' D_x \phi_t - F_t x \cdot D_x \phi_t \right. \\ &\quad \left. - g(t, x) \cdot D_x \phi_t - \text{Tr}(D_x g(t, x)) \right] \hat{q}_t^\theta dt, \end{aligned}$$

$$\hat{q}^\theta(x, 0) = \hat{q}_0^\theta(x).$$

By A1), this equation reduces to

$$\begin{aligned} d\hat{q}_t^\theta &= \frac{1}{2} \text{Tr}(G_t G_t' D_x^2 \hat{q}_t^\theta) dt - D_x \cdot (F_t x \hat{q}_t^\theta) dt \\ &+ \frac{\theta}{2} |Q_t^{1/2} (x - \hat{x}_t)|^2 \hat{q}_t^\theta dt \\ &+ H_t x \hat{q}_t^\theta \cdot C_t^{-1} dy_t - D_x \cdot (G_t \hat{q}_t^\theta \alpha_t') C_t^{-1} dy_t \\ &+ \left[-\frac{\partial}{\partial t} \phi_t - \frac{1}{2} \text{Tr}(G_t G_t' D_x^2 \phi_t) - \frac{1}{2} |G_t' D_x \phi_t|^2 \right. \\ &\quad \left. - F_t x \cdot D_x \phi_t \right] \hat{q}_t^\theta dt. \end{aligned} \quad (36)$$

Since the potential term of the above equation is a quadratic function of x , due to (20), we obtain

$$\begin{aligned} d\hat{q}_t^\theta &= \frac{1}{2} \text{Tr}(G_t G_t' D_x^2 \hat{q}_t^\theta) dt - D_x \cdot (F_t x \hat{q}_t^\theta) dt \\ &+ \frac{\theta}{2} \left(|Q_t^{1/2} (x - \hat{x}_t)|^2 - \frac{1}{\theta} \hat{Q}_t^\theta x \cdot x \right. \\ &\quad \left. - \frac{1}{\theta} 2\tilde{m}_t x - \frac{1}{\theta} \tilde{\delta}_t \right) \hat{q}_t^\theta dt \\ &+ H_t x \hat{q}_t^\theta \cdot C_t^{-1} dy_t - D_x \cdot (\hat{q}_t^\theta G_t \alpha_t') C_t^{-1} dy_t \end{aligned}$$

$$\hat{q}^\theta(x, 0) = \hat{q}_0^\theta(x). \quad (37)$$

When $\hat{Q}_t = 0$, $\tilde{m}_t = 0$, $\tilde{\delta}_t = 0$, (37) is the Feynman–Kac information state equation of the risk-sensitive filtering problem $dx_t = F_t x_t dt + G_t dw_t$, $dy_t = H_t x_t dt + \alpha_t dw_t + N_t^{1/2} db_t$, where $x(0)$ is a Gaussian random variable and, thus, an explicit solution is available (see [11] and [15]). When $\hat{Q}_t \neq 0$, $\tilde{m}_t \neq 0$, and $\tilde{\delta}_t \neq 0$, these terms contribute to the quadratic potential term and so (37) is again explicitly solvable. By the invertibility of the gauge transformation an explicit representation for q_t^θ is obtained and this can be verified by substituting (22) into (34).

The last equality in (28) is obtained by introducing the Radon–Nikodym derivative $d\hat{P}/dP = \exp(c_t)$ on $\mathcal{F}_{0,t}^y$. Note that \hat{P} is also absolutely continuous with respect to \tilde{P} when restricted $\mathcal{F}_{0,t}^y$. \square

Certainly, the finite-dimensional minimization problem (28), subject to (23), can be solved using the martingale approach introduced by Kumar and van Schuppen. Alternatively, one can derive a Hamilton–Jacobi equation satisfied by the optimal risk-sensitive conditional cost (see [11]).

Lemma 3.4: Let

$$W(x, t) = \exp \phi(x, t)$$

then $W_t \equiv W(x, t)$ is a solution of the PDE

$$\begin{aligned} \frac{\partial W_t}{\partial t} + \frac{\text{Tr}}{2} (G_t G_t' D_x^2 W_t) + F_t x \cdot D_x W_t \\ = \frac{1}{2} W_t \{ \hat{Q}_t x \cdot x + 2\tilde{m}_t x + \tilde{\delta}_t \}. \end{aligned} \quad (38)$$

Define

$$\begin{aligned} \Gamma(x, t) &\doteq \frac{1}{2} \Delta_t x \cdot x + x \cdot \zeta_t + \eta_t, \\ \Delta: [0, T] &\rightarrow \mathbb{R}^{n \times n}, \quad \zeta: [0, T] \rightarrow \mathbb{R}^n, \quad \eta: [0, T] \rightarrow \mathbb{R}. \end{aligned}$$

If there exists a $0 < \theta \leq \theta^*$ such that $H_t' C_t^{-1} H_t - \theta Q_t^\theta \geq 0$, $\forall t \in [0, T]$, then the Feynman–Kac information state $q^\theta(\cdot)$ admits explicit representations, at least for the following two classes of nonlinear function $g(\cdot)$, $h(\cdot)$.

Class 1—Rational Nonlinearities: Suppose $\Gamma(x, t) > 0$, $\forall(x, t) \in \mathbb{R}^n \times [0, T]$. A solution of (38) is $W_1(x, t) = \Gamma(x, t)$, which implies $g(\cdot)$, $h(\cdot)$ should be of the form

$$\begin{aligned} g(t, x) &= \frac{G_t G_t'}{\frac{1}{2} \Delta_t x \cdot x + x \cdot \zeta_t + \eta_t} (\Delta_t x + \zeta_t) \\ h(t, x) &= \frac{\alpha_t G_t'}{\frac{1}{2} \Delta_t x \cdot x + x \cdot \zeta_t + \eta_t} (\Delta_t x + \zeta_t) \end{aligned}$$

where

$$\begin{aligned} \dot{\Delta}_t + F_t' \Delta_t + \Delta_t F_t &= \tilde{\delta}_t \Delta_t \\ \dot{\zeta}_t + F_t' \zeta_t &= \tilde{\delta}_t \zeta_t \\ \dot{\eta}_t + \frac{1}{2} \text{Tr}(G_t G_t' \Delta_t) &= \tilde{\delta}_t \eta_t \\ \dot{Q}_t &= 0 \\ \dot{\tilde{m}}_t &= 0 \\ \dot{\tilde{\delta}}_t &= \text{arbitrary}. \end{aligned}$$

Class 2—Exponential Nonlinearities: Suppose $\gamma^1, \gamma^2: [0, T] \rightarrow \mathbb{R}$. A solution of (38) is $W_2(x, t) = \gamma_t^1 \exp(\Gamma(x, t)) + \gamma_t^2 \exp(-\Gamma(x, t))$, which implies that $g(\cdot)$, $h(\cdot)$ should be of the form

$$\begin{aligned} g(t, x) &= \frac{\gamma_t^1 \exp(\Gamma(x, t)) - \gamma_t^2 \exp(-\Gamma(x, t))}{\gamma_t^1 \exp(\Gamma(x, t)) + \gamma_t^2 \exp(-\Gamma(x, t))} \\ &\quad \cdot G_t G_t' (\Delta_t x + \zeta_t) \\ h(t, x) &= \frac{\gamma_t^1 \exp(\Gamma(x, t)) - \gamma_t^2 \exp(-\Gamma(x, t))}{\gamma_t^1 \exp(\Gamma(x, t)) + \gamma_t^2 \exp(-\Gamma(x, t))} \\ &\quad \cdot \alpha_t G_t' (\Delta_t x + \zeta_t) \end{aligned}$$

where

$$\begin{aligned} \dot{\Delta}_t + F_t' \Delta_t + \Delta_t F_t &= 0 \\ \dot{\zeta}_t + F_t' \zeta_t &= 0 \\ \dot{\eta}_t + \frac{1}{2} \text{Tr}(G_t G_t' \Delta_t) &= \frac{1}{2} \frac{d}{dt} \left(\log \frac{\gamma_t^1}{\gamma_t^2} \right) \\ \dot{Q}_t &= \Delta_t G_t G_t' \Delta_t \\ \dot{\tilde{m}}_t &= \Delta_t G_t G_t' \zeta_t \\ \dot{\tilde{\delta}}_t &= \zeta_t' G_t G_t' \zeta_t + \frac{d}{dt} (\log \gamma_t^1 \gamma_t^2). \end{aligned}$$

Proof: Substitute the solutions into the evolution equation of $W(\cdot)$ or $\phi(\cdot)$. \square

Example 3.5: Let $x_t \in \mathfrak{R}$, $z_t \in \mathfrak{R}$, $y_t \in \mathfrak{R}$, and consider the risk-sensitive estimation problem

$$\begin{aligned} dx_t &= \tanh(x_t) dt + dw_t^1, & x(0) &\in \mathfrak{R} \\ dz_t &= x_t dt + z_t dt + dw_t^2, & z(0) &\in \mathfrak{R} \\ dy_t &= x_t dt + z_t dt + dx_t + db_t, & y(0) &= 0 \in \mathfrak{R} \end{aligned}$$

where \hat{x}^* is a law minimizing (4). Setting $F(t) = \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}$, $G(t) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $\Delta_t = 0$, $\zeta_t = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\eta_t = 1$, in the equations of Class 2 nonlinearities, then $W_2(x, t) = \exp(x) + \exp(-x)$ is a solution of (38), with $\tilde{Q}_t = 0$, $\tilde{m}_t = 0$, $\tilde{\delta}_t = 1$, and so Theorem 3.1 holds with $H_t = [1 \ 1]$, $\alpha_t = [1 \ 0]$.

Similarly, the risk-sensitive estimation problem

$$\begin{aligned} dx_t &= \frac{\alpha}{2} x_t dt + \frac{\Delta x_t}{\frac{1}{2}\Delta x \cdot x + \eta} dt + dw_t^1 \\ x(0) &\in \mathfrak{R}, \alpha > 0, \Delta > 0 \\ dz_t &= x_t dt + z_t dt, & z(0) &\in \mathfrak{R} \\ dy_t &= x_t dt + z_t dt + \frac{\Delta x_t}{\frac{1}{2}\Delta x \cdot x + \eta} dt + dw_t^1 + db_t \\ y(0) &= 0 \in \mathfrak{R} \end{aligned}$$

with $\tilde{\delta}_t = \alpha$, $\zeta_t = 0$, $\eta_t = (1/2\alpha)\Delta$ is finite-dimensional.

Remark 3.6: Consider once more the risk-sensitive estimation problem (17) and (18) when the cost to be minimized, namely (4), is replaced by

$$\tilde{E} \left\{ \exp \left\{ \frac{\theta}{2} \left(\int_0^t [|Q_s^{1/2}(x_s - \hat{x}_s)|^2 + \ell(s, x_s)] ds \right) \right\} \right\} \quad (39)$$

and $\ell: [0, T] \times \mathfrak{R}^n \rightarrow \mathfrak{R}$ is Borel measurable with at most quadratic growth in the space variable.

If A1) and A2) hold with $\phi(\cdot)$ a solution of the modified second-order PDE

$$\begin{aligned} \frac{\partial}{\partial t} \phi_t + \frac{1}{2} \text{Tr}(G_t G_t' D_x^2 \phi_t) + \frac{1}{2} |G_t' D_x \phi_t|^2 + F_t x \cdot D_x \phi_t \\ = \frac{1}{2} (\tilde{Q}_t x \cdot x + 2\tilde{m}_t x + \tilde{\delta}_t) + \frac{\theta}{2} \ell(t, x) \end{aligned} \quad (40)$$

then Theorem 3.1 applies. This demonstrates that different choices of $\ell(\cdot)$ yield different solutions of (40) and, consequently, different nonlinear drift and observation functions $g(\cdot)$, $h(\cdot)$ for which explicit solutions of the Feynman–Kac information state equation are available. When $\ell \geq 0$, this method enables us to derive estimators having finite L^2 -gain because (39) is always an upper bound of $J(\hat{x}(\cdot))$. For example, if $f(x)$ is a monotonically increasing nonlinearity, then the scalar risk-sensitive estimation problem $dx_t = f(x_t) dt + dw_t$, $dy_t = x_t dt + dx_t + db_t$, with $\ell(t, x) = 1/\theta |f(x)|^2 + (1/\theta) D_x f(x)$, is finite-dimensional. Here $\tilde{Q} = \tilde{m} = \tilde{\delta} = 0$.

IV. SMALL NOISE LIMITS

In this section, we present small noise limit results. It is well known that, at least for linear systems, as the noise covariances tend to zero, the risk-sensitive cost is, within a logarithmic transformation, equivalent to a deterministic linear-quadratic minimax estimation problem with additive L^2 disturbances. Similar results for nonlinear systems are derived in [9]. In this case the risk-sensitive estimator becomes a minimax estimator. We wish to derive similar connections for the risk-sensitive estimators considered earlier. Our methodology which differs from that is applicable to general nonlinear systems. We keep our treatment simple by considering the uncorrelated version of (1) and (2); the correlated case can be treated in an analogous manner.

For the purpose of this section, let us rewrite our model as follows:

$$\begin{aligned} dx_t &= (F_t x_t + g^\varepsilon(t, x_t)) dt + \sqrt{\varepsilon} G_t dw_t \\ x^\varepsilon(0) &= x(0) + \sqrt{\varepsilon} \xi^\varepsilon, & x^\varepsilon(0) &\in \mathfrak{R}^n \\ dy_t &= H_t x_t dt + \sqrt{\varepsilon} N_t^{1/2} db_t \\ y(0) &= 0 \in \mathfrak{R}^d. \end{aligned} \quad (41)$$

Here $x^\varepsilon(0)$ has mean $x(0)$.

A3)

$$\begin{aligned} g^\varepsilon(t, x) &= \varepsilon G_t G_t' D_x \phi^\varepsilon(x, t) \\ \phi^\varepsilon(x, t) &= \frac{1}{\varepsilon} \tilde{\phi}^\varepsilon(x, t) + o(1) \end{aligned} \quad (43)$$

and $\phi^\varepsilon(x, t)$ is a solution of the following PDE:

$$\begin{aligned} \frac{\partial}{\partial t} \phi^\varepsilon(x, t) + \frac{1}{2} \varepsilon \text{Tr}(G_t G_t' D_x^2 \phi^\varepsilon(x, t)) \\ + \frac{1}{2} \varepsilon |G_t' D_x \phi^\varepsilon(x, t)|^2 + F_t x \cdot D_x \phi^\varepsilon(x, t) \\ = \frac{1}{2\varepsilon} (\tilde{Q}_t x \cdot x + 2\tilde{m}_t x + \tilde{\delta}_t). \end{aligned} \quad (44)$$

A4)

$$\begin{aligned} q_0^\theta(x) &= \exp(\phi^\varepsilon(x, 0)) \\ &\times \frac{\exp(-P_0^\varepsilon \cdot^{-1}(x - \xi) \cdot (x - \xi))}{(2\pi)^n / 2 |P_0^\varepsilon|^{1/2}}; \end{aligned} \quad (45)$$

$$\lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} P_0^\varepsilon = P_0, \quad \lim_{\varepsilon \rightarrow 0} x^\varepsilon(0) = x(0). \quad (46)$$

A5) $g^\varepsilon(\cdot, \cdot)$ satisfies a Lipschitz and linear growth condition uniformly on closed and bounded sets of $[0, T] \times \mathfrak{R}^n$.

In the small noise context, the risk-sensitive cost (under measure \tilde{P}) is

$$J^{\varepsilon, \theta}(\hat{x}(\cdot)) = \tilde{E} \left\{ \exp \left\{ \frac{\theta}{2\varepsilon} \left(\int_0^t |Q_s^{1/2}(x_s - \hat{x}_s)|^2 ds \right) \right\} \right\}. \quad (47)$$

Note that θ has been replaced by θ/ε . Now we introduce two consecutive measure transformations by defining measures P and P' in such a way that under P , the measurement process becomes a Wiener process independent of the state process, while the state process remains unchanged, and under P' , the state process becomes a Wiener. Using appropriate expressions for the Radon–Nikodym derivatives $dP/d\tilde{P}$ and dP'/dP , one can easily show that the risk-sensitive estimation problem is equivalent to minimizing over $\hat{x}(\cdot)$ the Wiener functional

$$\begin{aligned} \frac{\varepsilon}{\theta} \log J^{\varepsilon, \theta}(\hat{x}(\cdot)) \\ = \frac{\varepsilon}{\theta} \log E' \left\{ \exp \left\{ \frac{\theta}{2\varepsilon} \left(\int_0^t |Q_s^{1/2}(x_s - \hat{x}_s)|^2 ds \right) \right\} \right. \\ \times \exp \left\{ \frac{1}{\varepsilon} \left(\int_0^t (H_s x_s)' N_s^{-1} dy_s \right. \right. \\ \left. \left. - \frac{1}{2} \int_0^t |N_s^{-1/2} H_s x_s|^2 ds \right) \right\} \\ \times \exp \left\{ \frac{1}{\varepsilon} \left(\int_0^t (F_s x_s + g^\varepsilon(s, x_s))' (G_s G_s')^{-1} dx_s \right. \right. \\ \left. \left. - \frac{1}{2} \int_0^t |G_s^{-1} (F_s x_s + g^\varepsilon(s, x_s))|^2 ds \right) \right\} \end{aligned}$$

where E' denotes expectation with respect to measure P' .

Note that under P' we have

$$\begin{aligned} x_t &= x^\varepsilon(0) + \sqrt{\varepsilon} \int_0^t G_s dw_s \\ y_t &= \sqrt{\varepsilon} \int_0^t N_s^{1/2} db_s \\ \text{var}(x_t) &= \text{var}(x^\varepsilon(0)) + \varepsilon \int_0^t G_s G_s' ds \\ \text{var}(y_t) &= \varepsilon \int_0^t N_s ds. \end{aligned}$$

If we now assume existence of a unique viscosity solution of the second-order PDE (44), then

$$\begin{aligned} \lim_{\varepsilon \rightarrow 0} \tilde{\phi}^\varepsilon(x, t) &= \tilde{\phi}(x, t) \\ \lim_{\varepsilon \rightarrow 0} g^\varepsilon(t, x) &= G_t G_t' D_x \tilde{\phi}(x, t) \end{aligned} \quad (48)$$

uniformly on compact sets of $\mathbb{R}^n \times [0, T]$, where $\tilde{\phi}(\cdot)$ is the unique viscosity solution of the first-order PDE

$$\begin{aligned} \frac{\partial}{\partial t} \tilde{\phi}(x, t) + \frac{1}{2} |G_t' D_x \tilde{\phi}(x, t)|^2 + F_t x \cdot D_x \tilde{\phi}(x, t) \\ = \frac{1}{2} (\tilde{Q}_t x \cdot x + 2\tilde{m}_t x + \tilde{\delta}_t). \end{aligned} \quad (49)$$

From large deviations results of functionals of Wiener processes (see [16]), one obtains

$$\begin{aligned} \lim_{\varepsilon \rightarrow 0} \frac{\varepsilon}{\theta} J^{\varepsilon, \theta}(\hat{x}^*(\cdot)) \\ = \min_{\hat{x}(\cdot)} \sup_{y \in L^2([0, T]; \mathbb{R}^d)} \sup_{\hat{x} \in L^2([0, T]; \mathbb{R}^n)} \sup_{x \in \mathbb{R}^n} \\ \cdot \left\{ \frac{1}{\theta} \tilde{\phi}(x, 0) - \frac{1}{2\theta} |P_0^{-1/2}(x - x(0))|^2 \right. \\ + \frac{1}{2} \int_0^t |Q_s^{1/2}(x_s - \hat{x}_s)|^2 ds \\ - \frac{1}{2\mu} \int_0^t [|G_s^{-1}(\dot{x}_s - F_s x_s - G_s G_s' D_x \tilde{\phi}(x_s, s))|^2 \\ \left. + |N_s^{-1/2}(\dot{y}_s - H_s x_s)|^2] ds \right\}. \end{aligned}$$

Hence, in the limit as $\varepsilon \rightarrow 0$, the risk-sensitive estimation problem is logarithmically equivalent to the minimax dynamic game

$$\begin{aligned} J^{0, \theta}(\hat{x}^*(\cdot)) &= \min_{\hat{x}(\cdot)} \sup_{b^d \in L^2([0, T]; \mathbb{R}^d)} \sup_{w^d \in L^2([0, T]; \mathbb{R}^m)} \sup_{x \in \mathbb{R}^n} \\ &\cdot \left\{ \frac{1}{\theta} \tilde{\phi}(x, 0) - \frac{1}{2\theta} |P_0^{-(1/2)}(x - x(0))|^2 \right. \\ &+ \frac{1}{2} \int_0^t |Q_s^{1/2}(x_s - \hat{x}_s)|^2 ds \\ &\left. - \frac{1}{2\mu} \int_0^t [|w_s^d|^2 + |b_s^d|^2] ds \right\} \end{aligned} \quad (50)$$

subject to dynamics and observations

$$\begin{aligned} \dot{x}_t &= F_t x_t + G_t G_t' D_x \tilde{\phi}(x, t) + G_t w_t^d, \quad x(0) \in \mathbb{R}^n \\ \dot{y}_t &= H_t x_t + N_t^{1/2} b_t^d, \quad y(0) = 0 \in \mathbb{R}^d. \end{aligned} \quad (51)$$

Here $\tilde{\phi}(\cdot, \cdot)$ is a solution of (49). One may now proceed to show that this minimax dynamic estimation problem is finite-dimensional.

Remark 4.1: Finally, we point out that when the risk-sensitive cost (47) includes in the exponent the additional term $(1/2\varepsilon) \int_0^t \ell(s, x_s) ds$, then the right side of (44) includes the term $(\theta/2\varepsilon)\ell(t, x)$. In this case, the small noise limit yields (49) and (50), with the terms $(\theta/2)\ell(t, x)$, $(1/2\theta) \int_0^t \ell(s, x_s) ds$ entering their right sides, respectively. Thus, when $\ell \geq 0$, this leads to well-defined estimators which have finite L^2 -gain.

REFERENCES

- [1] V. Benes, "Exact finite-dimensional filters for certain diffusions with nonlinear drift," *Stochastics*, vol. 5, pp. 65–92, 1981.
- [2] S. Dey and J. Moore, "Risk-sensitive filtering and smoothing for hidden Markov models," *Syst. Contr. Lett.*, vol. 25, no. 5, pp. 361–366, 1995.
- [3] D. Jacobson, "Optimal stochastic linear systems with exponential performance criteria and their relation to deterministic differential games," *IEEE Trans. Automat. Contr.*, vol. 18, pp. 124–131, Feb. 1973.
- [4] P. Whittle, "A risk-sensitive maximum principle," *Syst. Contr. Lett.*, vol. 15, pp. 183–192, 1990.
- [5] ———, "A risk-sensitive maximum principle: The case of imperfect state observations," *IEEE Trans. Automat. Contr.*, vol. 36, pp. 793–801, July 1991.
- [6] A. Bensoussan and J. H. van Schuppen, "Optimal control of partially observable stochastic systems with an exponential-of-integral performance index," *SIAM J. Contr. Optimiz.*, vol. 23, no. 4, pp. 599–613, 1985.
- [7] W. H. Fleming and W. M. McEneaney, "Risk-sensitive control and differential games," in *Stochastic Theory and Adaptive Control*, T. E. Duncan and B. Pasik-Duncan, Eds. New York: Springer-Verlag, 1992, pp. 185–197.
- [8] M. James, J. Baras, and R. Elliott, "Risk-sensitive control and dynamic games for partially observed discrete-time nonlinear systems," *IEEE Trans. Automat. Contr.*, vol. 39, pp. 780–792, Apr. 1994.
- [9] C. Charalambous, "The role of informations state and adjoint in relating nonlinear output feedback risk-sensitive control and dynamic games," *IEEE Trans. Automat. Contr.*, vol. 42, pp. 1163–1170, Aug. 1997.
- [10] C. Charalambous, D. Naidu, and K. Moore, "Solvable risk-sensitive control problems with output feedback," in *33rd IEEE Conf. Decision and Control*, Lake Buena Vista, FL, Dec. 1994, pp. 1433–1434.
- [11] C. Charalambous, "Partially observable nonlinear risk-sensitive control problems: Dynamic programming and verification theorems," *IEEE Trans. Automat. Contr.*, vol. 42, pp. 1130–1138, Aug. 1997.
- [12] A. Bensoussan and R. Elliott, "General finite dimensional risk sensitive problems and small noise limits," *IEEE Trans. Automat. Contr.*, vol. 41, pp. 210–215, Feb. 1996.
- [13] C. Charalambous and R. Elliott, "Certain nonlinear stochastic optimal control problems with explicit control laws equivalent to LEQG/LQG problems," *IEEE Trans. Automat. Contr.*, vol. 42, pp. 482–497, Apr. 1997.
- [14] P. Kumar and J. H. van Schuppen, "On the optimal control of stochastic systems with an exponential-of-integral performance index," *J. Math. Analysis and Appl.*, vol. 80, pp. 312–332, 1981.
- [15] C. Charalambous and J. Hibey, "Minimum principle for partially observable nonlinear risk-sensitive control problems using measure-valued decompositions," *Stochastics and Stochastics Rep.*, vol. 57, pp. 247–288, 1996.
- [16] D. W. Stroock, *An Introduction to Large Deviations Theory*. New York: Springer-Verlag, 1984.