A finite dimensional filter
with exponential conditional density [*]

Damiano Brigo[†]

Department of Risk Management

CARIPLO Bank

via Boito 7

20121 Milano, Italy

e-mail : dbrigo@opoipi.it

July 1997

Abstract

In this paper we consider the continuous–time nonlinear filtering problem, which has an infinite–dimensional solution in general, as proved by Chaleyat–Maurel and Michel. There are few examples of nonlinear systems for which the optimal filter is finite dimensional, in particular Kalman's, Beneš', and Daum's filters. In the present paper, we construct new classes of scalar nonlinear filtering problems admitting finite–dimensional filters. We consider a given (nonlinear) diffusion coefficient for the state equation, a given (nonlinear) observation function, and a given finite–dimensional exponential family of probability densities. We construct a drift for the state equation such that the resulting nonlinear filtering problem admits a finite–dimensional filter evolving in the prescribed exponential family augmented by the observaton function and its square.

keywords Scalar Nonlinear Diffusion Processes, Finite Dimensional Families, Exponential Families, Stochastic Differential Equations, Scalar Nonlinear Filtering Problem, Finite–Dimensional Filters.

1 Introduction

In this paper we consider the (scalar) nonlinear filtering problem in continuous time. For a quick introduction to the filtering problem see Davis and Marcus (1981) [10]. For a more complete treatment see Liptser and Shiryayev (1978) [15] from a mathematical point of view or Jazwinski (1970) [13] for a more applied perspective.

The nonlinear filtering problem has an infinite–dimensional solution in general. Constructing nonlinear systems for which the optimal filter is finite dimensional is a problem which received considerable attention in the past. It turned out that such systems are quite rare.

---

[*]Available also at www.damianobrigo.it. Part of this paper has been published in "Statistics and Probability Letters" 49 (2000), pp. 127-134.

[†]A short version co-athored by Francois Le Gland has appeared on the proceedings of the conference 36th IEEE CDC San Diego, `http://www.irisa.fr/sigma2/legland/pub/cdc97-fdf.pdf`

Examples were given by Beneš [1] and Daum [9]. Instead, general results on nonexistence of such systems, based on Lie–algebraic techniques, were made available by Chaleyat–Maurel and Michel (1984) [8], and related works appeared for example in Ocone and Pardoux (1989) [16], Lévine (1991) [14]. In the present paper, we construct scalar nonlinear filtering problems admitting finite–dimensional filters. Our examples can be added to the contributions of Beneš and Daum of known examples of finite dimensional filters. The solution of the filtering problem is a Stochastic PDE which can be seen as a generalization of the Fokker–Planck equation expressing the evolution of the density of a diffusion process. This filtering equation is called Kushner–Stratonovich equation, and an unnormalized (simpler) version of it is known as the Duncan–Mortensen–Zakai Stochastic Partial Differential Equation (DMZ–SPDE). It is this second equation that we shall consider in the present article. In [7] and [5] the Fisher metric is used to project the Kushner–Stratonovich (or the Fokker–Planck) equation onto an exponential family of probability densities, yielding the new class of approximate filters called *projection filters*. In [3, 4] the Gaussian projection filter is studied in the small-noise setting. The results given in the present paper originate from ideas expressed in such past works, especially in [5] and [6].

Our approach to the construction of nonlinear filtering problems admitting finite–dimensional filters is the following. We model the state according to a SDE whose drift may depend on the observations. We consider a given (nonlinear) diffusion coefficient for the state equation, a given (nonlinear) observation function, and a given finite–dimensional exponential family of probability densities. We construct a drift for the state equation such that the resulting nonlinear filtering problem admits a finite–dimensional filter evolving in the prescribed exponential family augmented in a particular way (depending on h).

2   Problem formulation

We start by introducing the filtering problem for scalar continuous time systems.

On the probability space $(\Omega, \mathcal{F}, P)$ with the filtration $\{\mathcal{F}_t, t \in [0\ T]\}$ we consider the following scalar state and observation equations:

$$
\begin{aligned}
dX_t &= f_t(X_t, Y_t)\, dt + \sigma_t(X_t)\, dW_t, \quad X_0, \\
dY_t &= h(X_t)\, dt + dV_t, \quad Y_0 = 0\ .
\end{aligned}
\tag{1}
$$

We set

$$
a_t(\cdot) := \sigma_t(\cdot)^2\ .
$$

Time invariance of h is needed to simplify exposition. The reason why we require it will result clear later on. These equations are Itô stochastic differential equations (SDE's). We shall use both Itô SDE's (for example for the signal X) and McShane–Fisk–Stratonovich (MFS) SDE's (when dealing with densities). The MFS form will be denoted by the presence of the symbol '∘' in between the diffusion coefficient and the Brownian motion of a SDE. The noise processes $\{W_t, t \in [0\ T]\}$ and $\{V_t, t \in [0\ T]\}$ are two standard Brownian motions. Finally, the initial state X_0 and the noise processes $\{W_t, t \in [0\ T]\}$ and $\{V_t, t \in [0\ T]\}$ are assumed to be independent.

Notice that this model is different from the models given usually (Jazwinski [13], Davis and Marcus [10]) due to the presence of Y_t in the drift f_t of the state equation. However, this does not complicate matters. Indeed, in [15] a general formulation is given of which our model is a particular case. See also, for example, the application of SPDEs theory to filtering [11], where the used model class allows dependence of the coefficients f, a and h on the observation process

Y. Such models are often encountered in stochastic control, see for example the model (8.1.5), (8.1.10) in [2]. The nonlinear filtering problem consists in finding the conditional probability distribution π_t of the state X_t given the observations up to time t, i.e. $\pi_t(dx) := P[X_t \in dx \mid \mathcal{Y}_t]$, where $\mathcal{Y}_t := \sigma(Y_s, 0 \le s \le t)$. We shall say that a filtering problem such as (1) satisfies condition (A) if:

(A) for all $t \in [0\ T]$, the probability distribution π_t has an unnormalized density q_t w.r.t. the Lebesgue measure, and $\{q_t, t \in [0\ T]\}$ satisfies the Itô–type stochastic partial differential equation (Duncan–Mortensen–Zakai SPDE)

$$dq_t = \mathcal{L}_t^* q_t \, dt + q_t \, h \, dY_t \tag{2}$$

in a suitable functional space, where for all $t \ge 0$, the forward diffusion operator $\mathcal{L}_t^*$ is defined by

$$\mathcal{L}_t^* \phi = -\frac{\partial}{\partial x}[f_t \, \phi] + \tfrac{1}{2}\frac{\partial^2}{\partial x^2}[a_t \, \phi] \ ,$$

for any test function ϕ.

Conditions under which (A) holds, and more general results on solutions of the DMZ–SPDE can be found in [11]. The corresponding MFS form of the SPDE (2) is :

$$dq_t = \mathcal{L}_t^* q_t \, dt - \tfrac{1}{2} q_t \, |h|^2 \, dt + q_t \, h \, \circ dY_t \ .$$

It is known that in general the density q_t does not evolve in a finite–dimensional parametrized family, say $\{q(\cdot, \zeta), \quad \zeta \in U \subset \mathbf{R}^{m+1}\}$. In some particular cases this happens. For example, in the linear case (f linear, $\sigma_t(x) = \sigma_t$ for all x, h linear, and X_0 with Gaussian distribution) q_t evolves in the manifold of the unnormalized Gaussian densities. Some other examples of f, σ and h ensuring that q_t evolves in a finite–dimensional family are given in [1], [9]. Notice, however, that in these examples the drift f is not allowed to depend on the observation process Y.

In the present paper we focus on exponential families, according to the following

Definition 2.1 (Unnormalized exponential family) *Let $\{c_1, \cdots, c_m\}$ be scalar functions defined on $\mathbf{R}$, such that $\{1, c_1, \cdots, c_m\}$ are* linearly independent, *have at most polynomial growth and are twice continuously differentiable. Assume that the convex set*

$$\Theta_0 := \{\theta \in \mathbf{R}^m \, : \, \psi(\theta) = \log \int \exp[\theta^T c(x)] \, dx < \infty\} \ ,$$

has non–empty interior. *Then*

$$EU(c) \quad = \quad \{q(\cdot; \theta, \beta), \, (\theta, \beta) \in U\},$$

$$q(x; \theta, \beta)) \quad := \quad \exp[\theta^T c(x) + \beta] \ ,$$

where $(\theta, \beta) := [\theta_1, \ldots, \theta_m, \beta]^T$ and $U \subseteq \Theta_0 \times \mathbf{R}$ is open, is called an unnormalized exponential family of probability densities. The $m + 1$ quantities (θ, β) are called the canonical parameters for the unnormalized exponential family $EU(c)$.

Remark 2.2 *Given linearly independent $\{c_1, \cdots, c_m\}$, it may happen that the densities $\exp[\theta^T c(x)]$ are not integrable. However, it is always possible to extend the family so as to deal with integrable densities only. Indeed, since there exist $K > 0$ and $r \ge 0$ such that*

$$|c(x)| \le K\,(1 + |x|^r) \ ,$$

for all $x \in \mathbf{R}$, we can define $d(x) := |x|^s$ for all $x \in \mathbf{R}$, and some $s > r$. Then

$$EU([c \quad d]) \quad := \quad \{q'(\cdot; \theta, \mu, \beta), \, \theta \in \mathbf{R}^m, \mu > 0, \beta \in \mathbf{R}\},$$

$$q'(x; \theta, \mu, \beta) \quad := \quad \exp[\theta^T c(x) - \mu \, d(x) + \beta],$$

is an unnormalized exponential family of densities, with a non–empty open parameter set.

In this paper we solve the following problem. Consider the nonlinear filtering problem for the system (1). Suppose that the nonlinear coefficients $\sigma_t(\cdot)$ and $h(\cdot)$ are given a priori. Let be given an exponential family. Find a drift $f = u$ such that the resulting filtering problem has a solution in the given exponential family augmented by the functions h and h^2 in the exponent. More precisely:

Problem 2.3 *Let be given any Lipschitz continuous (uniformly in time) diffusion coefficient $a_t(\cdot) = \sigma_t(\cdot)^2$ and any Lipschitz–continuous observation function h (which has at most polynomial growth). Let be given any exponential family $EU(c)$, such that $EU(c^\bullet) := EU([h \; h^2 \; c^T]^T)$ is still an exponential family according to Definition 2.1. Let (ζ, β) be the $m + 3$ canonical parameters of $EU(c^\bullet)$. Set (as comes natural) $[\theta_1, \ldots, \theta_m] = [\zeta_3, \ldots, \zeta_{m+2}]$. Let be given an initial condition X_0 such that π_0 admits an unnormalized density $q_0 = q(\cdot; \zeta_0, \beta_0)$ in the extended family $EU(c^\bullet)$. Find a drift $f = u$ such that the filtering problem for the nonlinear system (1) admits a finite dimensional solution q_t evolving in $EU([h \; h^2 \; c^T]^T)$.*

3 Solution

In this section we solve Problem 2.3 by mean of the following

Theorem 3.1 (Solution of Problem 2.3) *Assumptions of Problem 2.3 in force. If the drift*

$$u_t(x, \zeta_t) \quad := \quad \tfrac{1}{2}\frac{\partial a_t}{\partial x}(x) + \tfrac{1}{2}a_t(x)\zeta_t^T \frac{\partial c^\bullet}{\partial x}(x),$$

with

$$\zeta_t^1 := Y_t + \zeta_0^1, \quad \zeta_t^2 := \zeta_0^2 - \tfrac{1}{2}t, \quad \zeta_t^i := \zeta_0^i, \; i = 3, \ldots, m, \tag{3}$$

satisfies (A) together with a_t and h, then it solves Problem 2.3. As a consequence, for the resulting nonlinear filtering problem with coefficients u_t, a_t and h, the optimal filter is expressed by the density $q_t = q(\cdot, \zeta_t, \beta_0)$ which evolves in the finite dimensional exponential family $EU(c^\bullet)$.

Remark 3.2 (Sufficient conditions for (A)) *For explicit conditions on u_t, a_t, h and X_0 under which (A) holds (and for general results on the solution of the DMZ–SPDE) see [11]. In our case a set of sufficient condition ensuring (A) is, as from Theorem 4.1 of [11]:*

(i) h is Lipschitz continuous;

(ii) $a_t, \partial_x a_t, a_t \, \partial_x c_j^\bullet$ are Lipschitz continuous, uniformly in t, for $j = 2, \ldots, m + 2$.

(iii) $a_t(x) \, y \, \partial_x h(x)$ is Lipschitz continuos in (x, y), uniformly in t.

(iv) $Y_0 = 0$ and $q_{X_0} \in EU(c^\bullet)$.

PROOF of the theorem: Consider a candidate drift $u_t(\cdot \; ; \zeta_t, \beta_t)$ and the associated forward differential operator

$$\mathcal{U}_t^* \phi = -\frac{\partial}{\partial x} \left[u_t(\cdot \; ; \zeta_t, \beta_t) \; \phi \right] + \tfrac{1}{2} \frac{\partial^2}{\partial x^2} \left[a_t \; \phi \right] .$$

The corresponding state equation originates the filtering problem for the system

$$dX_t = u_t(X_t; \zeta_t, \beta_t)dt \; + \sigma_t(X_t) \; dW_t, \quad X_0$$
$$dY_t = h(X_t)dt + dV_t .$$

In order to check under which conditions this filtering problem has solution $q_t = q(\cdot \; ; \zeta_t, \beta_t)$ we proceed as follows. Write the right–hand side of the DMZ–SPDE for the density $q(\cdot \; ; \zeta_t, \beta_t)$ and equate it to the differential (in time) of $q(\cdot \; ; \zeta_t, \beta_t)$ computed via the chain rule:

$$\sum_{i=1}^{m+2} \frac{\partial q(\cdot \; ; \zeta_t, \beta_t)}{\partial \zeta_i} \circ d\zeta_t{}^i \; + \frac{\partial q(\cdot \; ; \zeta_t, \beta_t)}{\partial \beta} \circ d\beta_t{}^i$$
$$= \mathcal{U}_t^* \, q(\cdot \; ; \zeta_t, \beta_t) \, dt - \tfrac{1}{2} \, q(\cdot \; ; \zeta_t, \beta_t) \, |h|^2 \, dt + q(\cdot \; ; \zeta_t, \beta_t) \, h \; \circ dY_t .$$

By dividing both sides by q and straightforward calculations, one obtains

$$\sum_{i=1}^{m+2} c_i^\bullet \circ d\zeta_t^i + d\beta_t = h \circ dY_t - \tfrac{1}{2}h^2 dt - u_t(\cdot \; ; \zeta_t, \beta_t)\zeta_t^T \frac{\partial c^\bullet}{\partial x}$$
$$- \frac{\partial u_t(\cdot \; ; \zeta_t, \beta_t)}{\partial x} + \frac{1}{2 \, q(\cdot \; ; \zeta_t, \beta_t)} \frac{\partial^2 [a_t \; q(\cdot \; ; \zeta_t, \beta_t)]}{\partial x^2}.$$

A first step in finding curves $t \mapsto \zeta_t$ and $t \mapsto \beta_t$ such that this equation is satisfied is to set (remember that $c_1^\bullet = h, \; c_2^\bullet = h^2$) $d\zeta_t^1 := dY_t, \quad d\zeta_t^2 := -\tfrac{1}{2}dt$. The above equation reduces then to the (linear) differential equation

$$\frac{\partial u_t(\cdot \; ; \zeta_t, \beta_t)}{\partial x} + \zeta_t^T \frac{\partial c^\bullet}{\partial x} \; u_t(\cdot \; ; \zeta_t, \beta_t)$$

$$= -\sum_{i=3}^{m+2} \dot{\zeta}_t^i c_i^\bullet - \dot{\beta}_t + \frac{1}{2 \, q(\cdot \; ; \zeta_t, \beta_t)} \frac{\partial^2 [a_t \; q(\cdot \; ; \zeta_t, \beta_t)]}{\partial x^2}.$$

By solving this last equation in u one obtains (θ_t as defined in Problem 2.3)

$$u_t(x; \zeta_t, \beta_t) \;\; = \;\; \tfrac{1}{2}\frac{\partial a_t(x)}{\partial x} + \tfrac{1}{2}a_t(x)\zeta_t^T \frac{\partial c^\bullet}{\partial x}$$

$$- \;\; \dot{\theta}_t^T \exp[-\zeta_t^T c^\bullet(x)] \; \int_{-\infty}^x c(z) \exp[\zeta_t^T c^\bullet(z)]dz$$

$$- \dot{\beta}_t \int_{-\infty}^x \exp[\zeta_t^T c^\bullet(z)]dz .$$

We are free to choose the curve $t \mapsto (\theta_t, \beta_t)$ as we wish, as long as it is regular. Set $\theta_t := \theta_0$ for all $t \in [0 \; T]$, and $\beta_t := \beta_0$ for all $t \in [0 \; T]$. $\qquad\square$

4 Examples

We present the following examples of applications of Theorem 3.1.

Example 4.1 Cubic observations. *The present example is inspired by the cubic sensor studied in [12], where it is proven that for the cubic sensor problem there exists no finite dimensional filter. Here we consider the case where the diffusion coefficient for the state equation is constant (say $a_t(x) = 1$ for all $t \in [0, T]$ and $x \in \mathbf{R}$) and with cubic observation function ($h(x) = x^3$ for all $x \in \mathbf{R}$). A straightforward application of Theorem 3.1 yields the following result.*

The filtering problem for the system

$$dX_t = [\frac{3}{2}Y_t X_t^2 - 3(1 + \frac{1}{2}t)X_t^5]dt + dW_t, \quad q_{X_0} \sim \exp[-x^6],$$

$$dY_t = X_t^3 dt + dV_t, \quad Y_0 = 0,$$

is finite dimensional and has conditional law with density

$$p_{X_t|\mathcal{Y}_t}(x) \propto \exp[Y_t x^3 - (1 + \frac{1}{2}t)x^6], \quad t \in [0, T], \quad x \in \mathbf{R}.$$

Example 4.2 Linear Observations. *We consider the case of linear observations. For simplicity, take h equal to the identity function. Theorem 3.1 yields the following result: The optimal filter for the filtering problem*

$$dX_t = \{\frac{1}{2}\frac{\partial a_t}{\partial x}(X_t) + \frac{1}{2}a_t(X_t)[Y_t + \frac{\mu_0}{v_0} - (t + \frac{1}{v_0})X_t]\}dt$$
$$+ \sigma_t(X_t)dW_t, \quad X_0 \sim \mathcal{N}(\mu_0, v_0), \quad a_t = \sigma_t^2,$$

$$dY_t = X_t dt + dV_t, \quad Y_0 = 0,$$

is finite dimensional and has conditional law with density

$$X_t|\mathcal{Y}_t \sim \mathcal{N}(\frac{\mu_0 + Y_t v_0}{1 + v_0 t}, \frac{v_0}{1 + v_0 t}).$$

5 Conclusion

It seems, at a first sight, that our result contradicts classical results on nonexistence of finite dimensional filters, such as for example Chaleyat–Maurel and Michel (1984) [8], and the related works Ocone and Pardoux (1989) [16], Lévine (1991) [14]. This contradiction appears a natural consequence of the arbitrariness of σ and h. Nonetheless, there is no real contradiction. Indeed, since $\{\zeta_t, \quad t \in [0 \ T] \}$ depends on the observation process Y, the drift itself depends on the observations. This assumption is not allowed in the works mentioned before, and indeed we cannot construct a nonlinear filtering problem with prescribed (nonlinear) σ and h, with drift *u which does not depend on the observation process Y* and whose solution remains finite dimensional. We have to allow for observations-dependent drifts in order to prove our result.

6 Acknowledgements

References

[1] V.E. Beneš, Exact Finite–Dimensional Filters for Certain Diffusions with Nonlinear Drift, *Stochastics* **5** (1981) 65–92

[2] A. Bensoussan, *Stochastic Control of Partially Observable Systems*, Cambridge University Press, Cambridge, 1992.

[3] D. Brigo, On the nice behaviour of the Gaussian Projection Filter with small observation noise, *Systems and Control Letters* **26** (1995) 363–370

[4] D. Brigo, New results on the Gaussian projection filter with small observation noise, to appear in *Systems and Control Letters*.

[5] D. Brigo, *Filtering by Projection on the Manifold of Exponential Densities*, PhD Thesis, Free University of Amsterdam, 1996.

[6] D. Brigo, On diffusions with prescribed diffusion coefficients whose densities evolve in prescribed exponential families, *Internal Report CNR–LADSEB* **02/96**, CNR–LADSEB, Italy, March 1996.

[7] D. Brigo, B. Hanzon, F. Le Gland, A differential geometric approach to nonlinear filtering: the projection filter, In *Proceedings of the Conference on Decision and Control*, IEEE-CSS, pp 4006–4011, New Orleans, 1995 (extended version available on the internet at URL: ftp://ftp.irisa.fr/techreports/1995/PI-914.ps.Z).

[8] M. Chaleyat-Maurel, D. Michel, Des resultats de non-existence de filtre de dimension finie, *Stochastics*, **13** (1984) 83–102.

[9] F. E. Daum, Exact Finite–Dimensional Nonlinear Filters, *IEEE Transactions on Automatic Control*, **31** (1986) 616–622.

[10] M. H. A. Davis, S. I. Marcus, An introduction to nonlinear filtering, in: M. Hazewinkel, J. C. Willems, Eds., *Stochastic Systems: The Mathematics of Filtering and Identification and Applications* (Reidel, Dordrecht, 1981) 53–75.

[11] I. Gyöngy, N.V. Krylov, Stochastic partial differential equations with unbounded coefficients and applications II, *Stochastics*, **32** (1990) 165–180.

[12] M. Hazewinkel, S.I.Marcus, and H.J. Sussmann, Nonexistence of finite dimensional filters for conditional statistics of the cubic sensor problem, *Systems and Control Letters* **3** (1983) 331–340.

[13] A. H. Jazwinski, *Stochastic Processes and Filtering Theory*, Academic Press, New York, 1970.

[14] J. Lévine, Finite dimensional realizations of stochastic PDE's and application to filtering, *Stochastics and Stochastic Reports* **43** (1991) 75–103.

[15] R.S. Liptser, A.N. Shiryayev, *Statistics of Random Processes I, General Theory* (Springer Verlag, Berlin, 1978).

[16] D. Ocone, E. Pardoux, *A Lie algebraic criterion for non-existence of finite dimensionally computable filters*, Lecture notes in mathematics 1390, 197–204 (Springer Verlag, 1989)