Old Dominion University

ODU Digital Commons

Electrical & Computer Engineering Theses & Dissertations

Electrical & Computer Engineering

Winter 1992

Topics in Nonlinear Stochastic Control, Estimation, and Decision, Using a Measure Transformation Approach

Charalambos Demetriou Charalambous
Old Dominion University

Follow this and additional works at: https://digitalcommons.odu.edu/ece_etds

Part of the Electrical and Computer Engineering Commons

Recommended Citation

Charalambous, Charalambos D.. "Topics in Nonlinear Stochastic Control, Estimation, and Decision, Using a Measure Transformation Approach" (1992). Doctor of Philosophy (PhD), Dissertation, Electrical & Computer Engineering, Old Dominion University, DOI: 10.25777/49xw-2678
https://digitalcommons.odu.edu/ece_etds/176

This Dissertation is brought to you for free and open access by the Electrical & Computer Engineering at ODU Digital Commons. It has been accepted for inclusion in Electrical & Computer Engineering Theses & Dissertations by an authorized administrator of ODU Digital Commons. For more information, please contact digitalcommons@odu.edu.

TOPICS IN NONLINEAR STOCHASTIC CONTROL, ESTIMATION, AND DECISION, USING A MEASURE TRANSFORMATION APPROACH

by

Charalambos Demetriou Charalambous

A Dissertation Submitted to the Faculty of
Old Dominion University in Partial Fulfillment of the
Requirements for the Degree of

DOCTOR OF PHILOSOPHY

ELECTRICAL ENGINEERING

OLD DOMINION UNIVERSITY
DECEMBER 1992

Approved by:

Joseph L. Hibey (Director)

Oscar R. Gonzalez

Griffith J. McRee

Carl A. Schulz

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

ABSTRACT

Topics in Nonlinear Stochastic Control, Estimation, and Decision, Using a Measure Transformation Approach

C. D. Charalambous

Old Dominion University, 1992

Director: Dr. Joseph L. Hibey

We discuss topics in the theory of nonlinear stochastic control, estimation, and decision via a probabilistic approach using measure transformations and martingale theory. First, we investigate the problem of estimating a diffusion process using coordinate transformations and measure transformations, both locally and globally; this is the analog of nonlinear coordinate and state feedback transformations used to obtain exact linearization in nonlinear deterministic control problems. Our results are new in that we use a probabilistic approach rather than a purely geometric one, and also in that we derive representations when the processes are defined locally rather than just globally. A gauge transformation then leads to a Feynman-Kac formula that is related to the unnormalized conditional density and subsequent bounds of filter estimates, where some of these bounds are extensions of pre-existing results while others are presented here for the first time. Second, we present new methods and new results in obtaining a minimum principle for partially observed diffusions using calculus of variations when the control variable is

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

present only in the drift coefficient and correlation exists between state and observation noise, and then when the control variable exists in both drift and diffusion coefficients and no correlation exists. Here the problem is formulated as one of complete information, but instead of considering the unnormalized conditional density as the new state, this density is decomposed into two measure-valued processes and leads to a separation principle reminiscent of the linear-quadratic-Gaussian problem and stochastic flows of Euclidean processes. Third, we study the decision problem using likelihood-ratio tests and evaluate the performance using Chernoff bounds. We present new results by expressing both likelihood-ratios and error-probabilities in terms of a ratio of two unnormalized conditional densities where each satisfies a stochastic differential equation that in some cases can be solved in closed form.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Dedicated to my family
Neophyta, Demetris, Antri and Athos

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

ACKNOWLEDGMENTS

First and foremost, I wish to convey my sincere gratitude to my advisor, Professor Joseph L. Hibey, who aroused my interest in stochastic system theory and provided the motivation, encouragement and continuous involvement throughout this research. It has been a privilege to work with Professor Hibey, as his teaching of the most important elements in my doctoral research has been essential for this work and to my future education.

I also want to thank Professors Oscar Gonzalez, Griffith McRee, and Carl Schulz, for being on my thesis committee and for reading and checking this work.

Special thanks are due to Professor Naidu D. S. for his guidance during part of my master's and one and a half years of doctoral work at NASA Langley Research Center on aeroassisted orbital transfer vehicles. It was the study of this stochastic control problem (reported in [107]) that motivated Dr. Hibey and me to pursue the subject matter reported in this dissertation.

I wish to thank NASA Langley Research Center for financial support provided under Grant number NAG-1-736.

I am indebted to my employers, George Pitsilides, owner of Captain George's Restaurant, and Kathy Pitsilides, Vice-President, for their support on this long term goal and for providing a flexible and enjoyable working environment.

I thank Alicia C. Walker for her excellent typing of this thesis.

Finally, thanks are due to my parents, Neophyta and Demetris Onisiphorou, for their love, support and encouragement over the long academic years.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

TABLE OF CONTENTS

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

LIST OF TABLES

TABLE

PAGE

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

CHAPTER 1

INTRODUCTION

We discuss topics in the theory of nonlinear estimation, nonlinear decision and nonlinear stochastic control with partial observations via a probabilistic approach in terms of measure transformations and martingale theory. First, we investigate the problem of estimating a diffusion process observed in white noise by using diffeomorphisms, measure and gauge transformations. In the case that the nonlinear stochastic system is modelled in terms of state-variables, an exact linearization of the diffusion process is obtained through stochastic differential rules and measure transformations; this is the analog of coordinate and nonlinear state-feedback transformations established for the exact linearization of nonlinear deterministic control systems as discussed by Isidori [79] (for the local case) and Dayawansa, Boothby and Elliott [41] (for the global case). Our results are new in that we use a probabilistic approach rather than a purely geometric one and also in that we derive representations when the processes are defined locally rather than just globally. A further measure and gauge transformation results in a Feynman-Kac formula related to the evolution of the unnormalized conditional density and subsequent bounds of filter estimates, where some of these bounds are extensions of pre-existing results while others are presented here for the first time. The main results are:

(a) A set of necessary and sufficient conditions for linearization of nonlinear diffusion processes both locally and globally;

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

(b) A set of finite-dimensional sufficient statistics for obtaining the unnormalized conditional density for the global case;

(c) By duality, a nonlinear degenerate stochastic control problem with explicit generalized solution;

(d) An initial-boundary value problem that describes the problem of locally estimating a diffusion process with termination;

(e) Approximate methods for obtaining lower and upper bounds on filter estimates using conditional correlation coefficients;

(f) Extension of the Bobrovsky and Zakai [23, 24] lower bound on the mean-square estimation error to degenerate diffusion processes.

We have learned that independently Cohen and Levine [34] have proved a theorem equivalent to our results (a) and obtained a finite-dimensional filter based on the linearization techniques. Besides their approach being completely different than ours in obtaining the finite-dimensional filter, they do not provide any of our results (c)-(f).

Second, we present a new method and derive new results in obtaining a minimum principle for partially observed diffusions when the control variable is present only in the drift coefficient and correlation between observation and state noise is present, and then when control variable is present both in the drift and diffusion coefficients and no correlation is allowed. The problem is formulated as one of complete information, but instead of considering the unnormalized density as the new state of the system, this density is decomposed into two measure-valued processes. This decomposition was first indroduced by Kunita [94, 96] to prove existence and uniqueness of solutions to stochastic partial differential equations. Here, this decomposition is used for the first time in the context of a stochastic control problem to serve as a separation principle similar to the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Wonham [130] separation principle of linear state-valued processes. The main results obtained using weak variations in L^2 space are:

(a) A rigorous derivation of Pontryagin's minimum principle in L^2 space and a formal derivation for the explicit representation of the adjoint-process and

(b) A rigorous derivation of Pontryagin's minimum principle in L^2 space and an explicit representation of the adjoint-process using the martingale representation theorem.

Third, we study the nonlinear decision problem when the unobserved process satisfies a diffusion process by observing only nonlinear functions of the unobserved process corrupted by white noise, so as to determine stochastic partial differential equations for computing the decision strategy and exact performance bounds. The decision strategy employed is the likelihood-ratio test and the performance bounds are due to Chernoff. Using the martingale representation theorem and measure transformations we prove that the likelihood-ratio test and performance bounds can be represented as a ratio of two unnormalized conditional densities integrated over the whole space, where each density satisfies a stochastic partial differential equation;to the best of our knowledge, these representations are new. The main results obtained through this approach are:

(a) A complete characterization of decision strategy and performance bounds by solving a single stochastic partial differential equation;

(b) A closed form solution for the evaluation of the decision strategy and performance bounds for linear and certain nonlinear decision problems.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

We now give a more detailed description of the methods employed in this study and show how they fit in with the historical development provided by other researchers.

1.1 THE NONLINEAR FILTERING PROBLEM

The nonlinear filtering problem involves the estimation of a stochastic process $\{x_t,\ t \geq 0\}$, called the state process, which cannot be observed directly. Information about x is available only through observing a related process $\{y_t,\ t \geq 0\}$, called the observation process. Given a complete probability space $(\Omega, \mathcal{F}, \mathcal{P})$ on which x, y are defined such that the y process satisfies

$$dy_t = h(t,x_t)dt + db_t \quad , \ y_{t_0} = 0 \tag{1.1.1}$$

where b_t is a noise process (usually an independent increment process), the goal is to compute, for each t, the least-squares estimate of functions $\varphi(\cdot)$ of x_t given the observation history $\{y_s,\ 0 \leq s \leq t\}$. Thus, one computes either the conditional expectation $\Pi_t(\varphi) \triangleq E\{\varphi(x_t)|\mathcal{F}_t^y\}$, where $\mathcal{F}_t^y \triangleq \sigma\{y_s, 0 \leq s \leq t\}$, is the σ-algebra generated by y, or the entire conditional distribution of x. Furthermore, this computation should be done recursively in terms of a statistic $\{\Theta_t,\ t \geq 0\}$ that evolves on a finite-dimensional manifold and which can be updated using new observations as

$$\Theta_{t+s} = \alpha(t,s,\Theta_t,\{y_u,t \leq u \leq t+s\}) \tag{1.1.2}$$

and from which estimates can be calculated in a "pointwise" fashion given by

$$\Pi_t(\varphi) = \beta(t,\varphi,y_t,\Theta_t). \tag{1.1.3}$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Generally, Θ_t is related to the conditional distribution of $\{x_t, t \geq 0\}$ given $\mathscr{F}_t^y$ and depends nonlinearly on the observations $\{y_s, 0 \leq s \leq t\}$; equation (1.1.2) is known as the nonlinear filter.

The most recent results of filtering theory which give a characterization of the optimal filters are formulated in two general situations. The first is a version of Bayes formula for $\Pi_t(\varphi)$ as presented by Kallianpur-Striebel [85], where $\Pi_t(\varphi)$ is represented by a functional space integration over the path space of $\{x_t, t \geq 0\}$ with $x_{t_0} = x_0$, $x_t = z$. This function space integration is interpreted as integration against conditional Wiener measure as explained by Mitter [106]. The function space integration is the most general since it is valid for minimum restrictions on x, h, φ as shown in a paper by Beneš and Karatzas [6], where the distribution of the initial state x_{t_0} is not necessarily Gaussian, but has finite first and second moments. The second characterization of optimal filters is valid when the $\{x_t, t \geq 0\}$ process is Markovian. As presented by Liptser and Shiryayev [103] and Fujisaki, Kallianpur and Kunita [63], $\Pi_t(\varphi)$ satisfies a nonlinear measure-valued stochastic differential equation [103, Theorem 8.3] whose solution is the conditional distribution. However, in general, $\Pi_t(\varphi)$ cannot be evaluated from this result because the filter equation depends on higher order moments which in turn require estimates of yet other functions, thus resulting in a system of equations which is infinite-dimensional.

The above two characterizations of nonlinear filters are a general extension of the celebrated Kalman-Bucy [83] filter which provides the solution of the model (1.1.1) when

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$\{x_t, t \geq 0\}$ is a Gaussian diffusion process, $h(t,x_t)$ is linear in x, $\varphi(x_t) = x_t$ and w_t is a Brownian motion. For this specific case the conditional distribution of the present state, given past and present observations, is shown to be Gaussian with nonrandom covariance, and the conditional mean vector satisfies a linear stochastic differential equation driven by the observations (or innovations), the "Kalman filter".

Both characterizations of the nonlinear filtering problem can be formulated in terms of an evolution-type linear stochastic partial differential equation (PDE) whose solution is the unnormalized version of the conditional density of the process $\{x_t, 0 \leq t \leq T\}$ given the observations $\mathcal{F}_t^y$, $t \in [0,T]$. This stochastic PDE will be referred to as the Duncan-Mortensen-Zakai (DMZ) equation and provides the most complete interpretation of the nonlinear filtering problem.

Despite extensive research in nonlinear filtering, generally speaking, the exact computation of the conditional distribution remains unsolved due to its mathematical intractability; thus, one is usually forced to seek suitable sub-optimal filters that can be solved using approximation techniques. Since the filtering problem posed in this thesis is investigated using as a primary tool the DMZ equation, we shall briefly present some of the recent advances in this field that use this approach.

1.1.1 Previous Method of Solution

The first explicit result for the nonlinear filtering problem was presented by Beneš [4] who considered a diffusion process satisfying

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$dx_t = f(t,x_t)dt + \sigma(t,x_t)dw_t, \quad x_{t_0} = x_0 \tag{1.1.4}$$

when observed through the noisy measurements (1.1.1). Using the Kallianpur-Striebel [85] formulation, Beneš [4] was able to solve the DMZ equation under the assumptions that (i) $f \in R^1$, $h \in R^1$, $f(x) = \nabla F(x)$, $h(x) = x$, (ii) $\sigma \in R^1$ is constant, and (iii) $\frac{\partial f}{\partial x} + f^2 + h^2 = c_1 x^2 + c_2 x + c$. With $I_{\{\cdot\}}$ denoting the indicator function of the event $\{\cdot\}$, his formulation requires the evaluation of

$$\tilde{p}(x, t) \triangleq \tilde{E}[I_{\{x_0+w_t \in dx\}} \exp \{\int_0^t h(x_s)dy_s - \frac{1}{2}\int_0^t |h(x_s)|^2 ds\}$$
$$\times \exp \{\int_0^t f(x_0+w_s)dw_s - \frac{1}{2}\int_0^t |f(x_0+w_s)|^2 ds\} | \mathcal{F}_t^y] \tag{1.1.5}$$

which is a consequence of the Radon-Nikodym Theorem (otherwise called a measure transformation). Whenever conditions (i) - (iii) are satisfied this expectation would produce the sufficient statistics computed in a recursive manner much as in the case of Kalman filter. As we shall see later, equation (1.1.5) is also the fundamental solution to the DMZ equation

$$\begin{aligned} d\rho(x,t) &= A(t)^* \rho(x,t)dt + h(x)\rho(x,t)dy_t \\ \lim_{t \downarrow t_0} \rho(x,t) &= p_{t_0}(x) \end{aligned} \tag{1.1.6}$$

where $A(t)^*$ is the adjoint of the second-order operator

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$A(t)=\frac{1}{2}\sigma(x)\sigma(x)^T\frac{\partial^2}{\partial x^2}+f^T(x)\frac{\partial}{\partial x}.$$

Introducing a gauge transformation the author has also shown that the solution $\rho(x,t)$, called the unnormalized conditional density, can be factored into two exponential terms, the first being a time-independent term which results from the gauge transformation and the second being a time-dependent term which is Gaussian. This approach is due to an earlier paper by Beneš and Shepp [2], where the solution to the density function of unconditioned diffusion satisfying (1.1.6) with $h(x) = 0$ is obtained under assumptions (i) - (iii). The multidimensional version of the scalar case above is also shown by Beneš [4] to produce finite-dimensional statistics.

In the same paper, Beneš also studied the interplay between finite-dimensional filters and the Lie algebra (LA) (estimation algebra) generated by

$$LA \triangleq \{A^* - \frac{1}{2}h^2, h\}. \tag{1.1.7}$$

The importance of LA was originally investigated by Brockett and Clark [28], and Brockett [26, 27, 29] as follows. Suppose (1.1.2), (1.1.6) are recast as Fisk-Stratonovich equations. Then if (1.1.2) evolves on a finite-dimensional manifold the conditional statistic $\Pi_t(\varphi)$ given by (1.1.3) is finite-dimensionally computable (i.e., computed pointwise). The system (1.1.2), (1.1.3) is a representation of the mapping from input functions y_t, $t \in [0,T]$ to output functions $\Pi_t(\varphi)$. Notice that (1.1.6) is a bilinear

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

differential equation in $\rho(x,t)$, and $\rho(x,t)$ is related to the normalize conditional density $p(x_t | \mathcal{F}_t^y)$

of x_t given $\mathcal{F}_t^y$ by

$$p(x_t | \mathcal{F}_t^y) = \frac{\rho(x,t)}{\int_{R^n} \rho(x,t)dx}. \tag{1.1.8}$$

Suppose now that some statistic $\Pi_t(\varphi)$ of the conditional distribution of $\{x_t,\ t \in [0,T]\}$ given the information $\mathcal{F}_t^y$, $t \in [0,T]$, can be calculated with a finite-dimensional recursive estimator of the form (1.1.2), (1.1.3); then this statistic can also be obtained from $\rho(x,t)$ through

$$\Pi_t(\varphi) = \frac{\int_{R^n} \varphi(x)\rho(x,t)dx}{\int_{R^n} \rho(x,t)dx}. \tag{1.1.9}$$

But the Fisk-Statonovich equation

$$d\rho(x,t) = (A^*(t) - \frac{1}{2}h^2(x))\rho(x,t)dt + h(x)\rho(x,t) \circ dy_t \tag{1.1.10}$$

together with (1.1.9) is a system representation of the same input/output mapping as (1.1.2), (1.1.3). Recalling that the Θ_t process involved in (1.1.2), (1.1.3) is required to be finite-dimensional, the analog $\rho(x,t)$ in (1.1.9), (1.1.10) is now seen to be infinite-dimensional, even though both have the same input/output map. Therefore, to gain insight into the solution of (1.1.10), we can view the y process in (1.1.10) as a control input and

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

appeal to the results covering solutions of finite-dimensional differential equations as described by Sussmann [125] and Hermann and Krener [73]. The importance of LA in (1.1.7) is now evident, the factors shown in its definition appear explicitly on the right side of (1.1.10). From the point of view of control theory, this algebra determines the region of accessibility of control functions u. As a consequence, LA governs the region accessible to the conditional density, and thus, this is an indication that some conditional statistic may be computed by an estimator of the form (1.1.2), (1.1.3). Indeed, this was the original modivation for studying the estimation algebra LA which reflects the complexity of the density equation $\rho(x,t)$ and the dimension of the manifold the density $\rho(x,t)$ lives on, which is also related to the number of sufficient statistics. Using this approach, Brockett and Clark [28] study the estimation of a finite state Markov process observed in additive Brownian motion leading to the discovery of new filters for the conditional distribution. For the linear case, the Lie algebraic approach provides results on how the evolution of differential equations for the sufficient statistics summarizes the relevance of past observations. The basic insight of the Lie algebraic approach is that the estimation algebra of (1.1.2) should be a homomorphic image of the estimation algebra generated by $\{A^* - \frac{1}{2}h^2, h\}$.

Ocone [109] studies the question of existence and representation of finite-dimensional filters using results established by Brockett and Clark [28], Brockett [26, 27, 29], and Mitter [106], who demonstrated the importance of Lie algebras to nonlinear filtering problems. Ocone establishes the connection between finite-dimensional

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

estimation algebras and finite-dimensional filters using the Wei and Norman [129] method. In fact, using the estimation algebra, he shows how one can integrate it, in the case when it is finite-dimensional, to recover the solution of the unnormalized conditional density. However, no extensions of Beneš results are obtained.

For the reader who is interested in applications of the Lie algebraic method to nonlinear filtering and smoothing, we provide the following bibliography: Roth and Loparo [117], Marcus [105], Hazewinkel and Marcus [70, 71], Krishnaprasad, Marcus and Hazewinkel [87, 88] Hazewinkel, Hazewinkel [69], Blankenship, Chang and Marcus [21], Benes [5], Ocone, Baras and Marcus [110].

In a series of papers by Pardoux [112, 113, 114, 115], the problem of existence, uniqueness and representation of the solution to the DMZ equation is investigated. His approach is mainly probabilistic. In the above cited references Pardoux uses probabilistic metho2ds to derive the equation that is adjoint to the DMZ equation and its robust version. The adjoint satisfies a backward stochastic PDE and its robust version satisfies a backward PDE equations. The backward and forward SDE's are interpreted in the nonlinear filtering set up as the backward and forward Kolmogorov equations for unconditioned diffusions. Furthermore Pardoux [112, 114] considers the problem of existence and uniqueness of the solution to the above stochastic PDE's by using variational method. The problem of prediction and smoothing is also investigated by the same author. Bensoussan [12], analyzes the evolutional properties of the above differential equations using the methodology presented by Pardoux [112, 114]. Similar results on existence and uniqueness are obtained by Kunita [93, 94, 96], Krylov and

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Rozovskii [90], Curtain [35], Sheu [120], Baras, Blankenship and Hopkins [1], whereas Davis [36, 37, 38, 39, 40] and Clark [33] emphasize pathwise solutions of the above partial differential equations using the result of Doss [44] and Sussmann [125].

Approximate methods to the nonlinear filtering problem are established in two directions. The first is based on approximating the DMZ equation (1.1.6) and the second is based on approximating the stochastic differential system (1.1.1), (1.1.4) using weak convergence techniques. The first approach is investigated by Elliott and Glowinski [48], Florhinger and Le Gland [59], Le Gland [102], Bensoussan, Glowinski and Ruscanu [13], Bensoussan [13]. The second approach is investigated by Kushner [99, 100].

1.1.2 Proposed Method of Solution

Our approach being of a probabilistic nature, most nearly resembles the one considered by Beneš [4], and Pardoux [113, 114, 115]. Thus, we also use the Radon-Nikodym Theorem as a basic tool to reduce the observation process to a pure Brownian motion, but unlike the authors above, we consider the theory of state-feedback linearization of control systems adapted to stochastic systems to introduce a new equivalent filtering problem. From the point of view of algebraic considerations it was originally shown by Brockett [27] that global coordinate transformations result in isomorphisms of estimation algebras. It is felt that the linearization approach allows us to provide a clear indication of the complexity that exists in determining finite-dimensional computable sufficient statistics. Whenever local coordinate transformation is under consideration we prove that the equation satisfied by the unnormalized conditional density is a stochastic PDE with split boundary conditions by assuming that

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

the diffusion process (1.1.4) terminates at the first exit from a bounded domain of interest. We further introduce a stochastic control problem having a generalized solution related to the filtering problem and determine bounds on certain statistical information associated with the nonlinear filtering problem. We now present a brief outline of the solution procedure, with details found in Chapter 3.

Step 1. We start with a stochastic differential system defined on a manifold M. We then perform local and global coordinate transformation through the use of stochastic differential rules, and measure transformations to represent the original system in a new coordinate system under a new probability law having a measure which is equivalent to the original measure. This transformation of measure allows us to redefine a new driving stochastic input that leads to a linear controllable system which is equivalent to the original one. The coordinate and measure transformations are viewed as the analog of state-feedback linearization of deterministic control systems investigated by Isidori [79] and Nijmeijer and Shaft [108] for the local case and Dayawansa, Boothby and Elliott [41] for the global case.

Step 2. Next, we derive necessary and sufficient conditions for local and global linearization of nonlinear stochastic differential systems having an equivalent controllable representation in the geometric content of the original system. Global results are only given for single-input/single-output systems whereas local results are applicable to multi-input/multi-output systems, where input denotes the martingale term of the state process and output denotes the observation process.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Step 3. We then investigate the nonlinear filtering problem that satisfies the necessary and sufficient conditions of Step 2, both globally and locally. That is, we recognize that the original and linearized filtering problems are equivalent in the sense that the conditional density of one can be obtained from the conditional density of the other. We proceed by introducing another measure transformation also used by the previous authors to reduce the observation equation (1.1.1) to $dy_t = db_t$ as in Zakai [134] and Wong [131]. Finally, the solution of a version of DMZ is shown to have a finite-dimensional solution which is interpreted as the degenerate version of Beneš' [4] finite-dimensional filtering example whenever global linearization is considered. For the case of local linearization, the problem becomes more complex since the evolution of the unnormalized conditional density satisfies a partial differential equation with split boundary conditions described by Friedman [60] as an initial-boundary value problem. Using the approach of Fleming [57] and Fleming and Mitter [58] a stochastic control problem is introduced. When the filtering problem is defined globally, the associated stochastic control problem is shown to have a generalized solution which is related to the solution of the filtering problem by a gauge transformation.

Step 4. Since we were unable to completely characterize the filtering problem by finite-dimensional statistics, we are forced to consider bounding techniques. Using a property of the correlation coefficient we succeed in obtaining bounds on certain statistics of the nonlinear filtering problem. Furthermore, we extend the result of Bobrovsky and Zakai [23, 24], who derived lower bounds on the minimum-mean-square-error, to the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

filtering problem of Step 2. Thus, filtering problems that are linearizable through coordinate and measure transformations admit such a lower bound estimate.

1.2 THE PARTIALLY OBSERVED STOCHASTIC CONTROL PROBLEM

For the system (1.1.1), (1.1.4) presented in Section 1.1 with correlation between measurement and process noise and dependence of f on the control functions $\{u_t, t \in [0,T]\}$, and then when control functions $\{u_t, t \in [0,T]\}$ are present on f, σ with the above correlation set to zero, an optimal control problem is formulated by specifying a performance criterion J(u) of the form

$$J(u) = E\{\int_0^T \amalg(t,x_t,u_t)dt + \kappa(x_T)\}. \tag{1.2.1}$$

The problem can be formulated as one of constraint optimization where u is chosen to minimize (1.2.1) subject to constraints (1.1.1), (1.1.4). However, even though deterministic constraint optimization is well understood, the presence of random disturbances makes the above problem very difficult to analyze.2

In order to establish a minimum principle for the above problem, we are naturally led into a consideration of a measure transformation commonly used in nonlinear filtering. As pointed out in Section 1.1.1 this measure transformation results in a stochastic PDE which in this case is pathwise dependent on the control variable u. The minimization problem is then cast as one of complete observations given by

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$J(u) = \tilde{E}\{\int_0^T \langle \Pi(t, x_t, u_t), \rho(x_t, t)\rangle + \langle k(x_T), \rho(x, T)\rangle\} \tag{1.2.2}$$

subject to the constraints

$$\begin{aligned} d\rho(x,t) &= L^u(t)^* \rho(x,t)dt + \sum_{k=1}^{d} M_k^u(t)^* \rho(x,t)dy_t^k \\ \lim_{t \downarrow t_0} \rho(x,t) &= p_{t_0}(x) \end{aligned} \tag{1.2.3}$$

where

$$\langle \alpha(x), \beta(x)\rangle \triangleq \int_{R^n} \alpha(x)\, \beta(x)dx$$

$$M_k^u(t) \triangleq h_k(t,x_t) + Y_k^u(t),$$

$$Y_k^u(t) \triangleq \sum_{i,j}^{n,m} \gamma^{kj} \sigma_j^i(t,x_t,u_t)\frac{\partial}{\partial x^i}, \quad k = 1, \dots, d.$$

The correlation between the state process and observation process is denoted by

$$\langle y_\cdot^k, w_\cdot^j\rangle_t = \int_0^t \gamma^{kj} ds, \quad j = 1, \dots, m, \quad k = 1, \dots, d.$$

Notice that if the diffusion coefficient σ of (1.1.4) is independent of the control variable u, the operator $M_k^u \to M_k$ and is independent of the control variable u. Moreover, if the correlation γ^{jk} is set to zero (i.e., w_t, b_t uncorrelated) then $M_k^u \to h_k(t,x)$ and is again independent of control variable u while (1.2.3) is now given by (1.1.6). The above problem governed by (1.2.2), (1.2.3) is converted into a problem of complete information

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

by considering the measure-valued function $\rho(x,t)$ as the new state of the system. This formulation was originated by Striebel [123] in a discrete-time set up and is sometimes referred to as the separated stochastic control problem of partially observed systems. We also note that the set of admissible controls considered in this thesis are of strict sense, that is, functions of the form

$$u(t) = u(t, \{y_s,\ 0 \le s \le t\}),$$

and not those of wide sense that is, functions of the form

$$u(t) = u(t, \{y_s,\ 0 \le s \le t\},\ \{v_s,\ 0 \le s \le t\}),\ v_t = \int_0^t u_s ds.$$

1.2.1 Previous Methods of Solution.

In recent papers, Bensoussan [10, 15], Haussmann [68] and Elliott and Yang [50] addressed the stochastic control problem above when the operator $M_k(t)$ is a zero-order differential operator, which corresponds to the case when state and observation noises are independent or simply uncorrelated.

Bensoussan [10], using weak variations, presents a minimum principle and an explicit representation of the adjoint process (Lagrange multiplier) when the control enters the drift coefficient f. His approach is based on the robust version of (1.1.6) which satisfies a linear PDE obtained via a gauge transformation $\rho(x,t) = \exp\{y_t h(t,x_t)\} q(x,t)$. Using the backward stochastic PDE's $V(x,t)$, and its robust version $u(x,t)$ which are the adjoint to $\rho(x,t)$, $q(x,t)$, respectively, he then obtains an equation for the adjoint-process

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

by applying the result of the Kunita and Watanabe [97] for martingales with values in L^2 spaces.

Bensoussan in a recent book [15] presents a new method for proving the equation satisfied by the adjoint process above. His method involves a Gelerkin approximation procedure using a finite-dimensional base vector for the Sobelev space H^1 defined by

$$H^1 \triangleq \{u \in L^2; \frac{\partial u}{\partial x^i} \in L^2, i = 1, \dots, n\}.$$

Haussmann [68], using strong variations, presents a minimum principle which depends on the representation of the adjoint-process when the control enters the drift coefficient. His adjoint-process is obtained by performing strong variations on the robust version of the conditional expectation which satisfies the backward PDE u(x,t). However, the description of the adjoint-process is given in terms of only a characterization of the conditional expectation, and this makes the result very difficult to implement especially for nonlinear systems.

Elliott and Yang [50] introduce a minimum principle by considering the differentiation of Blagoveščenskii and Freidlin [22] when the control enters the drift coefficient f and the diffusion coefficient σ. Their minimum principle depends on the description of the adjoint-process in terms of a conditional expectation, and therefore is again difficult to implement.

1.2.2 Proposed Method of Solution.

Our approach is completely different than the ones mentioned above due to the correlation between state and observation noise which prevents us from using the robust

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

version of the unnormalized conditional density since such version does not exist when the differential operator $M_k(t)$ is of the first-order. Moreover, our approach is different in that instead of considering perturbations associated with the measure-valued process $\rho(x,t)$ we introduce a further separation of the problem (1.2.2), (1.2.3) by decomposing $\rho(x,t)$ into two measure-valued processes μ_t, v_t satisfying

Case (i)

$$d\mu_t(\phi) = \sum_{k=1}^{d} \mu_t(M_k(t)\phi)\cdot dy_t^k \quad , \mu_{t_0} = \delta_x \tag{1.2.4(i)}$$

$$dv_t(\phi) = v_t(\mu_t L^u(t)\mu_t^{-1}\phi)dt \quad , v_{t_0} = \delta_x \tag{1.2.5(i)}$$

Case (ii)

$$d\mu_t(\phi) = \sum_{k=1}^{d} \mu_t(h_k(t,x_t)\phi)\cdot dy_t^k \quad , \mu_{t_0} = \delta_x \tag{1.2.4(ii)}$$

$$dv_t(\phi) = v_t(\mu_t L^u(t)\mu_t^{-1}\phi)dt \quad , v_{t_0} = \delta_x \tag{1.2.5(ii)}$$

where case (i) corresponds to control variable appearing only on the drift coefficient f with correlation between state and measurement process present and case (ii) corresponds to control variable appearing in both drift and diffusion coefficients f,σ, respectively when the above correlation is zero. As a result of the above decompositon any control variation would only affect the measure-valued process v_t which is the only process with explicit dependence on control functions u(t). This decomposition is used by Kunita [94, 96] to prove existence uniqueness and smoothness of the unnormalized conditional density $\rho(x,t)$. It is felt that the decomposition above allows us to generalize the Euclidean deterministic

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

variational methods given by Fleming and Rishel [56] and the Euclidean variational methods presented by Bensoussan [9, 11], where necessary conditions for state-valued completely observed stochastic processes are given to measure-valued processes in L^2 spaces. In order to successfully solve the optimization problem considered above we are required to determine equations that are satisfied by the inverse measure-valued processes $\mu_t^{-1}(f)$ and $v_t^{-1}(f)$, which are also shown to be the adjoints to the measure-valued processes $\mu_t(f)$, $v_t(f)$, respectively. We now give a brief outline of two solution procedures employed in obtaining a minimum principle and an explicit representation of the adjoint-process with values in L^2 spaces for case (i) which is also applicable to case (ii) once we establish certain technical conditions; details will be presented in Chapter 4.

Approach 1.

Case i

The approach we employ in obtaining the minimum principle is based on the previously given decomposition. However, to obtain the representation of the adjoint process we use the results of Bismut [16] and Kwakernaak [101] with the important difference that in our case the adjoint process is not originally assumed to be output feedback (i.e., $\mathcal{F}^{y}_{s,t}$-adapted) but it is formulated to satisfy this output feedback requirement.

Step 1. The minimization problem we consider is the one given by (1.2.2),(1.2.3) where $\rho_t \triangleq \rho(x,t)$ is given by the composition $v_t \circ \mu_t$ which we shall denote by $v_t\mu_t$ for

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

simplicity reasons. Thus, the measure-valued process described by (1.2.3) is now decomposed into (1.2.4(i)), (1.2.5(i)) or (1.2.4(ii)), (1.2.5(ii)). Due to this decomposition any control variation affects only the measure-valued process v_t; then a representation for the perturbed measure-valued process is obtained using Fleming and Rishel [56, Chp. 2, Thm. 10.2, pp. 38].

<u>Step 2</u>. We then derive the Gâteaux derivative of $J(u(\bullet))$, which is a function on the Hilbert space L^2 adapted to the filtration $\mathcal{F}^y_{s,t}$.

<u>Step 3</u>. Next, we introduce a process P_t which is the adjoint to the perturbed measure-valued process of Step 1 using an extension of Fleming and Rishel [56, Chp. 2 Thm. 11.1, p. 41] to process with values in L^2 spaces.

<u>Step 4</u>. At this point we perform the composition of the perturbed measure-valued process of Step 1 with its adjoint-process of Step 3. Then, substituting the result of the composition above into the variational cost of Step 2, the minimum principle is obtained in the L^2 space.

<u>Remark 1.2.1</u> The minimum principle we obtain in Step 4 is a general formulation of the minimum principle presented by Bensoussan [10, Theorem 2.1]. Moreover, the perturbed process considered by Bensoussan [15] is a special case of the process obtained by the composition of the perturbed process of Step 2 with the process described by (1.2.4(i)) when $M_k(t) \to h(t,x_t)$. Due to the approach taken, the adjoint process appearing in the minimum principle is an $\mathcal{F}^y_{s,t}$-adapted process (no output feedback assumption was needed).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Step 5. Next we present the stochastic PDE satisfied by the adjoint-process by using the result of Kwakernaak [101], which is an extension of Bismut [16], without having to assume that this process is output feedback as Kwakernaak does. We also show how one can recover the general necessary conditions presented by Kwakernaak [101] using our methodology.

Approach 2.

Case i

The approach we employ here requires no previously known result of stochastic partially observed diffusions. It is, however, more complete in the sense that we give a rigorous justification of the stochastic minimum principle without having to use the result of Kwakernaak [101] as an initial motivation.

Step 1. As a starting point we consider the perturbed measure-valued process of procedure 1, given in Step 1. That is, we do not use the equation satisfied by the unnormalized conditional density (1.2.3) but we treat the problem as one of the decomposed form presented by (1.2.4(i)), (1.2.5(i)).

Step 2. Next, we find a representation for the measure-valued process defined by the composition $\psi_{s,t} \triangleq z_{s,t}v_{s,t}^{-1}$. Then we express the variational cost of Approach 1, Step 2, in terms of the new process $\psi_{s,t}$.

Step 3. The minimum principle is then established by using a representation theorem stated by Liptser and Shiryayev [103, Chp.4, Thm. 4.6, pp. 128-130]. The above representation theorem allows us to express an $\mathcal{F}^{y}_{s,t}$-adapted process in terms of a

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

martingale adapted to the filtration $\mathcal{F}^y_{s,t}$ where the process $\{y_\tau, s \le \tau \le t\}$ is in this case a Wiener process.

Remark 1.2.2 The minimum principle obtained in Step 3 has the exact same form as the one given in Approach 1, Step 4. The adjoint-process as identified by the minimum principle is expressed in terms of a composition of three measure-valued processes.

Step 4. Finally, we derive the stochastic PDE satisfied by the adjoint-process identified in step 3. This adjoint-process satisfies the stochastic PDE of Approach 1, Step 5 having an additional term which is a driven by the observation $\{y_\tau, s \le \tau \le t\}$.

Remark 1.2.3 The two procedures outlined above can be adapted to handle the case when no correlation is present but f, σ are control dependent (i.e., case (ii)) by choosing the supremum norm as the new metric space.

Remark 1.2.4 It turns out that the adjoint-process of Approach 2 step 4 which is a measure-valued process satisfying a backward stochastic PDE is very similar to the state-valued adjoint-process given by Bensoussan [10, Sect. 4, p. 31, equation 4.17] for the case of a completely observed control problem in the Euclidean space. It is believed that the similarity is due to the decomposition (1.2.4) , (1.2.5) that allows us to solve the partially observed problem as a completely observed problem since v_t does not depend on the observations $\{y_\tau, s \le \tau \le t\}$ explicitly. The kind of separated control problem presented here in the L^2 space environment can be viewed as a generalization of the one established by Bensoussan [9, pp. 234-243] for processes with values in Euclidean space.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

1.3 THE NONLINEAR DECISION PROBLEM

We are given a measurable space $(\Omega, \mathcal{F}_t)$ and two probability measures $\mathcal{P}_0$, $\mathcal{P}_1$ defined on it. Given the d-dimensional vector $\{y_s; 0 \leq s \leq T\}$ of observations, determine, so as to minimize an expected risk function, which of the following hypotheses is true:

$$(\mathcal{P}_1; H_1): dy_t = h_t^1 dt + dw_t^1 \tag{1.3.1}$$

$$(\mathcal{P}_0; H_0): dy_t = h_t^0 dt + dw_t^0 \tag{1.3.2}$$

where w_t^i, $i = 0, 1$ are d-dimensional Wiener processes h_t^i, $i = 0, 1$ are d-dimensional random signal processes, and $\mathcal{P}_0$, $\mathcal{P}_1$ are induced by H_0, H_1, respectively.

The generalized likelihood-ratio (LR) for the above problem has been shown by Duncan [45] and Kailath [80, 81, 82] to be

$$\Lambda_T = e^{\int_0^T (\hat{h}_t^1 - \hat{h}_t^0)^T dy_t - \frac{1}{2} \int_0^T (|\hat{h}_t^1|^2 - |\hat{h}_t^0|^2) dt} \tag{1.3.3}$$

where $\hat{h}_t^i$ is the conditional expectation of h_t^i given the observations up to time t when hypothesis H_i is true. The likelihood-ratio (1.3.3) is obtained by assuming $\mathcal{P}_1$ is absolutely continuous with respect to $\mathcal{P}_0$ and defining

$$\Lambda_T \triangleq E_0 \left(\frac{d\mathcal{P}_1}{d\mathcal{P}_0} \middle| \mathcal{F}_T^y\right) \tag{1.3.4}$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

where E_0 denotes expectation with respect to measure $\mathcal{P}_0$ when restricted to the σ-algebra $\mathcal{F}_t^y$. A derivation of (1.3.2) based on martingale theory for both continuous and discontinuous process is given by Hibey [74, 75]. The likelihood-ratio test involves testing Λ_t against a given threshold γ, that is, performing the test

$$\Lambda_t \underset{H_0}{\overset{H_1}{\gtrless}} \gamma\,; \tag{1.3.5}$$

we decide in favor of hypothesis H_1 if Λ_t is greater than γ and decide in favor of hypothesis H_0 if Λ_t is less than γ. If $\Lambda_t = \gamma$, then H_0, H_1 are equally probable.

Now, once we adapt the decision strategy above, the next step is to study the error performance. In our case, there are two possibilities of making errors. The first is called "false alarm" and is denoted by P_F; it is the error of deciding hypothesis H_1 is true when in fact hypothesis H_0 is true. The second is called a "miss" and is denoted by P_M; it is the error of deciding hypothesis H_0 is true when in fact hypothesis H_1 is true. The detector peformance is completely characterized once P_F, P_M are known. However, in theory, to compute them one need the distribution of Λ_t, and this is often unknown. Hence, one is forced to consider bounding techniques which can be computed exactly.

The bounding technique we are concerned with is the so called Chernoff bound as given by Van Trees [128]. The extension of this bound to the nonlinear detection problem is presented in Evans [52] and Hibey [74]. The Chernoff bound on P_F is derived for $s > 0$ as follows:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$P_F = \text{Prob}\,\{\text{accept } H_1 | H_0 \text{ is true}\} \tag{1.3.6}$$
$$= E_0\,\{I_{\{\omega\in\Omega;\ \ln\Lambda_t(\omega) > \ln\gamma\}}\}.$$

The second equality follows from (1.3.5) since the logarithm is a monotonic increasing function of its argument. Proceeding by writing the moment generating function of $\ln \Lambda_t$ as

$$E_0[e^{s \ln \Lambda_t}] = \int_{-\infty}^{\infty} e^{s \ln \Lambda}\, p(\Lambda | H_0)\, d\Lambda$$
$$\geq \int_{\gamma}^{\infty} e^{s \ln \Lambda}\, p(\Lambda | H_0)\, d\Lambda,$$

we have

$$P_F \leq e^{-s \ln \gamma}\, E_0[\Lambda_t^s],\ s > 0. \tag{1.3.7}$$

The expectation E_0 is understood to be restricted to the σ-algebra $\mathcal{F}_t^y$. Therefore to obtain the tightest bound on P_F, a minimization leads to

$$P_F \leq \min_{0<s<1}\ e^{-s \ln \gamma}\, E_0[\Lambda_t^s]. \tag{1.3.8}$$

Similarly, to find the tightest bound for P_M one has to consider

$$P_M \leq \min_{-1<s<0}\ e^{-s \ln \gamma}\, E_1[\Lambda_t^s]. \tag{1.3.9}$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

It can be easily shown that the P_M can be expressed in terms of expectation with respect to measure $\mathcal{P}_0$ by

$$P_M \leq \min_{-1<s<0} e^{-s \ln \gamma} E_0[\Lambda_t^{s+1}]. \tag{1.3.10}$$

More detailed discussion on this subject is found in Hibey [74]. Before we present our approach let us present the method of solution as given by Van Trees [128], Evans [52], Hibey [74, 75], and Beneš [7].

1.3.1 Previous Method of Solution.

In the textbook by Van Trees [128, Chps. 1-2] the decision problem is treated for the case when the signal h_t^i is either a known function or some of its components are random such that no filtering estimate for h_t^i is required. In this case the LR test given by (1.3.5) is expressed as a ratio of two Gaussian densities. The performance bounds P_F, P_M are given by (1.3.8), (1.3.10) respectively, where Λ is again the ratio of two Gaussian densities.

Evans [52] considers the decision problem of (1.3.1), (1.3.2) when the signal $h^i(t, x_t^i)$ is a nonlinear random function with the state process x_t^i satisfying a nonlinear diffusion process. His concern is mainly the development of the expressions (1.3.8) and (1.3.9) as well as optimal and sub-optimal methods of solving for these bounds. Using some of the properties of Markov processes he is able to obtain an evolution type partial differential equation which is related to the evaluation of the Chernoff bounds P_M, P_F

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

above. The coefficients of the partial differential equation are in terms of optimal or sub-optimal estimates.

Hibey [74, 75 treats the nonlinear decision problem studied by Evans for both continuous and discontinuous processes, but unlike Evans, he uses the theory of martingales and measure transformations to extend the result of Evans [52]. He then presents a partial differential equation which is related to the evaluation of the Chernoff bounds P_M, P_F for both optimal and sub-optimal estimates. This partial differential equation is the adjoint equation to the one obtained by Evans [52].

Recently, Beneš [7] treated the sonar decision problem with emphasis on the use of nonlinear filtering techniques to formulate a general Bayesian model. His treatment considers two kinds of optimal detectors, fixed-time interval and sequential.

1.3.2 Proposed Method of Solution.

Our approach differs from the one taken by Evans [52] and Hibey [74, 75] above. Here, we formulate the nonlinear decision problem of evaluating (1.3.5), (1.3.8), (1.3.10) in terms of the unnormalized conditional density. However, we also consider as our basic tool the use of measure transformations which in this case are restricted to a bigger filtration $\mathcal{F}_t$ containing the filtration $\mathcal{F}_t^y$. It is felt that our approach would allow for the evaluation of the decision strategy (1.3.5) as well as the performance bounds P_F, P_M exactly, even for certain class of nonlinear systems. This is indeed an improvement over the approach taken by Evans [52] and Hibey [74, 75] since the PDE developed by the above authors is usually solvable for the linear Gaussian case only. We also notice that

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

the decision strategy and performance bounds we employ have as a special case the decision strategy and performance bounds obtained by Van Trees [128]. We now briefly present the approach to be followed in Chapter 5.

<u>Step 1</u>. Our initial concern is to express the well-known likelihood-ratio (1.3.3) in terms of yet another likelihood-ratio Ψ_t restricted to the filtration $\mathcal{F}_t \supset \mathcal{F}_t^y$ such that the LR Λ_t can be expressed as $\Lambda_t = E_0(\Psi_t | \mathcal{F}_t^y)$ where $\Psi_T = E_0\,[\frac{dP_1}{dP_0} | \mathcal{F}_T]$. This new representation of Λ_t allows us to obtain a new method for solving the decision problem without having to find equations that describe the evolution of the signal estimates $\hat{h}^i(t, x_t^i)$ as others do.

<u>Step 2</u>. Next we prove the Theorem that permits us to represent Λ_t in terms of a ratio of two conditional expectations with respect to the filtration $\mathcal{F}_t^y$ generated by the process $\{y_s, 0 \le s \le t\}$, which under this new measure becomes a standard Brownian motion. Then the likelihood-ratio test (1.3.5) is expressed as the ratio of two conditional densities integrated over the space $R^n \otimes R^n$. Each conditional density satifies a stochastic PDE with random coefficients much as in the case of the nonlinear filtering problem presented in Section 1.1.

<u>Step 3</u>. Having established the decison strategy, we next address the Chernoff bounds P_F, P_M as given by (1.3.8), (1.3.9) (or (1.3.10)), respectively. We first express P_F given by (1.3.8) using the formulation of Step 1. Similarly, we repeat the same procedure for P_M given by (1.3.10). Then by introducing another measure as in Step 2

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

we derive expressions for P_F, P_M as a ratio of two conditional densities integrated over the space $R^n \otimes R^n$. As in Step 2, each conditional density satisfies a stochastic PDE.

Step 4. Finally, we demonstrate the effectiveness of this approach by solving certain linear and nonlinear decision problems. We conclude that if the nonlinear filtering problem associated with the decision problem is solvable by means of solving a single stochastic PDE, the decision strategy and the exact Chernoff bounds can be evaluated. This is indeed the original goal for studying such decision problems.

Remark 1.3.1 In Step 4 we stated that if a single stochastic PDE is solvable the two questions concerning the decision problem can be answered. This is due to imbedding the two stochastic PDE's of Step 2 and Step 3 into a single stochastic PDE which then generates the solution of the above two stochastic PDE's.

1.4 ORGANIZATION OF THESIS

In this thesis we study topics in the theory of nonlinear estimation, nonlinear decision, and nonlinear stochastic control with partial observations.

In Appendix 7.A we discuss the mathematical concepts of differential geometry and its applications to linearization of nonlinear deterministic control systems. Here we define vector fields and introduce some notation. Our attention is focus on the Frobenius theorem and its application to local and global linearization of nonlinear control systems.

Appendix 7.B contains a survey of martingales. Here we define martingales, predictable processes and introduce notation. In addition we discuss stochastic integrals with respect to martingales and discuss topics such as Ito and Fisk-Stratonovich

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

differential rules in both forward and backward direction. We also provide the connection between Ito and Fisk-Stratonovich integrals.

Appendix 7.C serves as a brief introduction on strong and weak solutions to stochastic differential equations. While doing so, we present conditions of existence and uniqueness of strong and weak solutions.

Appendix 7.D contains a discussion of the stochastic differential geometry and its applications to stochastic differential equations. Here we present Fisk-Stratonovich differential rules when the drift and diffusion coefficients are represented using the tangent space basis. The connection between Ito and Fisk-Stratonovich integrals is also given. Next, we give the representation of a forward semimartingale in terms of the representation of a backward semimartingale using the definition of a differential maps. Finally, we indroduce the inverse map which is expressed in terms of the solution of a certain backward stochastic differential equation.

Appendix 7.E contains a discussion of the measure transformations, translation of martingales and sufficient conditions of absolute continuity of measures. Here we begin by introducing the exponential formula which allow us to define new measures. The connection between conditional and unconditional expectations under two different measures which are absolutely continuous is also presented here.

Finally, in Appendix 7.F we present some aspects of the semigroup theory of Markov processes. We begin with an introduction on transition functions of Markov processes and then present two families of operators which can be associated with

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

transition probabilities, and finally we give the backward and forward equations which are related to the above transition probabilities.

Chapter 2 contains a detailed exploration of representations of the unnormalized conditional density. Section 2.2 begins with a precise problem statement for the partially observed stochastic control problem.

In Section 2.3, we present a proof of the fundamental solution of the equation satisfied by the unnormalized conditional density (the DMZ equation) using the representation introduced by Friedman [60, 61, 62]. Then we give the PDE satisfied by the robust version of the DMZ equation and introduce the definition of the fundamental solutions to PDE's. Next, we represent the solutions of the backward SDE and its robust version which were originally derived by Clark [33] and Pardoux [113, 114], respectively, to their fundamental solutions.

In Section 2.4, we present the partially observed stochastic control problem by adapting the dynamic programming formulation used for completely observed systems. Here the problem is first converted into a completely observed stochastic control problem by viewing the unnormalized conditional density as the new state of the system.

Finally, in Section 2.5, we present the derivation of the unnormalized conditional density of a diffusion process when terminating at the first exit time from a domain of interest. The proof is an extension of the one presented by Hibey [74, 75] for unconditioned diffusion processes.

In Chapter 3, we present a detailed analysis in estimating diffusion processes by using diffeomorphism and measure transformations.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Section 3.1 begins with a precise problem statement and Section 3.2, Section 3.3 provide the necessary and sufficient conditions of local and global linearization of stochastic differential systems both locally and globally, respectively.

In Section 3.4, we present sufficient conditions for solving a Feynman-Kac formula whose solution is related to the solution of the DMZ equation using the developments of Section 3.3. Finally, we conclude this section by relating the solution to the Feynman-Kac formula in terms of a stochastic control problem using the methodology first introduced by Fleming [57].

In Section 3.5, we express the unnormalized conditional density of a diffusion process defined locally in terms of an initial-boundary value problem.

In Section 3.6, we derive lower and upper bounds on function of state estimates using the conditional correlation coefficient.

Finally, in Section 3.7, we extend the previous work of Bobrovsky and Zakai [23, 24] who derived lower bounds for nondegenerate nonlinear filtering problems to degenerate filtering problems which satisfy the linearization conditions of Sections 3.3, 3.4.

In Chapter 4, we present a detailed solution of the nonlinear partially observed stochastic control problem using two approaches, when the control variable is present in the drift coefficient and correlation between state process and observation process is allowed, and when the control variable is present in both drift and diffusion coefficients and no correlation is allowed. The methods and results appeared in this chapter are new.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Section 4.1 begins with a precise problem statement and a derivation of the equation satisfied by the unnormalized conditional density when correlation between state process and observation process is present and the diffusion process is of degenerate type.

Section 4.2 provides derivations of the decomposed measure-valued processes in both forward and backward direction. Equations satisfied by the inverse maps of the above measure-valued processes are also derived.

Section 4.3 presents the first approach to the stochastic control problem by providing a rigorous derivation for the minimum principle and a formal derivation for the equation satisfied by the adjoint-process. The second case is also discussed here.

Section 4.4 presents the second approach to the stochastic control problem by providing a rigorous derivation for both minimum principle and the equation satisfied by the adjoint-process. The second case is also discussed here.

In Chapter 5, we present a detailed solution of the decision problem. The representations and results of this chapter are new.

Section 5.1 begins with a precise problem statement of the problem we propose to solve. Then we give the derivation that allows us to represent the generalized likelihood-ratio restricted to the filtration generated by the observation in terms of a likelihood-ratio restricted to a bigger filtration.

In Section 5.2 we represent the generalized likelihood-ratio in terms of a ratio of two unnormalized conditional densities satisfying stochastic PDE's.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

In Section 5.3 we provide expressions for evaluating the error probabilities which are due to Chernoff in terms of a ratio of two unnormalized conditional densities satisfying stochastic PDE's.

Section 5.4 contains linear as well as nonlinear decision examples which can be solved exactly.

Finally, we present Chapter 6 which is the concluding chapter, where Section 6.1 provides a brief summary of our work and indicate its main contribution, and Section 6.2 contains some topics for further research.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

CHAPTER 2

EVOLUTION OF UNNORMALIZED CONDITIONAL DENSITY AND ITS REPRESENTATIONS

2.1 INTRODUCTION

Much of the modern nonlinear filtering and stochastic control problems are analyzed using the unnormalized conditional density that satisfies a linear stochastic partial differential equation (the DMZ equation). Answers to questions such as existence of admissible control to problems of a nonlinear nature were first presented by Fleming and Pardoux [53,54], and later by Bismut [20], Elliott and Kohlmann [49] and others through this unnormalized density. Also, exact solutions to linear filtering problems and linear partially observed control problems with non-Gaussian initial distributions are obtained by Beneš and Karatzas [6], using a version of the Kallianpur-Striebel formula [85] as a function space integral.

Results on the existence and uniqueness of the solution of DMZ equation are discussed by Pardoux [112,114] and Krylov and Rozovskii [90] when the Kolmagorov's operator is elliptic. Pardoux [112,114] proves existence and uniqueness of the solution using variational methods to stochastic PDE's while Krylov and Rozovskii [90] prove existence and uniqueness using Sobolev spaces. Furthermore Kunita [93,94,96] proves such existence and uniqueness of solution when the Kolmogorov's operator is degenerate by expressing the solution as the conditional expectation of some suitable stochastic

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

process. Then, using probabilistic methods and expressing any L^2 solution as an infinite sum of multiple Wiener Ito integrals, he proves uniqueness.

Derivations of the DMZ equation can be found in Zakai [134], Wong [131], Pardoux [113,114,115], Kunita [93,96] and Bensoussan [12]. Pardoux in his work [113,114,115] provides a pair of stochastic PDE's, one with respect to the forward variable and the other with the respect to the backward variable. The forward equation is the DMZ equation and the backward equation is its adjoint equation. In the filtering problem, they both play the role of the backward and forward Kolmogorov equations for unconditioned diffusions as presented in Appendix 7.F. He also presents the robust version of the forward and backward stochastic SDE's, with the forward one originally being formulated by Clark [33].

In this section we shall utilize certain results from the theory of partial differential equations and their fundamental solutions as given by Friedman [60,61,62] to construct fundamental solutions for the two pairs of stochastic and partial differential equations related to the nonlinear filtering problem. We shall show that the fundamental solution of the unnormalized density is exactly equal to a version of the Kallianpur-Striebel formula considered by Beneš [4] and Beneš and Karatzas [6] which were then used to obtain the finite-dimensional statistics associated with the nonlinear filtering problem and partially observed linear stochastic control problem with non-Gaussian initial conditions. Next, using the fundamental solution of the DMZ equation, we shall give a representation for the value function of a partially observed strict sense stochastic control problem in a dynamic programming formulation. Finally, we shall consider the filtering problem when

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

the unobserved state process is defined on an open bounded set in R^n having an exit time $\tau < T$ w.p.1. Using the approach of Dynkin [46], Prohorov and Rozanov [116], and Hibey [74], who considered the above problem when the state process is completely observed, we shall derive an evolution-type stochastic partial differential equation for the unnormalized conditional density for the partially observed problem. We shall also present a derivation which is based on the infinitesimal opearator of a diffusion process. This resulting stochastic differential equation can be transformed into a PDE which is classified by Friedman [60] as a first initial-boundary value type problem. This is indeed the equation we shall consider in our subsequent development to analyze the local equivalence of a nonlinear filtering problems.

2.2 PRECISE PROBLEM STATEMENT

We shall formulate the problem as one of partially observed nonlinear stochastic control. Notationwise, we shall address the nonlinear filtering problem by disregarding the performance index and ignoring the pathwise dependence on the control variable.

Consider the following minimization:

$$\min \left\{ J(u);\ u \in U_{ad} \right\}, \tag{2.2.1}$$

$$J(u) = E\left(\int_0^T \Pi(t, x_t, u_t)\, dt + \kappa(x_T) \right) \tag{2.2.2}$$

subject to constraints that describe the state process $\{x_s,\ 0 \le s \le t\}$ and observation process $\{y_s,\ 0 \le s \le t\}$ in terms of the stochastic differential equation and measurement equation

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$dx_t = f(t, x_t, u_t)dt + \sigma(t, x_t)dw_t \ , \ x_{t_0} = x_0 \tag{2.2.3}$$

$$dy_t = h(t, x_t)dt + db_t \qquad , \ y_{t_0} = 0. \tag{2.2.4}$$

We shall make the following assumptions:

(A1) $x_0 \in R^n$ is given;

(A2) $f:[0,T]xR^nxU \to R^n$ is Borel measurable, continuous, continuously differentiable in x, u, with U a Borel set and K_1, K_2 constants such that

$$|f_x(t, x, u)| + |f_u(t, x, u)| \leq K_1$$

$$|f(t, x, u)| \leq K_2(1 + |u| + |x|);$$

(A3) $\sigma:[0,T]xR^m \to R^n \otimes R^m$ is Borel measurable, continuous, continuously differentiable in x, u, and K_3 a constant such that

$$||\sigma(t, x)|| + ||\sigma_x(t,x)|| \leq K_3;$$

(A4) $h:[0,T]xR^n \to R^d$ is Borel measurable, continuous, continuously differentiable in x, and with constant K_4 such that

$$|h(t, x)| \leq K_4(1 + |x|);$$

(A5) ρ_0 is a probability measure on $(R^n, B_T{}^n)$, where $B_T{}^n$ is a family of Borel sets on R^n, and

$$\int_{R^n} |x|^q p_0(dx) < \infty;$$

(A6) $\kappa:R^n \to R$ is continuously differentiable in x with constant K_5 such that

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$|\kappa(x)| + |\kappa_x(x)| \leq K_5(1 + |x|^q)\ ,\ q < \infty;$$

(A7) $\pi:[0,T]xR^nxU \to R$ is Borel measurable continuously differentiable in x, u with constant K_6 such that

$$|\pi(t, x, u)| + |\pi_x(t, x, u)| \leq K_6(1 + |x|^q + |u|^q).$$

Consider the space $\Omega = R^n xC([0,T]; R^m)xC([0,T]; R^d)$, with coordinate functions (x_0, w, y), were w and y independent standard Brownian motions on $R^m xR^d$, and x_0 an independent random variable with density p_0. Let μ_w^k denote a Wiener measure on $C([0,T]; R^k)$ and B_T^K the Borel σ-fields on R^k; then the measure $\tilde{\wp}$ and σ-algebra $\mathscr{F}_t$ on Ω are defined by

$$\tilde{\wp} \triangleq \tilde{\wp}(x, dw, dy) \triangleq \wp_0(dx)\mu_w^m(dw)\mu_w^d(dy),$$

$$\mathscr{F}_T \triangleq B_T^n \otimes B_T^m \otimes B_T^d$$

Define $\mathscr{F}_T^y \triangleq \sigma\{y_s; 0 \leq s \leq T\}$ and $\mathscr{F}_T^w \triangleq \sigma\{w_s; 0 \leq s \leq T\}$.

The set of admissible control functions U_{ad} consists of the $\mathscr{F}_t^y, t \in [0,T]$-adapted functions

$$u:[0,T] \times C([0,T];R^d) \to U$$

which are Borel measurable such that

(A8) $u(t,y_t) \in L_y^2([0, T] \times C([0, T];R^d))$.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Let x_t^u be a unique strong solution of

$$\begin{aligned} dx_t &= f(t, x_t, u(t,y_t))dt + \sigma(t,x_t)dw_t, \quad t \geq 0 \\ x_0 &\sim p_0 \end{aligned} \tag{2.2.5}$$

on the probability space $(\Omega, \mathscr{F}_t, \tilde{\wp})$ for $u \in U_{ad}$, $x \in R^n$.

Suppose we define the process

$$m_t \triangleq \int_0^t h(s,x_s)dy_s \in M(\mathscr{F}_t, \tilde{\wp}) \tag{2.2.6}$$

where $h(t, x_t)$ is an adapted, predictable process. Then by defining the exponential formula as in Theorem 7.E.1,

$$\Lambda_T^u \triangleq \exp\{m_T - \frac{1}{2}\langle m, m\rangle_T\} = \exp\{\int_0^T h(s, x_s)dy_s - \frac{1}{2}\int_0^T |h(s, x_s)|^2 ds\}$$

and using Theorem 7.E.2, we can define the measure $\wp^u$

$$\Lambda_T^u \triangleq \tilde{E}\,[\frac{d\wp^u}{d\tilde{\wp}} \mid \mathscr{F}_T]$$

where $\wp^u$ is absolutely continuous with respect to measure $\tilde{\wp}$. Thus, because $h(t,x_t)$ satisfies a linear growth, $\sigma(t,x)$ is assumed to be bounded, and $f(t,x,u_t)$ satisfies (A2) (see Remark 7.E.2), we obtain $\tilde{E}\,[\Lambda_T^u] = 1$ ($\tilde{E}$ denotes expectation with respect to measure $\tilde{\wp}$) and therefore conclude that $\wp^u$ is a probability measure and, using an extension of

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Theorem 7.C.2 (x^u, y) is a weak solution of (2.2.3), (2.2.4) on the probability space $(\Omega, \mathcal{F}_t, P^u)$. The cost function (2.2.2) can be expressed as

$$
\begin{aligned}
J(u) &= E^u \{ \int_0^T \Pi(t, x_t^u, u(t,y_t))dt + K(x_T^u) \} \\
&= \tilde{E} \{ \Lambda_T^u (\int_0^T \Pi(t, x_t^u, u(t,y_t))dt + K(x_T^u)) \}
\end{aligned}
\tag{2.2.7}
$$

where E^u denotes expectation on $(\Omega, \mathcal{F}_t, P^u)$.

Remark 2.2.1 Superscript u on Λ, E, x indicates their dependence on control variable u. Whenever this dependence is removed, the problem is formulated as one of nonlinear filtering type, represented by (2.2.3), (2.2.4) only.

Remark 2.2.2

(i). under P^u the following properties hold:

(a) $\{w_s, 0 \le s \le t\}$, $\{b_s, 0 \le s \le t\}$ are independent Brownian motion processes;

(b) $\{b_s, 0 \le s \le t\}$ and $\{h(s, x_s^u), 0 \le s \le t\}$ are independent (i.e., the signal is independent of the measurement noise).

(ii). under $\tilde{P}$ the following properties hold:

(a) $\{w_s, 0 \le s \le t\}$, $\{y_s, 0 \le s \le t\}$ are independent Brownian motion processes;

(b) $\{y_s, 0 \le s \le t\}$ and $\{h(s,x_s), 0 \le s \le t\}$ are independent;

(c) $\{x_s^u, 0 \le t \le t\}$ has the same distribution as under measures P^u, $\tilde{P}$;

(d) $P^u \sim \tilde{P}$ with Radon-Nikodyrn derivative

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$\frac{d\rho^u}{d\tilde{\rho}} = \left(\frac{d\tilde{\rho}}{d\rho^u}\right)^{-1} = \mathscr{E}\left(\int_0^T h(s,x_s^u)dy_s\right).$$

Properties of (a) - (d) of Remark 2.2.2 (ii) are evident from the definition of the measure $\tilde{\rho}$ on $(\Omega, \mathscr{F}_t)$ which can be expressed as

$$\tilde{\rho}(A) = \int_A \frac{d\tilde{\rho}}{d\rho^u}(\omega)\ \rho^u(d\omega), \text{ for any } A \in \mathscr{F}_T$$

as in Theorem 7.E.2, where

$$\frac{d\tilde{\rho}}{d\rho^u} = \exp\left\{-\int_0^T h(s,x_s^u)db_s - \frac{1}{2}\int_0^T |(h(s,\ x_s^u)|)^2 ds\right\}$$

is a Radon-Nikodym derivative of $\tilde{\rho}$ with respect to ρ^u. By martingale translation, (Theorem 7.E.3),

$$\begin{aligned}\tilde{b}_t &= b_t - \langle b.,-\int_0^{\cdot} h(s,\ x_s^u)ds\rangle_t \\ &= b_t + \int_0^t h(t,x_t^u)dt \in M(\mathscr{F}_t,\ \tilde{\rho}).\end{aligned}$$

It then follows that $\tilde{b}_t$ is a Wiener process since $\langle\tilde{b},\ \tilde{b}\rangle_t = \langle b,\ b\rangle_t = t$ and the sample paths of b_t are continuous. The equivalence of measures ρ^u, $\tilde{\rho}$ (i.e. Remark 2.2.2 (d)) follows from Remark 7.E.3 since assumption (A4) implies

$$\tilde{\rho}\ \left(\int_0^T |h(t,\ x_t)|^2 dt < \infty\right) = 1.$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

2.3 DERIVATION OF FUNDAMENTAL SOLUTION ASSOCIATED WITH THE NONLINEAR FILTERING PROBLEM

Suppose we are concerned with the filtering problem (2.2.3), (2.2.4). If we define a new probability measure as in the previous section, then by (7.E.6), for any integrable function $\Phi(\cdot)$,

$$E\,[\Phi(x_t)\,|\,\mathscr{F}_t^y] = \frac{\tilde{E}\,[\Lambda_T\,\Phi(x_t)\,|\,\mathscr{F}_t^y]}{\tilde{E}\,[\Lambda_T\,|\,\mathscr{F}_t^y]}, \tag{2.3.1}$$

which relates the statistics of Φ evaluated under measure $\mathcal{P}$ to that of measure $\tilde{\mathcal{P}}$. Following the approach taken by Zakai [134] and Wong [131], rather than derive the DMZ equation, we shall instead derive the stochastic PDE satisfied by the fundamental solution of the DMZ equation. Thus, by reconditioning on x_t and x_{t_0},

$$E\,[\Phi(x_t)\,|\,\mathscr{F}_t^y] = \frac{\tilde{E}\,\{\tilde{E}\,(\Lambda_t\,\Phi(x_t)\,|\,\mathscr{F}_t^y,\,x_t,\,x_{t_0})\,|\,\mathscr{F}_t^y\}}{\tilde{E}\,\{\tilde{E}\,(\Lambda_t\,|\,\mathscr{F}_t^y,\,x_t,\,x_{t_0})\,|\,\mathscr{F}_t^y\}}. \tag{2.3.2}$$

Next, by defining $\Lambda(x_{t_0,t}\; y_{t_0,t}) \triangleq \Lambda_t$, and $\tilde{\Lambda}(z, x, y_{t_0,t}) \triangleq \tilde{E}(\Lambda_t\,|\,\mathscr{F}_t^y,\,x_t{=}z,\,x_{t_0}{=}x)$, and setting

$$P(z,t) = \mathrm{Prob}\{x_t \le z\},\quad P(x,t_0) = \mathrm{Prob}\{x_{t_0} \le x\}$$

(2.3.2) is expressed as

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$E\,[\Phi(x_t) \mid \mathcal{F}_t^y] = \frac{\int_{R^n \otimes R^n} \Phi(z)\tilde{\Lambda}(z, x, y_{t_0,t})\, \tilde{P}(dz, dx, t, t_0 \mid \mathcal{F}_t^y)}{\int_{R^n \otimes R^n} \tilde{\Lambda}(z, x, y_{t_0,t})\, \tilde{P}(dz, dx, t, t_0 \mid \mathcal{F}_t^y)}$$

where

$$\tilde{P}\,(z, x, t, t_0 \mid \mathcal{F}_t^y) = \mathrm{Prob}\{x_t \le z,\ x_{t_0} \le x \mid \mathcal{F}_t^y\} = P(z,x,t,t_0)$$

which follows by Remark 2.2.2 (ii), (b), (c). Thus,

$$E\,[\Phi(x_t) \mid \mathcal{F}_t^y] = \frac{\int_{R^n \otimes R^n} \Phi(z)\tilde{\Lambda}(z, x, y_{t_0,t})\, P(dz, dx, t, t_0)}{\int_{R^n \otimes R^n} \tilde{\Lambda}(z, x, y_{t_0,t})\, P(dz, dx, t, t_0)}.$$

Using the Markov property of x_t and the fact that $P(dz, dx, t, t_0)$ is absolutely continuous with respect to the Lebesque measure,

$$E\,[\Phi(x_t) \mid \mathcal{F}_t^y] = \frac{\int_{R^n \otimes R^n} \Phi(z)\tilde{\Lambda}(z, x, y_{t_0,t})\, p(z, t; x, t_0) p_{t_0}(x) dx dz}{\int_{R^n \otimes R^n} \tilde{\Lambda}(z, x, y_{t_0,t})\, p(z, t; x\ t_0) p_{t_0}(x) dx dz} \tag{2.3.3}$$

where $p(z, t; x, t_0)$ is the conditional density of $x_t = z$ given $x_{t_0} = x$.

Further, if $E\,[\Phi(x_t) \mid \mathcal{F}_t^y] = \int_A \Phi(z) p(z, t \mid \mathcal{F}_t^y) dz$, where $p(x, t \mid \mathcal{F}_t^y)$ is the normalized conditional density of x_t given $\mathcal{F}_t^y$, for any Borel set $A \in \mathcal{F}_t$, we have

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$\mathrm{Prob}\{x_t \in A \mid \mathcal{F}_t^y\} = \int_A p(z, t \mid \mathcal{F}_t^y)dz = E\,[I_{\{\omega;\, x_t(\omega)\in A\}} \mid \mathcal{F}_t^y]$ and from (2.3.3), since

$P(dz, dx; t, t_0)$ is absolutely continuous with respect to the Lebesque measure, we have

$$p\,(x_t,t \mid \mathcal{F}_t^y) = \frac{\int_{R^n} \tilde{\Lambda}(z, x, y_{t_0,t})\; p(z, t; x, t_0)\; p_{t_0}(x)dx}{\int_{R^n \otimes R^n} \tilde{\Lambda}(z, x, y_{t_0,t})\; p(z, t; x, t_0)\; p_{t_0}(x)dxdz}\,. \tag{2.3.4}$$

Remark 2.3.1

Recall that the unnormalized conditional density $\rho(\cdot, \cdot)$ of x_t given the filtration $\mathcal{F}_t^y$ satisfies the stochastic PDE (1.1.6). Referring to (2.3.4), we shall show that

$$r(z, t; x, t_0) = \tilde{\Lambda}(z, x, y_{t_0, t})\; p(z, t; x, t_0) \tag{2.3.5}$$

is the fundamental solution of the DMZ equation. Notice also that the numerator of (2.3.4) equals the unnormalized density $\rho(z,t)$. The denominator of (2.3.4) is its normalization part.

The following definition is from Friedman [60, pp.3].

Definition 2.3.1 A fundamental solution of the DMZ (1.1.6) is a function $r(z, t; x, s)$ defined for all (z, t) and (x, s) in $R^n x[0,T]$, $t > s$, satisfying

$$\rho(z, t) = \int_{R^n} r(z, t; x, s)p_s(x)dx. \tag{2.3.6}$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Theorem 2.3.1

Suppose a fundamental solution r(z, t; x, s) for the DMZ equation exists. Then r(z,t; x,s) as a function of (z, t) satisfies

$$\begin{aligned} dr(z,t;x,s) &= L(t)^* r(z,t;x,s)dt + h(t,z)r^*(z,t;x,s)dy_t, \quad s \le t \le T \\ \lim_{t \downarrow s} r(z,t;x,s) &= \delta(z - x) \end{aligned} \tag{2.3.7}$$

with

$$p(z,t \mid \mathcal{F}_t^y) = \frac{\int_{R^n} r(z,\ t;\ x,\ s)p_s(x)dx}{\int_{R^n \otimes R^n} r(z,\ t;\ x,\ s)p_s(x)dx\,dz}.$$

Proof: To prove (2.3.7) we follow the same approach as presented by Zakai [134] and Wong [131] up to the point where we define r(z,t; x,s) as in (2.3.5). Thus, applying the Ito differential rule to Λ_t and then taking conditional expectation, we obtain, respectively,

$$\Lambda_t = 1 + \int_{t_0}^{t} \Lambda_s h^T(s,x)dy_s,$$

$$\tilde{E}\,(\Lambda_t \mid \mathcal{F}_{t_0,t}^y, x_t, x_{t_0}) = 1 + \tilde{E}\,\{\int_0^t \Lambda_s h^T(s,x)\,dy_s \mid \mathcal{F}_{t_0,t}^y, x_t, x_{t_0}\}.$$

For $t \ge s \ge 0$ define

$B_1 = \mathcal{F}_{t_0,s}^{x,y} \triangleq \sigma\{x_u, y_u, t_0 \le u \le s\}, B_2 \triangleq \mathcal{F}_{s,t}^y \triangleq \sigma\{y_\tau - y_s, s \le \tau \le t\}$ and $B \triangleq \mathcal{F}_{t_0,s}^y \vee \{x_t, x_0\}$.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Since under measure $\tilde{\wp}$, $\mathcal{F}^{x}_{t_0,s}$ is independent of $\mathcal{F}^{y}_{t_0,t}$ and $\{y_s, t_0 \le s \le t\}$ is a Brownian motion, it follows that $\tilde{B}_1 \in \mathcal{F}^{x,y}_{t_0,s}$ and $\tilde{B}_2 \in \mathcal{F}^{y}_{s,t}$ are independent given B, i.e.,

$$\tilde{\wp}\{\tilde{B}_1\ \tilde{B}_2 \,|\, B\} = \tilde{\wp}\{\tilde{B}_1 \,|\, B\}\ \tilde{\wp}\{\tilde{B}_2 \,|\, B\}.$$

By Theorem 7.E.4, for any $A \in \mathcal{F}^{x,\,y}_{t_0,s}$

$$\tilde{\wp}\{A \,|\, B_2, B\} = \tilde{\wp}\{A \,|\, B\}.$$

Hence, $\tilde{E}\ (\Lambda_s\ h(s,\ x) \,|\, \mathcal{F}^{y}_{t_0,t}\ V\{x_t,\ x_{t_0}\}) = \tilde{E}\ (\Lambda_s\ h(s,\ x) \,|\, \mathcal{F}^{y}_{t_0,s}\ V\{x_t,\ x_0\})$.

Using the smoothing property of conditional expectation

$$\tilde{E}(\Lambda_s\, h(s,x) \,|\, \mathcal{F}^{y}_{t_0,s}\, V\{x_t,x_{t_0}\}) = \tilde{E}\{\tilde{E}(\Lambda_s\, h(s,x) \,|\, \mathcal{F}^{y}_{t_0,s}\, V\{x_t,x_s,x_{t_0}\}) \,|\, \mathcal{F}^{y}_{t_0,s}\, V\{x_t,x_{t_0}\}\}.$$

Because of the Markov property of (x_s, y_s), $\mathcal{F}^{x,y}_{t_0,s}$ and x_t are conditionally independent given $\{x_{t_0}, x_s\}$ and $\{y_s\}$, therefore B_1 and $\{x_t\}$ are also independent given $\{x_{t_0}, x_s\}$ and B, that is,

$$\tilde{\wp}(\tilde{B}_1 V\{x_t\} \,|\, \{x_{t_0},x_s\}\ V\mathcal{F}^{y}_{t_0,s}\}) = \tilde{\wp}(\tilde{B}_1 \,|\, \{x_{t_0},x_s\}\ V\mathcal{F}^{y}_{t_0,s})\tilde{\wp}(x_t \,|\, \{x_{t_0},x_s\}\ V\mathcal{F}^{y}_{t_0,s}).$$

Using Theorem 7.E.4, for any $A \in \mathcal{F}^{x,y}_{t_0,s}$

$$\tilde{\wp}\ (A \,|\, \mathcal{F}^{y}_{t_0,s}\ V\{x_{t_0},x_s\}\ V\{x_t\}) = \tilde{\wp}\ (A \,|\, \mathcal{F}^{y}_{t_0,s}\ V\{x_{t_0},\ x_s\})\ .$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Therefore,

$$\tilde{E}(\Lambda(x_{t_0,t}, y_{t_0,t}) \mid \mathcal{F}^y_t, x_t, x_{t_0}) = 1 + \int_{t_0}^{t} \tilde{E}\{h^T(s,x)\, \tilde{E}(\Lambda(x_{t_0,s}, y_{t_0,s}) \mid \mathcal{F}^y_{t_0,s}, x_{t_0}, x_s) \mid \mathcal{F}^y_{t_0,s}, x_{t_0}, x_t\} dy_s.$$

Thus, by the independence of x., y. (and using Fubini's Theorem, see Kunita [96]) we arrive at

$$\tilde{\Lambda}(z, x, y_{t_0,t}) = 1 + \int_{t_0}^{t} [\int_{R^n} h(s,\xi)\tilde{\Lambda}(\xi, x, y_{t_0,s}) P(d\xi, s; z, x, t, t_0)]^T dy_s. \tag{2.3.8}$$

Next, we define $r(z,t;x,t_0) \triangleq \tilde{\Lambda}(z,x,y_{0,t})\ p(z,t;x,t_0)$ and by (2.3.8) we have

$$r(z, t; x, t_0) = p(z, t; x, t_0) + \int_0^t [\int_{R^n} h(s, \xi)\ \tilde{\Lambda}(\xi, x, y_{t_0,s}) \tag{2.3.9}$$

$$\times\, p(\xi, s; z, x, t, t_0)\ p(z, t; x, t_0) d\xi]^T\ dy_s.$$

Using the Markov property of x_t,

$$p(\xi, s; z, t, x, t_0)\ p(z, t; x, t_0) = \frac{p(z, \xi, x, t, s, t_0)}{p(z, t, x, t_0)}\ p(z, t; x, t_0)$$

$$= p(z, t; \xi, s, x, t_0)\ p(\xi, s; x, t_0) = p(z, t; \xi, s,)\ p(\xi, s; x, t_0)$$

substituting into (2.3.9), and using the definition of $r(\xi, s; x, t_0)$, we get

$$r(z, t; x, t_0) = p(z, t; x, t_0) + \int_{t_0}^{t} [\int_{R^n} h(s,\xi)\ r(\xi, s; x, t_0)\ p(z, t; \xi, s) d\xi]^T dy_s. \tag{2.3.10}$$

The integral equation (2.3.10) is similar to the integral equation given in Wong [131] and Zakai [134]. Therefore, (2.3.7) follows by expanding (2.3.10) in the same

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

fashion described by Wong [131] followed by an integration by parts. The second part of Theorem 2.3.1 follows directly from (2.3.4) by setting $s = t_0$. Notice also that the integral equation (2.3.10) is similar to the integral equation obtained by Gihman and Skorohod [66, Chp. 1, pp. 69] for multiplicative functions. QED

2.3.1 Robust Version of DMZ Equation

Next, we shall consider the robust version q of the unnormalized density ρ defined as

$$q(z,t) \triangleq \rho(z,t)\, e^{-h(t,z)\, y_t}. \tag{2.3.11}$$

It was first shown by Clark in [33] and later by Pardoux [113] that q(z,t) satisfies the stochastic differential equation

$$\frac{\partial}{\partial t} q(z,t) = \check{L}(t)^* q(z,t) + e(t,z) q(z,t), \quad 0 \leq t \leq T \tag{2.3.12}$$

$$q(z,0) = p(z,0) \triangleq p_0(z)$$

where

$$\check{L}_t(\cdot) \triangleq L_t(\cdot) - y_t^T \nabla h(t,z) a(t,z) \frac{\partial}{\partial z}(\cdot)$$

$$e(t,z) \triangleq \frac{1}{2} y_t^T \nabla h(t,z) a(t,z) (\nabla h(t,z))^T y_t$$

$$- y_t^T \left(\frac{\partial h(t,z)}{\partial t} + L_t h(t,z)\right) - \frac{1}{2} |h(t,z)|^2$$

$$a(t,z) \triangleq \sigma(t,z)\, \sigma(t,z)^T.$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Remark 2.3.2 The robust version of the DMZ equation exists if the multiplications by h^i and h^j of the stochastic part of the DMZ equation commute. When there is correlation between measurement noise and state noise, this commutation property is violated, so no such robust representation exists.

The following proposition establishes the relation between (2.3.12) and its fundamental solution.

Proposition 2.3.1 Suppose a fundamental solution $\Gamma(z, t; x, s)$ for the forward partial differential equation (2.3.12) exists. Then $\Gamma(z, t; x, s)$ as a function of (z,t) in $R^n x[0,t]$, $0 \leq s \leq T$ satisfies

$$\begin{aligned} &\frac{\partial}{\partial t}\Gamma(z,t;x,s) = \check{L}(t)^* \, \Gamma(z,t;x,s) + e(t,z) \, \Gamma(z,t;x,s) \\ &\lim_{t \downarrow s} \Gamma(z,t;x,s) = \delta(z-x) \end{aligned} \tag{2.3.13}$$

where the solution to (2.3.12) is given by

$$q(z, t) = \int_{R^n} \Gamma(z, t; x, s) \, e^{-h(x,s)y_s} p_s(x) \, dx. \tag{2.3.14}$$

The relation between $r(z, t; x, s)$ and $\Gamma(z, t; x, s)$ is given by

$$r(z, t; x, s) = \exp \{h(t, z) \, y_t - h(s, x) y_s\} \Gamma(z, t; x, s). \tag{2.3.15}$$

One can easily show by direct substitution into (2.3.12) that (2.3.14) satisfies (2.3.12). Moreover, using the same approach it can also be shown that (2.3.15) satisfies (2.3.7).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Remark 2.3.3 The fundamental solution $r(z, t; x, s)$ or its robust version $\Gamma(z, t; x, s)$ is a version of the Kallinupur-Striebel formula considered by Beneš [4] and Beneš and Karatzas [6] to obtain the fundamental solution of the DMZ equation using a probabilistic approach. In their work [4,6], the function space integration

$$\tilde{E}\{I_{\{x_t \in dz\}} \Lambda_t | \mathcal{F}^y_{t_0,t}\}$$

is considered. This expectation, however, is exactly equal to the fundamental solution of the DMZ equation, namely the function $r(z, t; x, s)dz$.

2.3.2 Backward Stochastic Partial Differential Equation

Here we represent the solution to the backward stochastic differential equation given by Pardoux [113] in terms of its fundamental solution $r^*(z,t; x,T)$. We will also relate the solution to the adjoint of (2.3.13) in terms of the fundamental solution $r^*(z,t; x,T)$.

Again, referring to Pardoux [113], the adjoint equation to the DMZ equation is the following backward stochastic PDE:

$$\begin{aligned} &dV(z, t) + L(t)V(z, t)dt + h(t, z)V(z, t)\hat{d}y_t = 0 , \quad 0 \le t \le T, \\ &V(z, T) = \Phi(z) \end{aligned} \tag{2.3.16}$$

where $V(z, t) \triangleq \tilde{E}\{\Phi(x_T)\Lambda_{t,T} | \mathcal{F}^y_{t,T}, x_t\}$. If <.,.> denotes the L^2 norm, then

$$\langle \rho(z,T),\Phi(z)\rangle=\langle \rho(z,t),V(z,t)\rangle=\langle \rho(z,s),V(z,s)\rangle= \int_{R^n} p_{t_0}(z)\tilde{E}\{\Phi(x_T)\Lambda_{t_0,T} | \mathcal{F}^y_{t_0,T}, x_{t_0}=z\}dz \tag{2.3.17}$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Remark 2.3.4 The solution to (2.3.16) will be an $\mathcal{F}^y_T$-adapted process. The stochastic integral has to be considered as a backward Ito integral. If we define $\tilde{y}_t \triangleq y_t - y_T$, then $d\tilde{y}_t = dy_t$. Since $\tilde{y}_t$ is a backward $(\mathcal{F}^y_{t,T}, \tilde{\mathcal{P}})$ Wiener process, for all $t_1 < t_2 \le T$, $\tilde{y}_{t_1} - \tilde{y}_{t_2}$ is a Gaussian random variable with mean zero, and covariance $(t_2 - t_1)I$, independent of $\mathcal{F}^y_{t_2,T}$.

Definition 2.3.2 The fundamental solution of (2.3.16) in $R^n x[0,T]$ is a function $r^*(z, t; x, T)$ defined for all (z, t) and (x, T) in $R^n x[0,T]$, $t<T$, which satisfies for any continuous $\Phi(z)$

$$V(z, t) = \int_{R^n} r^*(z, t; \zeta, T)\Phi(\zeta)d\zeta. \tag{2.3.18}$$

Next, consider the adjoint equation to (2.3.12), given by

$$\begin{aligned} &\frac{\partial}{\partial t}u(z, t) + \check{L}u(z, t) + e(t, z)\, u(z, t) = 0, \quad 0 \le t \le T \\ &u(z, T) = \Phi(z)\, e^{y_T\, h(T, z)}. \end{aligned} \tag{2.3.19}$$

Again, $u(z,t)$ and $V(z,t)$ are related by

$$u(z, t) = V(z, t) \exp\{y_t\, h(t, z)\}. \tag{2.3.20}$$

Indeed, $u(z, t)$ is the adjoint of $q(z, t)$ since, when u,q have compact support in a domain G,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$\iint_G (V\check{L}^* u - u\check{L}V) dz dt = 0.$$

Alternatively, (for homogeneous equations) we say that u(z, t) is the adjoint of q(z, t) if

$$\frac{d}{dt}\int_{R^n} q(z, t) u(z, t) dz = 0.$$

In addition, carrying out this differentiation, we can easily derive (2.3.19).

Proposition 2.3.2 Suppose the fundamental solution of (2.3.16) and (2.3.19) exist. Then r*(z, t; ζ, T) and Γ*(z, t; ζ, T) as functions of (z, t) in $R^n x[0,T]$, $0 \le t \le T$ satisfy

$$\begin{aligned} &\frac{d}{dt} r^*(z, t; \zeta, T) + L(t) r^*(z, t; \zeta, T) + h(t, z) r^*(z, t; \zeta, T) \hat{d}y_t = 0 \\ &\lim_{t \uparrow T} r^*(z, t; \zeta, T) = \delta(\zeta - z) \end{aligned} \tag{2.3.21}$$

$$\begin{aligned} &\frac{\partial}{\partial t} \Gamma^*(z, t; \zeta, T) + \check{L}(t)\, \Gamma^*(z, t; \zeta, T) + e(t, z)\, \Gamma^*(z, t; \zeta\ T) = 0 \\ &\lim_{t \uparrow T} \Gamma^*(z, t; \zeta, T) = \delta(\zeta - z) \end{aligned} \tag{2.3.22}$$

$$V(z, t) = \int_{R^n} r^*(z, t; \zeta, T)\, \Phi(\zeta)\, d\zeta \tag{2.3.23}$$

$$u(z, t) = \int_{R^n} \Gamma^*(z, t; \zeta, T)\, \Phi(\zeta)\, e^{y_T h(T, \zeta)} d\zeta. \tag{2.3.24}$$

Proof: Substituting (2.3.23), (2.3.24) into (2.3.16), (2.3.19), respectively, and using (2.3.21), (2.3.22), equations (2.3.16), (2.3.19) are satisfied. QED

Remark 2.3.5 Equation (2.3.19) can be obtained by applying the backward differential rule to (2.3.20) as given in Theorem 7.B.3. One could obtain the same result

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

if (2.3.16) is expressed in the Fisk-Stratonovich form through (7.B.1). Finally the relation between $\Gamma^*(z, t; \zeta, T)$ and $r^*(z, t; \zeta, T)$ is given by

$$\Gamma^*(z, t; \zeta, T) = e^{y_t\, h(t, z) - y_T\, h(T, \zeta)}\, r^*(z, t; \zeta, T). \tag{2.3.25}$$

The following proposition is a consequence of analogous results for parabolic partial differential equations as presented by Friedman [60,Thm.15,pp.28-29].

Proposition 2.3.3 Suppose the fundamental solutions $r^*(z, t; \zeta, T)$ and $\Gamma^*(z, t; \zeta, T)$ of (2.3.16), (2.3.19), respectively, exist. Then they are related to the fundamental solutions $r(z, t; x, s)$ and $\Gamma(z, t; x, s)$ of the stochastic and PDE's of (2.3.7), (2.3.13), respectively, by

$$r(z, t; \zeta, \tau) = r^*(\zeta, \tau; z, t) \quad , t > \tau \tag{2.3.26a}$$

$$r(\zeta, \tau; z, t) = r^*(z, t; \zeta, \tau) \quad , \tau > t. \tag{2.3.26b}$$

with the above equalities satisfied if r, r^* are replaced by Γ, Γ^* respectively. Moreover, the solutions $V(z, t)$, $u(z, t)$ can be expressed as

$$V(z, t) = \int_{R^n} r(\zeta, T; z, t)\, \Phi(\zeta)\, d\zeta \tag{2.3.27a}$$

$$u(z, t) = \int_{R^n} \Gamma(\zeta, T; z, t)\, e^{y_T\, h(T, \zeta)}\, \Phi(\zeta) d\zeta. \tag{2.3.27b}$$

where $r(z, t; x, s)$, $\Gamma(z, t; x, s)$ are the fundamental solutions to (2.3.7), (2.3.13), respectively.

Proof: The proof of (2.3.26a), (2.3.26b) is an easy extension of the proof found in Friedman [60, Thm. 15, pp. 28-29]. Once we accept the relation (2.3.26a), then

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

(2.3.18) can be represented as given by (2.3.27a). To show (2.3.27b), we start with (2.3.18) and substitute for $r^*(z, t; \zeta, T)$ using (2.3.25); thus,

$$V(z, t) = \int_{R^n} \Gamma^*(z, t; \zeta, T) \exp\{- y_t\, h(t, z) + y_T\, h(T, \zeta)\}\, \Phi(\zeta)\, d\zeta .$$

By the relation (2.3.20) it follows that

$$u(z, t) = \int_{R^n} \Gamma^*(z, t; \zeta, T)\, e^{y_T h(T, \zeta)}\, \Phi(\zeta)\, d\zeta$$

and (2.3.27b) is a consequence of (2.3.26b). QED

Remark 2.3.6 What we have shown up to this point is that if we start with two stochastic differential equations for $\rho(z, t)$, $V(z, t)$, the first evolving forward in time and the second evolving backward in time, we can express their solutions in terms of the fundamental solutions of $\rho(z, t)$, $V(z, t)$, respectively. The same also holds for robust versions $q(z, t)$, $u(z, t)$, respectively. Moreover, if $\rho(z, t)$, $V(z,t)$ are adjoints of each other, in the sense defined earlier, we can always express the solutions $\rho(z,t)$, $V(z,t)$ in terms of the fundamental solutions of $V(z, t)$, $\rho(z, t)$ respectively. Finally, we conclude that stochastic PDE's have similar representations as nonstochastic PDE's.

Next, we shall give some results on the existence of the solution to (2.3.12), (2.3.19) which are characterized by Friedman [60] as Cauchy problems. The following proposition follows by the result of Friedman [61, Chp. 6, pp. 139-144].

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Proposition 2.3.4 Suppose assumptions (A2), (A3), (A4) are valid. Then there exist at most one solution q(z,t) and u(z,t) of the Cauchy problems (2.3.12), (2.3.19), respectively, satisfying

$$|q(z,\, t)| \le B\, e^{\beta |z|^2},\quad |u(z,\, t)| \le B\, e^{\beta |z|^2}$$

where B, β are some positive constants. In our case both solutions q(z, t) and u(z,t) are pathwise dependent on the trajectory of y, which may regarded as a parameter.

If, however, we assume that (2.3.12), (2.3.19) are parabolic and (A1), (A2), (A3) are satisfied, then there exist fundamental solutions $\Gamma(z,\, t;\, x,\, s)$ and $\Gamma^*(z,\, t;\, \zeta,\, T)$, respectively, satisfying the inequalities

$$\left|\frac{\partial^m}{\partial z^m}\, \Gamma(z,\, t;\, x,\, s)\right| \le C_1(t-s)\, e^{-(n+|m|)/2}\, e^{-C_2\frac{|z-x|^2}{t-s}}$$

$$\left|\frac{\partial^m}{\partial x^m}\Gamma^*(x,\, t;\, \zeta,\, T)\right| \le C'_1(T-t)\, e^{-(n+|m|)/2}\, e^{-C'_2\frac{|x-z|^2}{T-t}},$$

for $|m| = 0, 1$, where C'_1, C'_2, C_1, C_2 are positive constants. Based on the above results the existence and uniqueness of the DMZ equation follows.

Remark 2.3.7 It should be noted that both ρ,V (also, r^*, r) are measure-valued semimartingales, as defined in Appendix 7.B.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

2.4 PARTIALLY OBSERVED STOCHASTIC CONTROL AND ITS REPRESENTATIONS

Here we consider the stochastic control problem described by (2.2.1) - (2.2.4), and attack it by adapting the dynamic programming approach used for complete observation (see, e.g. Elliott [47]) to our problem of partial observation. First, we wish to convert the partially observed stochastic control problem to one which is completely observed and where the DMZ equation is regarded as the new state equation.

If $u \in U$, then the total expected cost is given by (2.2.2). However, if the dynamic programming approach is used and control $v \in U_{0,s}$ is used on the interval (0,s) and control $\tilde{u} \in U_{s,T}$ is used during the interval (s,T], then the expected cost at time s, given the observations up to time s, is expressed as

$$\tilde{J}(s, u) = E^u \{ \int_0^T \text{Ж}(t, x_t^u, u_t) dt + \kappa(x_T^u) \,|\, \mathcal{F}_s^y \}. \tag{2.4.1}$$

Using the smoothing property of conditional expectation,

$$\begin{aligned} \tilde{J}(s, u) &= E^u \{ E^u [\int_0^T \text{Ж}(t, x_t^u, u_t) dt + \kappa(x_T^u) \,|\, \mathcal{F}_T^y] \,|\, \mathcal{F}_s^y \} \\ &= E^u \{ \int_0^T E^u [\text{Ж}(t, x_t^u, u_t) \,|\, \mathcal{F}_T^y] dt + E^u [\kappa(x_T^u) \,|\, \mathcal{F}_T^y] \,|\, \mathcal{F}_s^y \}. \end{aligned}$$

If we define a new measure as done earlier, then by using the concept of conditional independence as is done in the proof of Theorem 2.3.1, we obtain

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$\tilde{J}(s,u) = E^u\{\int_0^T \frac{\langle \Pi(t, x_t^u, u_t), \rho(x,t)\rangle}{\langle 1, \rho(x,t)\rangle}dt + \frac{\langle \kappa(x_T^u), \rho(x,T)\rangle}{\langle 1, \rho(x,T)\rangle} \,|\, \mathcal{F}_s^y\}.$$

Again, applying (7.E.5),

$$\tilde{J}(s,u) = \frac{\tilde{E}\{\Lambda_T^u\,[\int_0^T \frac{\langle \Pi_t, \rho_t\rangle}{\langle 1,\rho_t\rangle}dt + \frac{\langle \kappa_T, \rho_T\rangle}{\langle 1,\rho_T\rangle}] \,|\, \mathcal{F}_s^y\}}{\tilde{E}[\Lambda_T^u \,|\, \mathcal{F}_s^y]}.$$

By the smoothing property of conditional expectation,

$$\tilde{J}(s,u) = \frac{\tilde{E}\{\tilde{E}[\int_0^T \Lambda_t^u \frac{\langle \Pi_t, \rho_t\rangle}{\langle 1,\rho_t\rangle}dt + \Lambda_T^u \frac{\langle \kappa_T, \rho_T\rangle}{\langle 1, \rho_t\rangle} \,|\, \mathcal{F}_T^y] \,|\, \mathcal{F}_s^y\}}{\tilde{E}\{\tilde{E}[\Lambda_{0,s}^v\, \Lambda_{s,T}^{\bar{u}} \,|\, \mathcal{F}_s^y]\}}.$$

Thus,

$$\tilde{J}(s,u) = \frac{\tilde{E}\{\int_0^T \frac{\langle 1,\rho_t\rangle\langle \Pi_t, \rho_t\rangle}{\langle 1, \rho_t\rangle}dt + \langle 1,\rho_T\rangle \frac{\langle \kappa_T, \rho_T\rangle}{\langle 1,\rho_T\rangle} \,|\, \mathcal{F}_s^y\}}{\tilde{E}[\Lambda_{0,s}^v \,|\, \mathcal{F}_s^y]}$$

$$= \frac{\tilde{E}\{\int_0^T \langle \Pi_t, \rho_t\rangle dt + \langle \kappa_T, \rho_T\rangle \,|\, \mathcal{F}_s^y\}}{\langle 1, \rho_s\rangle}.$$

Since we are interested in minimizing controls used during the interval (s,t), then, as in dynamic programming, we need not be concerned with controls $v \in U_{0,s}$. Therefore, by an abuse of notation (i.e., without introducing a new symbol for $\tilde{J}(s,u)$), we have

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$\tilde{J}(s, u) = \tilde{E}\{\int_s^T \langle \Pi(t, x_t^{\tilde{u}}, \tilde{u}_t), \rho(x, t)\rangle dt + \langle \kappa(x_T^{\tilde{u}}), \rho(x, T)\rangle \,|\, \mathcal{F}_s^y\}. \tag{2.4.2}$$

This is in fact the cost functional we shall minimize over all $\tilde{u} \in U_{s,T}$.

If we define

$$\varphi_{s,T} \triangleq \int_s^T \langle \Pi(t, x_t^{\tilde{u}}, \tilde{u}), \rho(x, t)\rangle dt + \langle \kappa(x_T^{\tilde{u}}), \rho(x, T)\rangle, \tag{2.4.3}$$

then

$$\begin{aligned} \varphi_{s,T} &= \tilde{E}\{\kappa(x_{s,T}^{\tilde{u}})\, \Lambda_{s,T}^{\tilde{u}} + \int_s^T \Lambda_{s,t}^{\tilde{u}}\, \Pi(t, x_{s,t}^{\tilde{u}}, \tilde{u}_t) \,|\, \mathcal{F}_{s,T}^y\} \\ &= \tilde{E}\{\tilde{E}[\kappa(x_{s,T}^{\tilde{u}})\, \Lambda_{s,T}^{\tilde{u}} + \int_s^T \Lambda_{s,t}^{\tilde{u}}(t, x_{s,t}^{u}, \tilde{u}_t) dt \,|\, \mathcal{F}_{s,T}^y, x_s] \,|\, \mathcal{F}_{s,T}^y\} \\ &= \langle \bar{V}(z, s), \rho(z, s)\rangle = \int_{R^n} \bar{V}(z, s)\, \rho(z, s) dz \end{aligned} \tag{2.4.4}$$

where

$$\bar{V}(z, s) \triangleq \tilde{E}\{\kappa(x_{s,T}^{\tilde{u}}\, \Lambda_{s,T}^{\tilde{u}} + \int_s^T \Lambda_{s,T}^{\tilde{u}}\, \Pi(t, x_{s,t}^{\tilde{u}}, \tilde{u}_t) dt \,|\, \mathcal{F}_{s,T}^y, x_s = z\}. \tag{2.4.5}$$

The last equality of (2.4.4) is a consequence of (2.3.17).

<u>Theorem 2.4.1</u>

The expectation (2.4.5) satisfies the stochastic backward PDE

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$
\begin{aligned}
&d\bar{V}(z, s) + L(t)\bar{V}(z, s)ds + h(s, z)\hat{d}y_s + \text{π}(s, x, \bar{u}) = 0,\ 0 \le s \le T \\
&\lim_{s\uparrow T}\bar{V}(z, s) = \kappa(z).
\end{aligned} \qquad (2.4.6)
$$

Moreover, if the fundamental solution $r^*(z, s; \zeta, T)$ of (2.4.6) exists in $R^n x[0,T]$, defined for all (z, s, ζ, T) in $R^n x[0,T] x R^n x[0,T]$, $s < T$, then it must satisfy (2.3.21).

Using proposition (2.3.3), $\bar{V}(z, s)$ can be written as

$$
\bar{V}(z, s) = \int_{R^n} r(\zeta, T; z, s)\, \kappa(\zeta)\, d\zeta + \int_s^T \int_{R^n} r(\zeta, t; z, s)\, \text{π}(\zeta, t, \bar{u}_t) d\zeta dt. \quad (2.4.7)
$$

Proof: First, consider the proof of (2.4.5). Since $\mathcal{F}^{x,y}_{s,t}$ and $\mathcal{F}^{y}_{t,T}$ are conditionally independent given $\mathcal{F}^{y}_{s,t}\ V\{x_s\}$, then Theorem 7.E.4, (2.4.5) can be written as

$$
\begin{aligned}
\bar{V}(z, s) &= \tilde{E}\{\kappa(x^{\bar{u}}_{s,T})\, \Lambda^{\bar{u}}_{s,T} \mid \mathcal{F}^{y}_{s,T} x_s = z\} + \tilde{E}\{\int_s^T \Lambda^{\bar{u}}_{s,t} \text{π}(t, x^{\bar{u}}_{s,t}, \bar{u}_t) dt \mid \mathcal{F}^{y}_{s,t}\ V\{x_s\}\} \\
&= \tilde{E}\ \{\kappa(x^{\bar{u}}_{s,T})\, \Lambda^{\bar{u}}_{s,T} + \int_s^T \Lambda^{\bar{u}}_{s,t} \text{π}(t, x^{\bar{u}}_{s,t}, u_t) dt \mid x_s = z\}.
\end{aligned}
$$

The last equality can be shown as follows: Define $B \triangleq \{x_s\}$, $B_1 \triangleq \mathcal{F}^{y}_{s,t}$, $B_2 = \mathcal{F}^{x,y}_{s,t}$. Then,

$$
\begin{aligned}
\rho(B_2 \mid B, B_1) &= \frac{\rho(B, B_1, B_2)}{\rho(B, B_1)} \\
&= \frac{\rho(B_1, B_2 \mid B)\, \rho(B)}{\rho(B_1 \mid B)\rho(B)} \\
&= \frac{\rho(B_2 \mid B)}{\rho(B_1 \mid B)} = \frac{\rho(B_2 \mid B)}{\rho(B_1)} = \rho(B_2 \mid B)
\end{aligned}
$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

where last equality follows since $\mu_y^d(w)$ is a Wiener measure.

Applying the Ito differential rule to $\bar{V}(z, s)$ as presented by Krylov [89,Chp. 1] we deduce (2.4.6) which is also given in Bensoussan [10] and Haussmann [68]. Substituting (2.4.7) into (2.4.6), one can show that (2.4.7) is the correct representation of the solution $\bar{V}(z, s)$ by using (2.3.21) which also follows by Freidman [61,Chp. 6, pp. 139-144]. QED

The robust version of (2.4.6) is obtained through the gauge transformation

$$\bar{u}(z, s) = \bar{V}(z, s)\, e^{y_t h(s, z)}.$$

Proposition 2.4.1 The random function $\bar{u}(z,s)$ satisfies the partial differential equation

$$\frac{\partial}{\partial s}\bar{u}(z,s) + \check{L}_s\bar{u}(z,s) + e(s,z)\bar{u}(z,s) + \text{II}(s, z, \bar{u})e^{y_s h(s,z)} = 0, \quad 0 \le s \le T \tag{2.4.8}$$
$$\bar{u}(z,T) = \kappa(z)e^{y_T h(T,z)}.$$

Furthermore, if a fundamental solution $\Gamma^*(z, t; \zeta, T)$ of (2.4.8) exists in $R^n x[0,T]$, defined for all (z, s, ζ, T) in $R^n x[0,T]xR^n x[0,T]$, $s < T$, it satisfies (2.3.22),

$$\frac{\partial}{\partial s}\Gamma^*(z, s; \zeta, T) + \check{L}(s)\Gamma^*(z, s; \zeta, T) + e(s, z)\,\Gamma^*(z, s; \zeta, T) = 0, \quad 0 \le s \le T \tag{2.4.9}$$
$$\lim_{s \uparrow T} \Gamma^*(z, s, \zeta, T) = \delta(\zeta - z).$$

The solution $\bar{u}(z,s)$ can be represented as

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$\bar{u}(z,s) = \int_{R^n} \kappa(\zeta)\Gamma(\zeta,T;z,s)e^{y_T h(T,\zeta)}d\zeta + \int_s^T \int_{R^n} \Gamma(\zeta,t;z,s)e^{y_t h(t,\zeta)}\,\Pi(t,\zeta,\bar{u})d\zeta dt. \tag{2.4.10}$$

Proof: It follows as in Theorem 2.4.1. QED

Theorem 2.4.2 The cost function (2.4.2) can be expressed as

$$\begin{aligned}\tilde{J}(s, u) &= \tilde{E}\{\langle \bar{V}(z, s), \rho(z, s)\rangle \mid \mathcal{F}_s^y\} \\ &= \tilde{E}\{ \int_{R^n} \langle r(\zeta, T; z, s), \rho(z, s)\rangle\, \kappa(\zeta)d\zeta \\ &+ \int_s^T \int_{R^n} \langle r(\zeta, t; z, s), \rho(z, s)\rangle\, \Pi(\zeta, t, u)d\zeta dt \mid \mathcal{F}_s^y\}\end{aligned}$$

where the expectation $\tilde{E}$ is with respect to the Wiener measure $\mu_y^d\,(dy)$.

Proof: The first equality is just (2.4.4) and the second equality follows from (2.4.7). QED

2.5 DERIVATION OF THE DMZ EQUATION WITH TERMINATION

Suppose the state and observation processes satisfy (2.2.3), (2.2.4), respectively, where in this case the Markov process $\{x_t, 0 \le t \le T\}$ terminates at some time $\tau > s$ w.p. 1 and is interpreted as the first exit time of the state process x_t from some open bounded set D with C^2-boundary ∂D. For this version of the stopping problem the only available information is given through the noisy measurements $\{y_t, 0 \le t \le T\}$. Our problem is to derive an evolution-type stochastic PDE that describes the behavior of the unnormalized conditional density.

Thoughout the remainder of this section we shall make the following assumptions.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

(a) The coefficient f, σ, h of the stochastic differential equations (2.2.3), (2.2.4) are bounded, twice continuously differentiable having bounded first derivatives.

(b) The initial density satisfies assumption (A5) given earlier.

(c) The function $\Phi(x)$ is bounded with compact support in (0,T)xD.

We start by noting that the time $\tau(\omega)$, $\omega \in \Omega$ must be an $\mathscr{F}_t^y$-adapted process. Therefore, by the definition of stopping times given in Appendix 7.B, it is a stopping time with respect to family of σ-fields $\{\mathscr{F}_t^y, t \in [0, T]\}$ since the event $\{\tau \leq t\} \in \mathscr{F}_t^y$. Moreover, by $\mathscr{F}_t^y \subset \mathscr{F}_t$, it follows that τ is also a stopping time with respect to the family $\{\mathscr{F}_t, t \in [0, T]\}$. Using the exponential formula of Theorem 7.E.1 we define

$$\Lambda_\tau \triangleq \mathscr{E}\left(\int_0^\tau h^T(s, x_s) dy_s\right)$$

with

$$\Lambda_\tau \triangleq \tilde{E}\left[\frac{dP_{\tau(y)}}{d\tilde{P}_{\tau(y)}} \middle| \mathscr{F}_\tau(y)\right]$$

where $P_{\tau(y)}$, $\tilde{P}_{\tau(y)}$ denote the restriction of measures P, $\tilde{P}$ (i.e., the measures introduced earlier for the case $\tau > t$) on the σ-algebra $\mathscr{F}_\tau$. Due to the bounded assumption on $h(t, x_t)$, it follows that $P_{\tau(y)}\{\int_0^{\tau(y)} |h(t, x_t)|^2 dt < \infty\} = 1$, a.s.; therefore, by Girsanov's Theorem

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

(Theorem 7.E.2) $\tilde{E}[\Lambda_\tau] = 1$ (that is $\rho_{\tau(y)} << \tilde{\rho}_{\tau(y)}$) and by Remark 7.E.3 $\tilde{\rho}_{\tau(y)} << \rho_{\tau(y)}$, so $\rho_{\tau(y)} \sim \tilde{\rho}_{\tau(y)}$. At this point we like to follow the approach presented by Pardoux [113,114] to first derive a backward PDE the adjoint of the DMZ equation. Thus, we formulate the filtering problem by considering

$$E[\Phi(x_\tau)|\mathcal{F}^y_\tau] = \frac{\tilde{E}[\Phi(x_\tau)\ \Lambda_\tau|\mathcal{F}^y_\tau]}{\tilde{E}[\Lambda_\tau|\mathcal{F}^y_\tau]}$$

where we drop the dependence of τ on y by keeping in mind that under measure $\tilde{\rho}_\tau$ the stopping time τ is a function of y, a standard Brownian motion process. By the proof of Theorem 2.4.1 we are left to consider the quantity:

$$V(x,s) = \tilde{E}[\Phi(x_\tau)\ \Lambda_{s,\tau}\ |\ \mathcal{F}^y_{s,\tau}, x_s] \tag{2.5.1}$$

where, $\tau > s$ a.s..

Next, defining $u(x,s) = V(x,s)\, e^{y_s h(s,\, x_s)}$ as in (2.3.20), we are required to determine a partial differential equation satisfied by

$$u(x,s) = \tilde{E}\{\Phi(x_\tau)\, e^{y_\tau h(\tau,x_\tau)}\, e^{\int_s^\tau e(\sigma,\, x_\sigma,\, y_\sigma)\, d\sigma} |x_s = x\}. \tag{2.5.2}$$

Notice also that the drift coefficient of (2.2.3) is now replaced by $f(t,x_t) - a(t, x_t)(\nabla h(t,x_t))^T y_t)$.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The approach to be followed is based on the work of Dynkin [46, Chps. 9, 10], Prohorov and Rozanov [116, pp. 277-282], Gihman and Skorohod [66, pp. 63-75], Hibey [74, pp. 39-44] where the above authors treated the same problem when the state process is completely observed and the second exponential term of (2.5.2) is a standard multiplicative functional of the state process x_t, that is, the functional of x_t defined by

$$\mu_{s,t} \triangleq e^{-\int_s^t e(x_u)du}$$

satisfying $\mu_{s,\tau}\, \mu_{\tau,t} = \mu_{s,t}$ for $s \le \tau \le t$ and $0 \le \mu_{s,t} \le 1$. In our case $\mu_{s,t}$ is not standard in that it is not homogeneous and is not bounded above by one because u(x,s) is an unnormalized conditional expectation. Dynkin [46, pp. 283] calls such a function a quasi-transition function.

Following the same approach as in Hibey [74, pp. 39], suppose we are given function $\xi(\omega)$, $\omega \in \Omega$ taking values in the interval $[0, \infty]$. The Markov process $x_t(\omega)$ defined on the space Ω, where $t \in [0, \xi(\omega)]$ for each $\omega \in \Omega$, takes values in the measurable space (R^n, B^n), the state space having a transition function P(x, t; x, s). Suppose a new process $\bar{x}_t(\omega)$ is obtained by terminating the process x_t at some random time $\tau(\omega) \le \xi(\omega)$, where $\tau(\omega)$ is an $\mathcal{F}_t^y$-adapted stopping time. Then the time set $t \in [0, \xi]$ on which the original process was defined can be replaced by the time set $t \in [0, \tau(y)]$.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Assume there exists a constant c such that $-\bar{e} \triangleq e(t, x_t, y_t) + c \leq 0$; then (2.5.2) can be expressed as

$$u(x, s) = \tilde{E}\left[\Phi(x_\tau)\, e^{y_\tau h(\tau, y_\tau)}\, e^{-\int_s^\tau \bar{e}^{\,y}(u, x_u)\, du} \,|\, x_s = x\right] e^{-c(\tau - s)}.$$

where the $\mathscr{F}_\tau^y$-adapted process $\tau(y)$ can be removed from the expectation via (2.5.1).

Now, let each trajectory of the original process $x_t(\omega)$, when taking the value $x_t(\omega) = x \in (R^n, B^n)$ terminate during the following period of time Δt with probability $\bar{e}^{\,y}(t, x)\, \Delta t + 0(\Delta t)$, $\bar{e}^{\,y} \geq 0$ which is pathwise dependent on the trajectory of the process $y_t(\omega)$. This density is called the termination density of the new process. The new process $\bar{x}_t(\omega)$ obtained in this way will terminate at some random time $\tau(y)$ where $0 \leq \tau(y) \leq \xi(\omega)$ for all $\omega \in \Omega$. The function $\xi(\omega)$ is called the lifetime of the original process $x_t(\omega)$, or the terminal time. If $\xi(\omega) = \infty$ then we say that $x_t(\omega)$ is non-terminating for all $\omega \in \Omega$. The terminated Markov process $\bar{x}_t(\omega)$ will have the same drift and diffusion coefficients as the original process x_t. Denote the completion of B^n by $\bar{B}^n$. Then, for $\omega \in \Omega$, if we define $\bar{\xi}(y) \triangleq \tau(y)$, $\bar{x}_t(\omega) \triangleq x_t(\omega)$ for $0 \leq t \leq \tau(y)$ and $\bar{\wp}_{\tau,x} \sim \tilde{\wp}_{\tau,x}$ where subscript x denotes conditional probability measure on x_s, then x_t is also a Markov process in the state space $(R^n, \bar{B}^n)$ of lifetime $\bar{\xi}(y)$ with measure $\bar{\wp}_{\tau,x}$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

(see Dynkin [46, pp. 301]). That is, if we denote by ξ, $\bar{\xi}$ the duration of the processes x_t, $\bar{x}_t$, respectively, then $\bar{x}_t$ can be obtained by the truncation of the duration of the process x_t.

At this point we adapt the methodology presented by Prohorov and Rozanov [116, pp. 278-281] by defining the probability of surviving until time t, given that the process moves along the paths of x_t, y_t by

$$\bar{P}(\tau > t \mid x_u, y_u, 0 \le u \le t) \triangleq e^{-\int_s^t \bar{e}^{\,y}(u,\, x_u)\, du} .$$

By the equivalence of measures $\bar{\wp}_{\tau,x} \sim \wp_{\tau,x}$ we have for any $\beta \in \bar{B}^n$

$$\begin{aligned}\bar{P}_{\tau,x}(x_t \in \beta \mid x_s = x) &= \bar{P}_{\tau,x}\{x_t \in \beta, \tau > t \mid x_s\} = \bar{E}[I_{\{\omega; x_t \in \beta\} \cap \{\omega; \tau > t\}} \mid x_s = x] \\ &= \bar{E}\,[\bar{E}(I_{\{\omega; x_t \in \beta\} \cap \{\omega; \tau > t\}} \mid x_u, y_u, s \le u \le t) \mid x_s = x)]\end{aligned}$$

where $\bar{E}$ denotes expectation with respect to measure $\bar{\wp}_{\tau,x}$.

But, under measure $\tilde{\wp}_{\tau,x}$, x. and y. are independent, so

$$\wp_{\tau,x}\,(dx, dy) = \wp_{\tau,x}\,(dx)\ \mu_y^d\ (dy)$$

where μ_y^d is a Wiener measure. Therefore,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$\bar{P}_{\tau,x}\{x_t \in \beta,\ \tau > t | x_s = x\} = \int_{\{x_t \in \beta,\ t \le \tau\}} e^{-\int_s^t \bar{e}\, y(u, x_u) dw} P_{\tau,x}(d\omega). \tag{2.5.3}$$

By the definition of u(x, s) it follows that

$$u(x, s) = e^{-c(\tau - s)} \bar{u}(x, s)$$

where

$$\bar{u}(x,s) \triangleq \bar{E}[e^{y_\tau h(\tau, x_\tau)} e^{-\int_s^\tau \bar{e}\, y(u, x_u) du} | x_s = x].$$

Moreover, the random process $\bar{u}(x, s)$ satisfies the functional relation

$$\bar{u}(x, s) = \int_D \bar{u}(z, t)\, \bar{P}_\tau(dz, t; x, s)$$

in which $\bar{P}_\tau(x_t \in \beta;\ x_s = x)$ denotes the transition function given by (2.5.3) and $\bar{u}(z,t) = e^{y_t h(t, y_t)} \Phi(z)$. The function $\bar{u}(x, s)$ satisfies the Feynman-Kac equation

$$\frac{\partial \bar{u}(x, s)}{\partial s} + \check{L}_s \bar{u}(x, s) - \bar{e}\, y(x, s)\, \bar{u}(x, s) = 0 \quad , (s, x) \in [0,t] \mathrm{x} D$$

$$\bar{u}(x, t) = \Phi(x) e^{y_t h(t, x_t)} \quad , x \in D \tag{2.5.4}$$

$$\bar{u}(x, s) = 0 \quad , (s, x) \in [0,\ t), \mathrm{x}\partial D.$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Since, $u(x,s)$, $\bar{u}(x,s)$ are related by the factor $e^{-c(\tau-s)}$ if follows from (2.5.4) that $u(x,s)$ satisfies

$$\frac{\partial u(x,s)}{\partial s} + \check{L}_s\, u(x,s) + e(x,s)\, u(x,s) = 0 \quad , (s,x) \in [0,t]xD$$

$$u(x,t) = \Phi(x)e^{y_t\, h(t,x_t)} \quad , x \in D \tag{2.5.5}$$

$$u(x,s) = 0 \quad , (x,s) \in [0,t),x\partial D.$$

The resulting boundary value problem (2.5.5) is investigated by Friedman [60, Chp. 3] when the pathwise dependence on the process $y_t(\omega)$ is eliminated. The same author classifies such problems as first initial-boundary value problems. Existence and uniqueness results on the solution to (2.5.5) are given by Friedman [60, Chp. 3].

Remark 2.5.1 Suppose a solution $u(x, s)$ to (2.5.5) exists. Then, using an earlier formula stated in this section when $\tau > t$, we have

$$E[\Phi(x_t) \mid \mathscr{F}_t^y] = \int_D u(x,0)\, p_0(x)dx.$$

Using the methodology described earlier for nonterminating diffusion processes we can derive the adjoint equation to (2.5.5) which satifies a forward PDE. The results are summarized in the following theorem.

Theorem 2.5.1

Suppose the state process x_t is defined on an open bounded domain $D \subset R^n$ having a C^2-boundary ∂D (i.e., ∂D is a manifold of class C^2). Assume that f, σ (i.e drift and diffusion coefficients of x_t process) are bounded and twice continuously differentiable

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

with respect to x and the operator $\check{L}$ is of parabolic type. Then the robust version of the DMZ equation satisfies

$$\frac{\partial q}{\partial t}(z,t) + \check{L}(t)^{*}q(z, t) + e(t, z_t)\, q(z,t) = 0 \quad ,(t,z)\in [0,t]xD$$

$$q(z,0) = p_0(z) \qquad ,x\in D \tag{2.5.6}$$

$$q(z,t) = 0 \qquad ,(z,t) \in [t,0)x\partial D,$$

where q(z,t) has a unique continuously differentiable solution q(z, t) for any smooth initial density $p_0(z)$.

Proof: The existence and uniqueness of the solution q(z, t) follows by direct application of the results presented by Friedman [60, Chp. 3, Thm. 16, pp. 82].

Remark 2.5.2 The physical interpretation of the stopping time τ associated with the state process x_t is interpreted as the hitting time of a suitable boundary by the conditional distribution q(z, t).

Remark 2.5.3 One could also derive the result of Theorem 2.5.1 by modyfying the approach considered in proving Theorem 2.3.1 as follows. First replace (2.3.1) by

$$E[\Phi(x_t)I_{\{\omega;x_t(\omega)\in D\}}|\mathcal{F}_t^y] = \frac{\tilde{E}[\Lambda_t\Phi(x_t)I_{\{\omega;x_t(\omega)\in D\}}|\mathcal{F}_t^y]}{\tilde{E}[\Lambda_t|\mathcal{F}_t^y]}$$

and use the smoothing property of conditional expectation thus,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$E[\Phi(x_t)I_{\{\omega;x_t(\omega)\in D\}}|\mathscr{F}_t^y] = \frac{\tilde{E}[\tilde{E}(\Lambda_t\Phi(x_t)I_{\{\omega;x_t(\omega)\in D\}}|\mathscr{F}_t^y, x_t, x_{t_0})|\mathscr{F}_t^y]}{\tilde{E}[\tilde{E}(\Lambda_t|\mathscr{F}_t^y, x_t, x_0)|\mathscr{F}_t^y]}$$

$$= \frac{\tilde{E}[I_{\{\omega;x_t(\omega)\in D\}}\Phi(x_t)\,\tilde{E}(\Lambda_t|\mathscr{F}_t^y, x_t, x_{t_0})|\mathscr{F}_t^y]}{\tilde{E}[\,\tilde{E}(\Lambda_t|\mathscr{F}_t^y, x_t, x_{t_0})|\mathscr{F}_t^y]}.$$

Since the domain of the infinitesimal operator of the nonterminating diffusion process $\{x_t, t \geq 0\}$ given by (2.3.3) is contained in the domain of the characteristic operator of the terminating process $\{x_t, 0 \leq t \geq \tau\}$ $(\tau \triangleq \inf\{t \geq 0;\ x_t \notin D\})$

(see Dynkin [46, pp. 143,Vol. 1]) defined as

$$L(s)f(x) = \lim_{D \downarrow x} \frac{E[f(x_\tau)|x_s = x] - f(x)}{E[\tau|x_s = x]}$$

and the above limit coincides with the limit of (7.F.5) for any C^2-functions f, it follows that for diffusion processes satisfying the strong Markov property (see, Wong and Hajek [132, pp. 19 for definition), L(s) of (7.B.5) takes the form of the second order operator given by (7.F.6).

Therefore, we have

$$E[\Phi(x_t)I_{\{\omega;x_t(\omega)\in D\}}|\mathscr{F}_t^y] = \frac{\int_D \Phi(z)\tilde{\Lambda}(z, x, y_{t_0,t})p(dz, dx, t, t_0)}{\int_D \tilde{\Lambda}(z, x, y_{t_0,t})p(dz, dx, t, t_0)}$$

where

$$P(A, x, t, t_0) = \mathrm{Prob}\{x_t \in A, x_{t_0} = x\}, \text{ for any Borel set } A \in D.$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

It is now clear that the proof of Thorem 2.3.1 could be adapted to show that

$$\rho(z,s) = \int_D \tilde{r}(z,t;\, x,s) P_s(x) dx \tag{2.5.7}$$

where $\tilde{r}(z,t;\, x,s)$ is defined by (2.3.5) and satisfies the stochastic PDE (2.3.7) on the domain $[0,T]\mathrm{x}D$ and

$$\begin{aligned} \rho(z,s) &= p_s(z) \;, \quad x \in D \\ \rho(z,t) &= 0, \quad\;\; , \; (t,x) \in [0,t)\mathrm{x}\partial D. \end{aligned} \tag{2.5.8}$$

As a consequence of (2.5.7), (2.5.8), the representations derived earlier for nonterminating processes could also be modified to cover the case of the current section.

2.5.1 An Upper bound on DMZ Equation with Termination

We shall now find an upper bound for the unnormalized conditional density $\rho(z,t)$ given by (2.5.7), (2.5.8) using the representation methods introduced in this Chapter.

Suppose $r(z,t;\, x,s)$ is the fundamental solution of the DMZ equation. Then using Friedman [62, pp. 346] the function $\tilde{r}(z,t;x,s)$ of (2.5.7), (2.5.8) is called a Green's function for (1.1.6) in the domain $(0, \infty)\mathrm{x}D$. If $\tilde{r}(z,t;\, x,s)$ is unique we can construct the function $\tilde{r}(z,t;\, x,s)$ by

$$\tilde{r}(z,t;x,s) = r(z,t;x,s) + r_1(z,t;x,s) \tag{2.5.9}$$

where $r(z,t;\, x,s)$ is the fundamental solution of the DMZ equation satifying (2.3.7) as shown by Theorem 2.3.1. Notice also that $r_1(z,t;\, x,s)$ should be the solution to

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$\begin{aligned}
&dr_1(z,t;x,s) = L(t)^* r_1(z,t;x,s)dt + h^T(t,z) r_1(z,t;x,s)dy_s &&, [0,T]xD \\
&\lim_{t \downarrow s} r_1(z,t;x,s) = 0 &&, z \in D \qquad (2.5.10) \\
&r_1(z,t;x,s) = -r(z,t;x,s) &&, (t,z) \in (0,T)x\partial D.
\end{aligned}$$

That is, the unnormalized conditional density given by (2.5.7) can always be constructed by (2.5.9), (2.5.10). An application of maximum principle (see Friedman [Chp. 6, pp. 132-133]) implies that $r_1(z,t;x,s)$ is always less than zero. Therefore,

$$\tilde{r}(z,t;x,s) \le r(z,t;x,s)$$

which provides an upper bound on u(z,t) also. Thus, if we can solve (2.3.7) we can determine the upper bound of the unnormalized conditional density of a diffusion process terminated at the first exist time from D.

We conclude this chapter by presenting the following table to assist the reader in referencing the several processes we have introduced.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table of Measure-Valued Processes and their Representation						
Process	Defining equation	Differential operator	Time Equation	Adjoint Process	Fundamental Solution	Representation
ρ	1.1.6	L^*	forward	V	r	$\rho=\int r\rho_s dx$
V	2.3.16	L	backward	ρ	r^*	$V=\int r^*\Phi d\zeta$
q	2.3.12	$\check{L}^*$	forward	u	Γ	$q=\int \Gamma e^{-hy}\rho_s dx$
u	2.3.19	$\check{L}$	backward	q	Γ^*	$u=\int \Gamma^*\Phi e^{hy} d\zeta$
r	2.3.7	L^*	forward	r^*	-	-
r^*	2.3.21	L	backward	r	-	-
Γ	2.3.13	L^*	forward	Γ^*	-	-
Γ^*	2.3.22	$\check{L}$	backward	Γ	-	-

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

CHAPTER 3

NONLINEAR FILTERING PROBLEM: LINEARIZATION OF STOCHASTIC DIFFERENTIAL EQUATIONS

3.1 PRECISE PROBLEM STATEMENT

Suppose the state process $\{x_t, t \in [0,T]\}$ is defined by the stochastic differential equation

$$dx_t = f(x_t)dt + \sigma(x_t)dw_t \quad , \; x_{t_0} = x_0 \tag{3.1.1}$$

and is observed via

$$dy_t = h(x_t)dt + db_t \qquad , \; y_{t_0} = y_0 = 0. \tag{3.1.2}$$

For the moment we shall assume that conditions (A1)-(A5) of Section 2.2 and conditions (i), (ii) of Remark 2.2.2 are satisfied, thus the unnormalized density of x_t given the filtration $\mathcal{F}_t^y$ satisfies the stochastic partial differential equation (1.1.6).

In the ensuing sections we shall investigate the nonlinear filtering problem by first considering a measurable mapping Φ of a state space (E, B) of a Markov process x onto the space $(\tilde{E}, \tilde{B})$ such that $\Phi(B) \subset \tilde{B}$, where P_x denotes the probability measure of the x_t-process. The above mapping is assumed to be a local diffeomorphism with $\Phi(x)$ defined by $\Phi: U \rightarrow V \in \tilde{E}$ with U an open set in E. Now, if we define $z_t \triangleq \Phi(x_t)$ and

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$\mathscr{F}_t^z \triangleq \sigma\{z_s,\ o \le s \le t\}$, the σ-algebra generated by events $\{z_t \in A,\ t \ge 0\}$ for $A \in \tilde{B}$, then

we can define measures $\tilde{\rho}_z$ on $\mathscr{F}_t^z$ such that

$$\tilde{\rho}_z\{z_t \in A\} = \rho_{\Phi^{-1}(z)}\ \{x_t \in \Phi^{-1}(A)\}$$

where z_t also forms a Markov process. Generalizing the above formulation to measures restricted on certain Borel sets (i.e., Borel sets generated by observations) we have a way of recognizing equivalent filtering problems related by local diffeomorphic transformations. When U, V are replaced by R^n and $\Phi(x)$ is a global diffeomorphism onto R^n, then Φ becomes a measurable mapping of the measurable space (R^n, B^n) onto itself.

Diffeomorphic transformations of the above type appear to be significant in relating equivalent filtering problems. That is, once the conditional density of z is obtained, the conditional density of x can be determined and vice-versa. Moreover, if a second type of transformation is used, that is, if a change of scale on the unnormalized density $\rho(x,t)$ is performed by $\rho(x,t) \rightarrow \Psi(x)\ \tilde{\rho}(x,t)$, then $\rho(x,t)$ can be obtained from $\tilde{\rho}(x,t)$ as long as $\Psi(x)$ is a nonnegative function. It is shown by Brockett [27] that the estimation algebra is invariant under the above two types of transformations, with the second transformation sometimes being called a "gauge" transformation, often used in physics (classical mechanics).

The following remark relates conditional densities defined under two coordinate systems.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Remark 3.1.1 Suppose that $\Phi: R^n \to R^n$ is a one-to-one and invertible mapping such that Φ and Φ^{-1} have continuous partial derivatives with respect to the coordinates. Furthermore, suppose that x_t has a conditional density $\rho(x, t)$. Then $z_t = \Phi(x_t)$ also has a density function, which is given by

$$\rho(z,t) = \rho(\Phi^{-1}(z,t)\,|J(z)|$$

where J is the Jacobian matrix of first partials and $|J(z)|$ is the absolute value of its determinant (see Wong and Hajek [132, pp. 9-10]).

3.2 LOCAL LINEARIZATION OF STOCHASTIC DIFFERENTIAL SYSTEMS: NECESSARY AND SUFFICIENT CONDITIONS

In this section, we shall consider the stochastic analog of the control system defined in Appendix 7.A where local state-feedback linearization is considered to transform a nonlinear control problem into an equivalent linear controllable system. We shall show that the stochastic system 3.1.1 is linearizable up to a stopping time τ if and only if the deterministic control system obtained by replacing w_t with u_t is linearizable. The stopping time $\tau(x_0)$ is such that $\{\tau(x_0) \le t\} \in \mathscr{F}_t$, for each $t \in [0,T]$, thus τ is a stopping time with respect to the family of σ-fields $\{\mathscr{F}_t,\ t \in [0,T]\}$.

Suppose the control input of (7.A.1) is replaced by a more general input such as white noise. Then, by rewriting (7.A.1) in terms of an Ito differential equation, we have

$$(\Omega, \mathscr{F}_t, P)\ \Sigma_1:\ dx_t = \tilde{f}(x_t)dt + \sum_{j=1}^{m} \sigma_j(x_t)dw_t^j, \quad 0 \le t \le T \tag{3.2.1}$$

where

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$\tilde{f}^i(x) \triangleq f^i(x) + \frac{1}{2}\sum_{j,k}^{n,m} \sigma_k^j \frac{\partial \sigma_k^i}{\partial x_j};$$

the differential term in $\tilde{f}^i$ is sometimes called the "correction term" (see, e.g., Wong and Hajek [132, pp. 155-163]). Furthermore, assume that the vector fields f, σ_j, j=1, ..., m are bounded on an n-dimensional manifold M. By Friedman [61, p.104-105], there exists a unique solution $x_t \sim \mathcal{F}_t$, $0 \le t < \tau$. We shall first consider the case when m=1 and then generalize to the multi-input case. We are interested in the behavior of (3.2.1) when defined in an open neighborhood $U^0 \subset M$ of some point $x_0 \in M$ where $\tau_U{}^0$ denotes the first exit time of x_t for U^0.

We start by introducing a new probability measure $\tilde{P}$ on $(\Omega, \mathcal{F}_t)$ defined by

$$\Lambda_t^{-1} \triangleq E[\frac{d\tilde{P}}{dP}|\mathcal{F}_t]$$

where

$$\Lambda_t^{-1} = \mathcal{E}\left(\int_0^{t\wedge\tau_{U^0}} \alpha(x_s)dw_s\right) = e^{\int_0^{t\wedge\tau_{U^0}} \alpha(x_s)dw_s - \frac{1}{2}\int_0^{t\wedge\tau_{U^0}} |\alpha(x_s)|^2 ds}.$$

Thus,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$\Lambda_t = e^{-\int_0^{t\wedge\tau_{U^0}} \alpha(x_s)d\tilde{w}_s - \frac{1}{2}\int_0^{t\wedge\tau_{U^0}} |\alpha(x_s)|^2 ds} \tag{3.2.2}$$

where $\alpha(x_t)$ is an adapted, predictable process, and by Girsanov's Theorem (Theorem 7.E.2), under measure $\tilde{P}$, the process $\tilde{w}_t = w_t - \int_0^{t\wedge\tau_{U^0}} \alpha(x_s)ds \in M_{loc}(\mathscr{F}_t, \tilde{P})$ with $\langle \tilde{w}_., \tilde{w}_. \rangle_t = t$ so that $\tilde{w}$ is a standard Brownian motion (see Remark 7.B.4). Therefore, system Σ_1 is transformed to

$$(\Omega, \mathscr{F}_t, \tilde{P}) \quad \Sigma_2:\ dx_t = (\tilde{f}(x_t) + \sigma(x_t)\,\alpha(x_t))dt + \sigma(x_t)d\tilde{w}_t, \quad 0 \le t < \tau_{U^0}. \tag{3.2.3}$$

Moreover, the $\mathscr{F}_t$-adapted process x_t, $0 \le t \le \tau_{U^0}$ with stopping time τ_{U^0} taking values in $U^0 \subset M$ is the solution of

$$x_t = \int_0^t (\tilde{f}(x_s) + \sigma(x_s)\ \alpha(x_s))ds + \int_0^t \sigma(x_s)d\tilde{w}_s, \quad 0 \le t < \tau_{U^0}.$$

When Σ_2 is expressed in terms of a Fisk-Stratonovich representation (using 7.B.1), (3.2.3) is equivalent to

$$(\Omega, \mathscr{F}_t, \tilde{P})\ \Sigma_3:\ dx_t = (f(x_t) + \sigma(x_t)\alpha(x_t))dt + \sigma(x_t)\circ d\tilde{w}_t, \quad 0 \le t < \tau_{U^0}. \tag{3.2.4}$$

Remark 3.2.1 System Σ_3 has the same exact form as (7.A.2) with $\beta(x) = 1$ that was obtained in Appendix 7.A by applying a nonlinear feedback transformation $u = \alpha(x) + v$ to (7.A.1) (i.e., both are in F-S form, but the input in (7.A.3) is a deterministic control whereas the input in (3.2.4) is a random process). Thus, we can

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

view the martingale translation $dw_t = d\tilde{w}_t + \alpha(x_t)dt$ as the dual of nonlinear feedback transformation, and the rest of the analysis from this point forward will, for the most part, be identical to the analysis following for deterministic systems (see, Isidori [79, Chp. 4]).

Although the deterministic linearization is well understood by now, (see, for example, Isidori [79, Chp. 4-5] and references cited in Section 2.1), we shall present some of the concepts involved in obtaining equivalent linear controllable stochastic systems assuming a nonlinear stochastic system is given.

Suppose the measurable map Φ is such that the following conditions are satisfied for all $x^0 \in U^0 \subset U$:

$$\begin{aligned}
&\sum_{j=1}^{n} \frac{\partial \phi_1}{\partial x_j}(f^j(x) + \sigma^j(x)\ \alpha(x) + \sigma^j(x)\cdot d\tilde{w}) = \phi_2(x),\\
&\sum_{j=1}^{n} \frac{\partial \phi_2}{\partial x_j}(f^j(x) + \sigma^j(x)\ \alpha(x) + \sigma^j(x)\cdot d\tilde{w}) = \phi_3(x),\\
&\vdots\\
&\sum_{j=1}^{n} \frac{\partial \phi_{n-1}}{\partial x_j}(f^j(x) + \sigma^j(x)\ \alpha(x) + \sigma^j(x)\cdot d\tilde{w}) = \phi_n(x)
\end{aligned} \tag{3.2.5}$$

where f^j, σ^j are the j-th components of the vector fields f,σ, respectively.

If we use the notation $\langle .,.\rangle_1$ introduced in Appendix 7.A, we can rewrite (3.2.5) as

$$\langle d\phi_i,\ f + \sigma\ \alpha + \sigma\cdot d\tilde{w}\rangle_1 = \phi_{i+1},\ i = 1, \dots, n-1. \tag{3.2.6}$$

Since the first n - 1 equations are not excited by the noise term, we shall have

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$\langle d\phi_i, \sigma \rangle_1 = 0 \tag{3.3.7}$$

$$\langle d\phi_i, f \rangle_1 = \phi_{i+1} \quad , \quad i = 1, \dots, n-1 \tag{3.2.8}$$

and

$$\langle d\phi_n, f + \sigma\alpha + \sigma \cdot dw \rangle_1 = \langle d\phi_n, f \rangle_1 + \langle d\phi_n, \sigma\alpha \rangle_1 + \langle d\phi_n, \sigma \rangle_1 \cdot d\tilde{w}. \tag{3.2.9}$$

Next, using the identity

$$\begin{aligned} \langle d\lambda, [f + \sigma\alpha, \sigma] \rangle_1 &= \langle d\langle d\lambda, \sigma \rangle_1, f + \sigma\alpha \rangle_1 - \langle d\langle d\lambda, f + \sigma\alpha \rangle_1, \sigma \rangle_1 \\ &= L_{f+\sigma\alpha} L_\sigma \lambda - L_\sigma L_{f+\sigma\alpha} \lambda \end{aligned} \tag{3.2.10}$$

given by Isidori [79, pp. 10], which is an application of the Leibnitz rule, we rewrite conditions (3.2.7) - (3.2.8) in terms of ϕ_1 as follows:

$$\langle d\phi_1, ad_f^i(\sigma) \rangle_1 = 0 \quad i = 0, 1, \dots, n-2; \tag{3.2.11}$$

$$\langle d\phi_1, ad_f^{n-1}(\sigma) \rangle_1 \neq 0. \tag{3.2.12}$$

Define

$$\alpha(x) \triangleq - L_f^n \phi_1(x)(L_\sigma L_f^{n-1} \phi_1(x))^{-1}.$$

The last two conditions imply that the existence of the scalar field ϕ_1 is necessary for transforming system Σ_3 to Σ_4 given by:

$$(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})\,\Sigma_4: \begin{bmatrix} dz_{1t} \\ dz_{2t} \\ \vdots \\ dz_{nt} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & 0 \\ \vdots & & & & \vdots \\ 0 & \dots & & \dots & 0 \end{bmatrix} \begin{bmatrix} z_{1t} \\ z_{2t} \\ \vdots \\ z_{nt} \end{bmatrix} dt + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ L_\sigma L_f^{n-1} \phi_1 \end{bmatrix} \cdot d\tilde{w}_t \tag{3.2.13}$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

where the new coordinate system is related to the original coordinate system by $z_{t\wedge\tau_{U^0}} = \Phi(x_{t\wedge\tau_{U^0}})$. Moreover, it can be shown (see Isidori [79, Chp. 4, Lemma 2, pp. 148-149]) that a necessary condition for (3.2.11) - (3.2.12) to hold is the linear independence of the n vector fields

$$\sigma(x),\ ad_f(\sigma)(x),\ \ldots,\ ad_f^{n-1}(\sigma)(x) \tag{3.2.14}$$

for all $x \in U$. But, if condition (3.2.14) is satisfied, then the distribution

$$\Delta = \text{span}\{\sigma, ad_f(\sigma), \ldots, ad_f^{n-2}\}(x) \tag{3.2.15}$$

is nonsingular and (n-1)-dimensional for all $x \in U$. Thus, since Δ is nonsingular for $x^0 \in U$, condition (3.2.11) implies that distribution Δ is completely integrable for $x^0 \in U$, so there exists a scalar field ϕ_1. Finally, due to Frobenius Theorem, (Theorem 7.A.1) the distribution Δ is involutive for $x^0 \in U$, thus proving necessity of involutive distribution Δ. The equivalence of systems Σ_2, Σ_4 is understood up to the stopping time τ_{U^0}. That is, the solution x_t of Σ_2 can be obtained from the solution of Σ_4 at least up to a stopping time τ_{U^0}, where $z_{\tau_{U^0}} = \Phi(x_{\tau_{U^0}})$ if $x_{\tau_{U^0}} \in U^0 \subset U$ and $z_{\tau_{U^0}} = 0$ otherwise, because of the assumption that Φ is a local diffeomorphism on U and $z_0 = \Phi(x_0) \in U^0 \subset U$.

Conditions (3.2.11) - (3.2.12) are also sufficient as shown by Isidori [79, Lemma 2.5, pp. 165].

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Remark 3.2.2 For the rest of this chapter we shall assume that $L_\sigma L_f^{n-1}\phi_1$ is independent of x, so, without loss of generality, we can set $L_\sigma L_f^{n-1}\phi_1 = 1$.

Definition: Locally Linearizable Stochastic Systems.

The nonlinear single-input stochastic Ito differential system Σ_1 defined on a probability space $(\Omega, \mathcal{F}_t, P)$ is said to be locally linearizable if given an initial condition x_0, there exist a neighborhood U^0 of x_0, a local diffeomorphism Φ defined on U^0, and a Wiener process $\tilde{w}_t = w_t - \int_0^{t\wedge\tau_{U^0}} \alpha(x_s)ds$ also defined on U^0, up to a stopping time $\tau_{U^0} \triangleq \inf\{x_t \notin U^0;\ x_0 \in U^0\}$, such that the corresponding stochastic equation when defined on a new probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}_t, \tilde{P})$ is linear and of the form

$$z_{t\wedge\tau_{U^0}} = \Phi(x_0) + \int_0^{t\wedge\tau_{U^0}} Az_s\,ds + \int_0^{t\wedge\tau_{U^0}} Bd\tilde{w}_s. \tag{3.2.16}$$

Furthermore, the nxn matrix A and nx1 vector B are a controllable pair.

Theorem 3.2.1

Suppose that the single-input stochastic system Σ_1 is defined on a C^∞-manifold M of dimension n, with f, σ, both bounded on M of C^∞-class. Then, system Σ_1 is locally linearizable in a neighborhood U^0 of x_0 if and only if the deterministic control system is locally linearizable. That is, if and only if the following conditions are satisfied:

(i) $\{\sigma, ad_f(\sigma), \ldots, ad_f(\sigma)\}\,(x_0) = T_{x_0}U^0$;

(ii) $\Delta = \text{span}\,\{\sigma, ad_f(\sigma), \ldots, ad^{n-2}{}_f(\sigma)\}\,(x_0)$ is involutive near x_0.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Proof: Follows from the deterministic analog by choosing $U^0 \subset U$. QED

The Multidimensional Case. The result of Theorem 3.2.1 can be extended to stochastic differential equations with more than one input. This follows from its deterministic analog stated by Isidori [79, Chp. 5]. Once we assume deterministic linearization we proceed in the exact manner as presented for the single-input case. Thus we introduce the coordinate transformation

$$z^i = (\phi_i, L_f \phi_i, \dots, L_f^{r_i-1} \phi_i)^T = \Phi^i, \quad i = 1, 2, \dots, m, \; r_1 + r_2 + \dots + r_m = n$$

and define the exponential formula as in Theorem 7.E.2 by

$$\Lambda_t^{-1} \triangleq e^{\sum_{j=1}^{m} \int_0^{t \wedge \tau_{U^0}} \alpha^j dw_s^j - \int_0^{t \wedge \tau_{U^0}} \frac{1}{2} \sum_{j=1}^{m} |\alpha^j|^2 ds} \tag{3.2.17}$$

Therefore, by martingale translation, Theorem 7.E.3,

$$\tilde{w}_t^j = w_t^j - \langle w_\cdot^j, \sum_{i=1}^{m} \int_0^{\cdot} \alpha^i dw_s^i \rangle_{t \wedge \tau_{U^0}}, \quad j = 1, 2, \dots, m \tag{3.2.18}$$

is a standard Brownian motion with respect to measure $\tilde{\mathcal{P}}$. Writing (3.2.18) in vector form we have,

$$\tilde{w}_t = w_t - \int_0^{t \wedge \tau_{U^0}} I_m \alpha_s ds \in M_{loc}(\mathcal{F}_t, \tilde{\mathcal{P}})$$

where $\tilde{w}_t$, w_t, α are mx1 vectors and I_m is the mxm identity matrix. If we define

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$\alpha(x) \triangleq \beta^{-1}(x)\,\tilde{\alpha}(x) \triangleq - \begin{bmatrix} L_{\sigma_1} L_f^{r_1-1} \phi_1 & \cdots & L_{\sigma_m} L_f^{r_1-1} \phi_1 \\ \vdots & & \\ L_{\sigma_1} L_f^{r_m-1} \phi_m & \cdots & L_{\sigma_m} L_f^{r_m-1} \phi_m \end{bmatrix}^{-1} \begin{bmatrix} L_f^{r_1} \phi_1 \\ \vdots \\ L_f^{r_m} \phi_m \end{bmatrix}(x) \tag{3.2.19}$$

then, under probability space $(\tilde{\Omega}, \mathcal{F}_t, \tilde{P})$ the original system (3.2.1) is transformed to

$$dz_t^i = \bar{A}_i z_t^i dt + \bar{B}_i d\bar{w}_t^i \quad , \; \bar{w}_t^i = \sum_{j=1}^{m} \beta^{ij} d\tilde{w}_t^j \tag{3.2.20}$$

where β^{ij} denotes the (i,j)-th component of matrix β.

As in the deterministic case, the r_i x r_i matrix $\bar{A}_i$ and r_i x 1 vector $\bar{B}_i$ are of the form

$$\bar{A}_i = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 \cdots & \vdots \\ \vdots & & & \ddots & 1 \\ 0 & 0 & \cdots & & 0 \end{bmatrix}, \; \bar{B}_i = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}, \quad 1 \le i \le m. \tag{3.2.21}$$

<u>Remark 3.2.3</u> With appropriate modifications, Theorem 3.2.1 for single-input systems also applies to multi-input systems.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

3.3 GLOBAL LINEARIZATION OF STOCHASTIC DIFFERENTIAL SYSTEMS: NECESSARY AND SUFFICIENT CONDITIONS

Here we will use the results of the preceding section and the conditions of global linearization of deterministic control systems stated in Theorem 7.A.2 to establish necessary and sufficient conditions for global linearization of stochastic differential systems.

Suppose $U^0 = U = M = R^n$ and we require Φ to be a global diffeomorphism onto R^n. Then for the single-input case, we define a new measure $\tilde{\mathcal{P}}$ as in Section 3.2 and, if we can show that it is a probability measure, then $\tilde{\mathcal{P}} << \mathcal{P}$. We start by assuming that f, σ satisfy globally the Lipschitz and linear growth condition of Theorem 7.C.1.

Definition: Global Linearization of Stochastic Systems.

The nonlinear stochastic system (3.2.1) is said to be globally linearizable if there exist a global diffeomorphism Φ onto R^n and a martingale translation $d\tilde{w}_t = dw_t - \alpha(x_t)\,dt$ such that $\tilde{w}_t$ is a Brownian motion under a new measure $\tilde{\mathcal{P}} << \mathcal{P}$, where $\alpha(x)$ is an adapted, predictable process on the new probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}_t, \tilde{\mathcal{P}})$ and coordinate $z_t = \Phi(x_t)$ satisfies (3.2.16) globally with (A, B) a controllable pair.

Theorem 3.3.1

Suppose the stochastic differential system (3.2.1) with m=1 is defined on R^n with f,σ satisfying a global Lipschitz and linear growth condition. Then this system is globally linearizable if and only if the deterministic control system given by 7.A.1 with m=1 and g replaced by σ is globally linearizable and there exists a $\delta > 0$ such that

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$\sup_{t\in[0,T]} E\, e^{\delta\alpha^2(x)} < \infty$, where $\alpha(x) = - L_f^n \phi_1(x)$. That is, conditions (i)-(iii) of Theorem 7.A.2(b) are satisfied with the additional condition:

$$\sup_{t\in[0,T]} E\, e^{\delta\alpha^2(x)} < \infty, \text{ for some } \delta > 0.$$

Proof: The proof follows that of the deterministic control problem and is given by Dayawansa, Boothby and Elliott [41]. The sup bound is a sufficient condition for measure $\tilde{\mathcal{P}}$ to be absolutely continuous with respect to $\mathcal{P}$ as stated in Remark 7.E.2 and given by Gihman and Skorohod [65, Thm. 3, pp.90] or Liptser and Shiryayev [103, pp. 220]. This condition is satisfied when f is Lipschitz and has a linear growth, σ is bounded, and $\alpha(x)$ has a linear growth, as given by Liptser and Shiryayev [103, Thm. 4.7, pp. 137-139]; in our case if $\phi_1(x)$ has a of linear growth, this condition is always satisfied. QED

3.4 FINITE DIMENSIONAL FILTERS: GLOBAL CASE

In this section we shall derive a set of sufficient statistics for obtaining the unnormalized conditional density for the nonlinear filtering problem (3.1.1), (3.1.2) stated in Section 3.1. Throughout this section we shall assume that conditions of Theorem 3.3.1 are satisfied so that (3.1.1) is linearizable through the global diffeomorphism map $\Phi \in R^n$. Even though we restrict ourselves to this class of linearizable problems, they are by far less restrictive than the current existing filtering problems that admit finite-dimensional filters.

We start by defining the transformation

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$z_t \triangleq \Phi(x_t) = (h, L_f h, \ldots, L_f^{n-1} h)^T$$

and noting that, since $\phi_1 = h$ (the signal component of (3.1.2)) the condition given in Theorem (3.3.1) regarding linear growth in fact becomes a requirement on h. Proceeding, we obtain the equivalent filtering problem

$$(\Omega, \mathcal{F}_t, P): \begin{cases} dz_{1t} = z_{2t}\, dt \\ dz_{2t} = z_{3t}\, dt \\ \quad\vdots \\ dz_{(n-1)t} = z_{nt}\, dt & , z_0 = \Phi(x_0) \\ dz_{nt} = L_f^n h(x) dt + dw_t \\ dy_t = h(x_t) dt + db_t & , y_0 = 0 \end{cases} \tag{3.4.1}$$

where the only nonlinearity appears in the last component of state coordinate z. We shall consider two measure transformations. The first will express the state process z in terms of a certain linear stochastic Ito system, and the second will reduce the observation equation to $y_t = b_t$.

First consider the following martingale

$$m_t \triangleq \int_0^t \alpha'(x_s) dw_s \quad \in M(\mathcal{F}_t, P)$$

where $\alpha'(x)$ is some arbitrary adapted, predictable process that will be chosen to satisfy a linear growth condition. If we introduce a new measure P' on $(\Omega, \mathcal{F}_t)$ defined by

$$L_T \triangleq E[\frac{dP'}{dP} | \mathcal{F}_T] \text{ where } L_T = \mathcal{E}(\int_0^T \alpha'(x_s) dw_s) \tag{3.4.2}$$

Then if $P << P'$,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$\frac{d\mathcal{P}}{d\mathcal{P}'} = e^{-\int_0^t \alpha'(x_s)dw_s + \frac{1}{2}\int_0^t |\alpha'(x_s)|^2 ds}.$$

Consequently, by the martingale translation, Theorem 7.E.3, under measure $\mathcal{P}'$ the process

$$dw'_t \triangleq dw_t - \alpha'(x_t)dt \in M(\mathcal{F}_t, \mathcal{P}') \tag{3.4.3}$$

is a standard Brownian motion with respect to probability measure $\mathcal{P}'$, which follws from Remark 7.B.4. Therefore, substituting (3.4.3) into (3.4.1), the last component z_{nt} satisfies

$$dz_{nt} = \alpha(x)dt + dw' \tag{3.4.4}$$

where we define $\alpha(x) \triangleq L_f^n h(x) + \alpha'(x)$. Substituting (3.4.3) into the expression of $\frac{d\mathcal{P}}{d\mathcal{P}'}$ and using (3.4.4), after some simple algebra we can write

$$\frac{d\mathcal{P}}{d\mathcal{P}'} = e^{\int_0^T (L_f^n h(x) - \alpha(x_s))dz_{ns} - \frac{1}{2}\int_0^T (|L_f^n h(x_s)|^2 - |\alpha(x_s)|^2)ds} \tag{3.4.5}$$

Note that under the new probability measure $\mathcal{P}'$, b_t remains a standard Brownian motion independent of z_t, but the distribution of z_t is changed. Also, $\alpha(x)$ will be chosen to satisfy a linear growth condition so that $E[L_T] = 1$ and thus, $\mathcal{P}' << \mathcal{P}$.

A different approach to the definition of measure $\mathcal{P}'$ can be established along the lines of weak solutions to stochastic differential equations as discussed in Appendix 7.C.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

If we assume Φ: $(\Omega, \mathcal{F}_t) \rightarrow (\tilde{\Omega}, \tilde{\mathcal{F}}_t)$ under probability measure $\wp$, system 3.4.1, when viewed in terms of the coordinate variable $z = (z^1, z^2)^T$, $z^1 \in R^{n-1}$, $z^2 \in R^1$, can be expressed as

$$\begin{aligned} dz_t^1 &= f^1(z_t^1, z_t^2)dt \\ dz_{nt} &= f^2(z_t^1, z_t^2)dt + \bar{\sigma}(z_t^1, z_t^2)dw_t \end{aligned}$$

where $z^2 \triangleq z_n$, $z^1 = (z_1, z_2, ..., z_{n-1})^T, \bar{\sigma}$ is nonsingular, and f^1 is Lipschitz in z_t^1, uniformly in z_t^2. Then for each trajectory $z_{nt}(\tilde{\omega})$, $\tilde{\omega} \in \tilde{\Omega}$, there is a unique solution $z_t^1 = \xi(z_{nt}(\tilde{\omega}))$ and the differential equation of z_{nt} can be written as

$$dz_{nt} = f^2(\xi(z_{nt}), z_{nt})dt + \bar{\sigma}(\xi(z_{nt}), z_{nt})dw_t.$$

This equation has a weak solution, as shown by Davis [37] and presented by Elliott [47, pp. 224]. Thus, by the definition of weak solutions given in Appendix 7.C, we must have $\wp\{\tilde{\omega}, \int_0^T |\bar{\sigma}(z_{nt}(\tilde{\omega}))|^2 dt < \infty\} = 1$ a.s. and $\wp\{\tilde{\omega}; \int_0^T |f^2(z_{nt}(\tilde{w}))|^2 dt < \infty\} = 1$ a.s. which imply $\wp \ll \wp'$. However, by Remark 7.E.3 it follows that if $\wp'\{\tilde{\omega}; \int_0^T |\bar{\sigma}^{-1}(\tilde{\omega}) f^2(\tilde{\omega})|^2 dt < \infty\} = 1$ a.s. we must have $\wp' \ll \wp$. Therefore we conclude that $\wp \sim \wp'$.

Next we define another new probability measure $\tilde{\wp}$ on $(\Omega, \mathcal{F}_t)$ to be absolutely continuous with respect to $\wp'$ in the following manner. Consider the martingale

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$m_t \triangleq -\int_0^t h(x_s)db_s \in M(\mathcal{F}_t, \mathcal{P}')$$

where $h(x_t)$ is an adapted, predictable process. Then by defining $\mathcal{E}(m_T) \triangleq E'[\frac{d\tilde{\mathcal{P}}}{d\mathcal{P}'}|\mathcal{F}_T]$,

where E' denotes expectation with respect to measure $\mathcal{P}'$, we have

$$\frac{d\tilde{\mathcal{P}}}{d\mathcal{P}'} = e^{-\int_0^T h(x_s)db_s - \frac{1}{2}\int_0^T |h(x_s)|^2 ds}.$$

Consequently,

$$\frac{d\mathcal{P}'}{d\tilde{\mathcal{P}}} = e^{\int_0^T h(x_s)dy_s - \frac{1}{2}\int_0^T |h(x_s)|^2 ds} \tag{3.4.6}$$

where, because of the martingale translation, theorem, the process y_t, $t \in [0,T]$ is a standard Brownian motion as was shown in Remark 2.2.2(ii)(a).

Therefore by (7.E.6) for any integrable function $\Psi(x)$

$$E[\Psi(x_t)|\mathcal{F}_t^y] = \frac{\tilde{E}[\Psi(\Phi^{-1}(z))\,\frac{d\mathcal{P}}{d\mathcal{P}'} \times \frac{d\mathcal{P}'}{d\tilde{\mathcal{P}}}|\mathcal{F}_t^y]}{\tilde{E}[\frac{d\mathcal{P}}{d\mathcal{P}'} \times \frac{d\mathcal{P}'}{d\tilde{\mathcal{P}}}|\mathcal{F}_t^y]}. \tag{3.4.7}$$

Now, if we define $\Lambda_T \triangleq \frac{d\mathcal{P}}{d\mathcal{P}'} \times \frac{d\mathcal{P}'}{d\tilde{\mathcal{P}}}$ it follows that

$$\Lambda_t = \mathcal{E}(\int_0^T h(\Phi^{-1}(z_s))dy_s) \times \frac{d\mathcal{P}'}{d\mathcal{P}}. \tag{3.4.8}$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Remark 3.4.3 The sufficient condition for $\tilde{\mathcal{P}}$ to be absolutely continuous with respect to $\mathcal{P}'$ again follows from the existence of some $\delta > 0$ such that $\sup_{t\in[0, T]} E e^{\delta|h(x_t)|^2} < \infty$, where $h(x_t)$ satisfies a linear growth condition. Moreover, $\tilde{\mathcal{P}} << \mathcal{P}'$ if the weaker condition $\mathcal{P}'\{\tilde{\omega}: \int_0^T |h(x_t(\omega))|^2 dt < \infty\} = 1$ a.s. is satisfied. The proof is found in Liptser and Shiryayev [103, Example 4, pp. 221-222] and is due to $h(x_t)$ being independent of b_t. Finally, if we choose $\alpha(x) = \sum_{i=1}^{n} c_i L_f^{i-1} h(x)$ the filtering problem (3.4.1) is equivalent to

$$(\tilde{\Omega}, \tilde{\mathcal{F}}_t, \tilde{\mathcal{P}}): dz_t = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 \dots & 0 \\ \vdots & & \ddots & & \vdots \\ & & & & 1 \\ c_1 & c_2 & & \dots & c_n \end{bmatrix} z_t dt + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} dw'_t \tag{3.4.9}$$

where any conditional estimate with respect to $\mathcal{F}_t^y$ under measure $\mathcal{P}$ is related to the conditional estimate under measure $\tilde{\mathcal{P}}$ through (3.4.7) with Λ_t defined by (3.4.8).

The next step is to find the density of Λ_t. Suppose we assume the existence of a scalar potential $V(z)$, $\forall z \in R^n$ such that

$$(L_f^n h(x) - \sum_{j=1}^{n} c_j L_f^{j-1} h(x))|_{x=\Phi^{-1}(z)} = \frac{\partial V(z)}{\partial z_n} \tag{3.4.10}$$

$$\frac{\partial V(z)}{\partial z_i} = 0, \ 1 \le i \le n-1 \ . \tag{3.4.11}$$

Applying the Ito differential rule to $V(z_t)$, we find

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$V(z_t) - V(z_0) = \int_0^t (L_f^n h(\Phi^{-1}(z_s)) - \sum_{j=1}^n c_j z_{js})\, dz_{ns}$$

$$+ \frac{1}{2}\int_0^t [\frac{\partial}{\partial z_n}(L_f^n h(\Phi^{-1}(z_s)) - \sum_{j=1}^n c_j z_{js})]\, ds. \tag{3.4.12}$$

Furthermore, integrating $\int_0^t h_s dy_t$ by parts, yields

$$\int_0^t z_{1s} dy_s = z_{1t} y_t - \int_0^t z_{2s} y_s ds. \tag{3.4.13}$$

Substituting (3.4.12), (3.4.13) into (3.4.8), Λ_t is expressed as

$$\Lambda_t = e^{V(z_t) - V(z_0) + z_{1t} y_t}\, e^{-\frac{1}{2}\int_0^t [\frac{\partial}{\partial z_n}(L_f^n h(\Phi^{-1}(z)) - \sum_{j=1}^n c_j z_{js})] ds}$$

$$\times e^{-\frac{1}{2}\int_0^t (|L_f^n h(\Phi^{-1}(z))|^2 - \sum_{j=1}^n |c_j z_{js}|^2) ds - \int_0^t z_{2s} y_s ds - \frac{1}{2}\int_0^t z_{1s}^2 ds}.$$

By (2.3.2), the numerator of (3.4.7) can be written as

$$\tilde{E}\,\{\Psi(\Phi^{-1}(z_t))\, e^{V(z) - V(z_0) + z_1 y}\, \tilde{E}\,(e^{-\int_0^t Q(s,z_s) ds} \,|\, \mathcal{F}_t^y, z_t = z, z_{t_0} = z_0)\,\} \tag{3.4.14}$$

$$= \int_{R^n} \Psi(\Phi^{-1}(z))\, e^{V(z) - V(z_0) + z_1 y} q(z, t) dz$$

where q(z,t) satisfies the Feynman-Kac formula

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$\frac{\partial}{\partial t} q(z,t) = L(t)^* q(z,t) - Q(t,z_t)q(z,t), \quad 0 \le t \le T$$
$$q(z,0) = p_0(z) \tag{3.4.15}$$

with L(t) the Kolmogorov's operator associated with Markov process z described by the linear system (3.4.9) and Q(t,z) defined by

$$Q(t,z) \triangleq \frac{1}{2} \frac{\partial}{\partial z_n} (L_f^n h(\Phi^{-1}(z)) - \sum_{j=1}^{n} c_j z_j)$$
$$+ \frac{1}{2}(|L_f^n h(\Phi^{-1}(z))|^2 - \sum_{j=1}^{n} |c_j z_j|^2) + y_t z_2 + \frac{1}{2} z_1^2 . \tag{3.4.16}$$

The Feynman-Kac formula (3.4.15) follows by a slight modification of the proof presented in Theorem 2.3.1. Finally, if the solution to (3.4.15) can be determined, the unnormalized conditional density of $\{z_t, 0 \le t \le T\}$ is given by

$$\rho(z,t) = e^{V(z) - V(z_0) + z_1 y} q(z,t) \tag{3.4.17}$$

which is related to the unnormalized conditional density of $\{x_t, 0 \le t \le T\}$ through Remark 3.1.1. The following theorem provides a set of sufficient condition for obtaining a finite-dimensional filter, i.e., the sufficient statistics are finite-dimensional and evolve on a finite dimensional manifold.

Theorem 3.4.1

Suppose we are given the nonlinear system of Section 3.1.1. If we assume that the necessary and sufficient conditions of Theorem 3.3.1 are satisfied with

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$\Phi(x) = (h(x), L_f h(x), \ldots, L_f^{n-1} h(x))^T$, then, if there exists a scalar potential function satisfying (3.4.10), (3.4.11) with the additional condition

$$\frac{\partial}{\partial z_n} L_f^n h(\Phi^{-1}(z)) + |L_f^n h(\Phi^{-1}(z))|^2 = z^T \tilde{A} z + \tilde{B}^T z + \tilde{C}, \qquad (3.4.18)$$

the unnormalized conditional density of the state process $\{z_t, 0 \leq t \leq T\}$ is given by

$$\rho(z, t) = e^{V(z) - V(z_0) + z_1 y} e^{-\frac{1}{2}(z - \mu_t)^T \Sigma_t^{-1} (z - \mu_t)} \qquad (3.4.19)$$

and is determined in terms of two statistics μ_t, Σ_t satisfying

$$\dot{\mu}_t = (A - \Sigma_t \tilde{A}') \mu_t - \frac{1}{2} \Sigma_t \tilde{B} - \frac{1}{2} \Sigma_t H^T y_t \quad , \mu_0 = \Phi(x_0) \qquad (3.4.20)$$

$$\dot{\Sigma}_t = A \Sigma_t + \Sigma_t A^T - \Sigma_t \tilde{A}' \Sigma_t + B B^T \quad , \Sigma_0 = 0 \qquad (3.4.21)$$

Proof: As given above, with (3.4.19) as proved in the Appendix 7.G. QED

Remark 3.4.4 The conditions of Theorem 3.4.1 are satisfied when the nonlinear component of (3.4.1) satisfies $L_f^n h(\Phi^{-1}(z)) = \tanh(z_n)$. Thus the conditional density of $\{z_t, 0 \leq t \leq T\}$ is finite-dimensional and is given by (3.4.19).

3.4.1 The Stochastic Control Problem

The connection between filtering and stochastic control was established by Fleming and Mitter [58] by considering the transformation s(z,t) = - log q(z,t) when the filtering problem is nondegenerate. This transformation was also used by Fleming [57]

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

to derive the Ventcel-Freidlin estimates for diffusion processes depending on small parameters.

Here we shall use the above logarithmic transformation to obtain the dynamic programming equation of an optimal control problem.

Suppose we define

$$q(z,t) = e^{-s(z,t)}$$

where q(z,t) satisfies (3.4.15), then

$$\frac{\partial}{\partial t}s(z,t) = -\frac{1}{q(z,t)}\frac{\partial}{\partial t}q(z,t).$$

After some simple algebra we deduce that s(z,t) satisfies the nonlinear degenerate parabolic equation

$$\begin{aligned}\frac{\partial}{\partial t}s(z,t) &= \frac{1}{2}\mathrm{Tr}(BB^T\frac{\partial^2}{\partial z^2}s(z,t)) - \frac{1}{2}(\frac{\partial}{\partial z}s(z,t))^T BB^T\frac{\partial}{\partial z}s(z,t)\\ &- (Az)^T\frac{\partial}{\partial z}s(z,t) + (Q(t,z) + \mathrm{Tr}(A)).\end{aligned}$$

Let us now replace t by T-s so that a backward PDE for s(z,t) is obtained. Then

$$\begin{aligned}-\frac{\partial}{\partial s}s(z,T-s) &= \frac{1}{2}\mathrm{Tr}(BB^T\frac{\partial^2}{\partial z^2}s(z,T-s)) - \frac{1}{2}(\frac{\partial}{\partial z}s(z,T-s))^T BB^T\frac{\partial}{\partial z}s(z,T-s)\\ &- (Az)^T\frac{\partial}{\partial z}s(z,T-s) + (Q(z,T-s) + \mathrm{Tr}(A)).\end{aligned}$$

If we set $\bar{s}(z,s) = s(z,T-s)$ we deduce

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$\frac{\partial}{\partial s}\bar{s}(z,s) + \frac{1}{2}\mathrm{Tr}(BB^T\frac{\partial^2}{\partial z^2}\bar{s}(z,s)) - \frac{1}{2}(\frac{\partial}{\partial z}\bar{s}(z,s))^T BB^T\frac{\partial}{\partial z}\bar{s}(z,s)$$
$$-(Az)^T\frac{\partial}{\partial z}\bar{s}(z,s) + (Q(z,T-s) + \mathrm{Tr}(A)) = 0 \tag{3.4.22a}$$
$$\bar{s}(s,T) = -\log p_0(z). \tag{3.2.22b}$$

It can be easily shown that (3.4.22) is the dynamic programming equation to the stochastic control problem

$$d\xi_t = -A\xi_t dt + Bu_t dt + Bdw_t, \quad 0 \le t \le T \tag{3.4.23}$$

$$J(u) = E[\int_0^T \{\frac{1}{2}u_t^T BB^T u_t + Q(\xi, T-t) + \mathrm{Tr}(A)\}dt + \bar{s}(\xi,T)|x] \tag{3.4.24}$$

where $u_t \in R^1$, $B = [0, 0, ..., 1]$, A the nxn matrix identified by (3.4.9). That is,

$$\bar{s}(z,s) = \min_{u \in U_{ad}} J(u). \tag{3.4.24}$$

Notice that (3.4.22a) is degenerate parabolic therefore we can no longer assume that a solution $\bar{s}(z,t) \in C^{1,2}$ exist as required by the Verification theorem given by Fleming and Rishel [56, Thm. 4.1, pp. 159]. Thus, no existence of control $u(t, \xi_t)$ can be deduce. As explained in Remark 7.C.3 we shall consider generalized (weak) solutions to (3.4.22). However, since the transition density of

$$d\xi_t = -A\xi_t dt + Bdw_t$$

satisfy

$$\int_D |p(y,t;x,s)|^k dtdx < \infty, \quad 0 \le s \le t \le T$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

then by Fleming and Rishel [56, Chp. VI, pp. 177-178] any Borel measurable feedback control law $u(t,\xi_t)$ is admissible. Equivalently one can show that there is a solution ξ_t of (3.4.23) which is unique in probability law using the approach presented in Chapter 2. Finally, excluding the details which are found in Fleming and Rishel [56, Chp. VI, pp. 177-178] we conclude that if $\bar{s}(s,t)$ is a solution of (3.4.22) in some L^P space, then

(i) $\bar{s}(s,t) \le J(u)$ for some bounded Borel measurable u,

(ii) for the optimal control $u^* = -\frac{\partial}{\partial \xi_n}\bar{s}(\xi,t)$ we have $\bar{s}(s,t) = \min J(u^*)$,

(iii) the optimal cost $\bar{s}(s,t)$ is related to the solution of the filtering problem of Theorem 3.4.1 by

$$s(z,t) = -\log q(z,t).$$

3.5 FINITE DIMENSIONAL FILTERS: LOCAL CASE

Here we consider the filtering problem stated in Section 3.4 with the assumption of a global diffeomorphism map Φ onto R^n removed. Thus we use Theorem 3.2.1 and assume the existence of a local diffeomorphism $\Phi: U^0 \to V^0$. The motivation for investigating filtering problems of this nature stems from the fact that the local results of stochastic linearization (Theorem 3.2.1) are less restrictive than the global results. In addition, no explicit recursive generated set of statistics are currently known for the representation of the unnormalized conditional density when the Beneš [4] multi-

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

dimensional conditions: a) $\sigma(x)$ is constant, b) $h(x) = Hx$, and c) $f(x) = \nabla F(x)$, where $F(x)$ is a function on all of R^n satisfying

$$\Delta F(x) + ||\nabla F(x)||^2 + ||Hx||^2 = x^T\tilde{A}x + \tilde{B}^Tx + \tilde{C}\ ,\ A \geq 0,$$

are satisfied only locally on an open bounded domain of R^n (Δ is the Laplacian operator).

Here we shall give conditions similar to Section 3.4 for the existence of finite-dimensional statistics for the representation of the unnormalized density. This unnormalized conditional density is given in terms of a backward PDE restricted to an open bounded neighborhood V^0 of z_0 as described in Section 2.5. The derivation of the above backward PDE is similar to the derivation of (2.5.5). The adjoint equation to this backward PDE can be obtained as in Theorem 2.5.1 or Remark 2.5.3.

We start by considering the filtering problem (3.1.1), (3.1.2) represented by

$$(\tilde{\Omega}, \mathcal{F}_t, \tilde{P}): \begin{cases} dz_t = A\ z_t + B\,L_f^n h(\Phi^{-1}(z_t))\,dt + Bdw_t & ,\ 0 \leq t < \tau_{V^0} \\ dy_t = z_{1t}dt + db_t \end{cases} \qquad (3.5.1)$$

which is the result of the local diffeomorphism Φ chosen as in Section 3.3.4, where U^0 is an open neighborhood of x_0. Let $\{z_s, 0 \leq s < \tau_{V^0}\}$ be the solution of (3.5.1) defined up to time τ_{V^0}, the time it takes for the process z to hit the boundary ∂V^0 of class C^2. If we define the Radon-Nikodym derivative Λ as in (3.4.8) with T replaced by τ_{V^0}, then

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

we can define a new measure $\tilde{P} \sim P$. By 7.E.6, for any bounded function Ψ, the numerator of (3.4.7) can be expressed as

$$\tilde{E}[\Psi(\Phi^{-1}(z_t))\, I_{\{\omega:\, \tau_{V^0}(\omega)>t\}} \Lambda_t | \mathcal{F}_t^y] \tag{3.5.2}$$

which is the result of Section 2.5. If we assume conditions (3.4.9), (3.4.10) are satisfied locally, then (3.5.2) can be expressed as

$$\tilde{E}\,(\Lambda_{0,s}\tilde{E}[I_{\{\omega:\, \tau_{V^0}(\omega)>t\}}\, \Psi(z_t)\, e^{V(z_t)-V(z_0)+z_{1t}-z_{1s}y_s}\, e^{-\int_s^t Q(s,\, z_s)ds} |z_s]|\mathcal{F}_t^y) \tag{3.5.3}$$

for $s<\tau_{V^0}$ a.s. where the definitions of V(z) and Q(t,z) are given as in Section 3.4. If we define

$$u(z,s) \triangleq e^{z_{1s}y_s}\tilde{E}[I_{\{\omega;\, \tau_{V^0}(\omega)>t\}}\, \Psi(z_t) e^{V(z_t)-V(z_s)+z_{1t}y_t-z_{1s}y_s}\, e^{-\int_s^t Q(s,z)ds} |z_s) \tag{3.5.4}$$

as was similarly done in Section 2.5, we observe that (3.5.4) satisfies the backward PDE

$$\begin{aligned}
&\frac{\partial u}{\partial s}(z,s) + L(t)u(z,s) - Q(s,z)u(z,s) = 0 \quad , (s,x)\in[0,t]xV^0 \\
&u(z,t) = \Psi(z)e^{V(z) - V(z_0) + z_1y_t} \quad , x\in V^0 \\
&u(z,s) = 0 \quad , (s,x)\in[0,t)x\partial V^0.
\end{aligned} \tag{3.5.5}$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Suppose that the initial density of z_0 is $p_0(z)$, and (3.5.5) can be solved; then (3.5.2) is equal to

$$\int_{V^0} \Psi(\Phi^{-1}(z_t))\ u(z,0)p_0(z)dz. \tag{3.5.6}$$

Remark 3.5.1 An equation for the adjoint process q(x,t) to u(x,t) could also be derived using the methodology of Remark 2.5.3. Therefore, if we can solve the initial-boundary value problem above, the unnormalized density can be obtained.

3.6 LOWER AND UPPER BOUNDS ON FUNCTIONS OF STATE ESTIMATES

In this section we shall derive lower and upper bounds on some of the components of the minimum-mean-square-error (mmse) state estimate $E[x_t|\mathscr{F}_t^y]$ and its corresponding error-covariance. Furthermore, lower and upper bounds on nonlinear functions of the state such as $E[\Psi(x_t)|\mathscr{F}_t^y]$ will also be obtained. Throughout this section we assume that the filtering problem (3.1.1), (3.1.2) satisfies the necessary and sufficient conditions stated in Theorem 3.3.1.

Our starting point is (3.4.9) defined on the probability space $(\tilde{\Omega}, \tilde{\mathscr{F}}_t, \tilde{\mathcal{P}})$, where the exponential formula is defined as in (3.4.8) and is equal to

$$\frac{d\mathcal{P}}{d\tilde{\mathcal{P}}} \triangleq \Lambda_T = \mathscr{E}\Big(\int_0^T [L_f^n h(\Phi^{-1}(z_s)) - \sum_{j=1}^n c_j z_{js}]dw'_s\Big)x\mathscr{E}\Big(\int_0^T z_{1s}dy_s\Big). \tag{3.6.1}$$

Recall (7.E.4) that for any integrable function Ψ

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$E[\Psi(x_t) \mid \mathcal{F}_t^y] = \frac{\tilde{E}[\Psi\ (\Phi^{-1}(z_t))\ \Lambda_t \mid \mathcal{F}_t^y]}{\tilde{E}[\Lambda_t \mid \mathcal{F}_t^y]} \tag{3.6.2}$$

and

$$\tilde{E}[\Psi(\Phi^{-1}(z_t)) \mid \mathcal{F}_t^y] = \frac{E\,[\Psi(x_t)\Lambda_t^{-1} \mid \mathcal{F}_t^y]}{E\ [\Lambda_t^{-1} \mid \mathcal{F}_t^y]}. \tag{3.6.3}$$

Bayes formula (3.6.2) relates conditional estimates of measure $\mathcal{P}$ to that of measure $\tilde{\mathcal{P}}$ as in (3.4.7), whereas Bayes formula (3.6.3) relates conditional estimates of measure $\tilde{\mathcal{P}}$ to that of $\mathcal{P}$. Next, we define the conditional correlation coefficients of Ψ, Λ and Ψ, Λ^{-1} as

$$P_{\Psi\Lambda}^{\tilde{\mathcal{P}}} \triangleq \frac{\tilde{E}[\Psi(\Phi^{-1}(z_t))\Lambda_t \mid \mathcal{F}_t^y] - \tilde{E}[\Psi(\Phi^{-1}(z_t)) \mid \mathcal{F}_t^y]\,\tilde{E}[\Lambda_t \mid \mathcal{F}_t^y]}{\sigma_\Psi^{\tilde{\mathcal{P}}}\,\sigma_\Lambda^{\mathcal{P}}} \tag{3.6.4}$$

$$P_{\Psi\Lambda^{-1}}^{\mathcal{P}} \triangleq \frac{E[\Psi(x_t)\Lambda_t^{-1} \mid \mathcal{F}_t^y] - E[\Psi(x_t) \mid \mathcal{F}_t^y]\,E[\Lambda_t^{-1} \mid \mathcal{F}_t^y]}{\sigma_\Psi^{\mathcal{P}}\,\sigma_{\Lambda^{-1}}^{\mathcal{P}}} \tag{3.6.5}$$

where superscript $\mathcal{P}$, $\tilde{\mathcal{P}}$ denotes the probability measure and $\sigma_{(\cdot)}^{\tilde{\mathcal{P}}}$, $\sigma_{(\cdot)}^{\mathcal{P}}$ denotes the conditional standard deviation of process (•) under probability measure $\tilde{\mathcal{P}}$, $\mathcal{P}$, respectively.

Substituting (3.6.4), (3.6.5) into (3.6.2), (3.6.3) respectively, we obtain

$$E\,[\Psi(x_t) \mid \mathcal{F}_t^y] = \tilde{E}\,[\Psi(\Phi^{-1}(z_t)) \mid \mathcal{F}_t^y] + \frac{P_{\Psi\Lambda}^{\tilde{\mathcal{P}}}\ \sigma_\Psi^{\tilde{\mathcal{P}}}\ \sigma_\Lambda^{\tilde{\mathcal{P}}}}{\tilde{E}[\Lambda_t \mid \mathcal{F}_t^y]} \tag{3.6.6}$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$E[\Psi(x_t) \mid \mathcal{F}_t^y] = \tilde{E}[\Psi(\Phi^{-1}(z_t)) \mid \mathcal{F}_t^y] - \frac{P^{\rho}_{\Psi\Lambda}\,\sigma^{\rho}_{\Psi}\,\sigma^{\rho}_{\Lambda^{-1}}}{E[\Lambda_t^{-1} \mid \mathcal{F}_t^y]}. \tag{3.6.7}$$

Defining $k_t^{1\tilde{\rho}} \triangleq \dfrac{\sigma^{\tilde{\rho}}_{\Psi}\,\sigma^{\tilde{\rho}}_{\Lambda}}{\tilde{E}[\Lambda_t \mid \mathcal{F}_t^y]}$, $k_t^{2\rho} \triangleq \dfrac{\sigma^{\rho}_{\Psi}\,\sigma^{\rho}_{\Lambda^{-1}}}{E[\Lambda_t^{-1} \mid \mathcal{F}_t^y]}$ we have the following proposition.

<u>Proposition 3.6.1</u> Using the property $|P^{(\cdot)}_{(\cdot)}| \leq 1$ we obtain the following lower and upper bounds:

$$\text{for } P^{\tilde{\rho}}_{\Psi\Lambda} \leq 1 \;:\; E[\Psi(x_t) \mid \mathcal{F}_t^y] \leq \tilde{E}[\Psi(\Phi^{-1}(z_t))] + k_t^{1\tilde{\rho}} \tag{3.6.8}$$

$$\text{for } P^{\tilde{\rho}}_{\Psi\Lambda} \geq -1 \;:\; E[\Psi(x_t) \mid \mathcal{F}_t^y] \geq \tilde{E}[\Psi(\Phi^{-1}(z_t))] - k_t^{1\tilde{\rho}} \tag{3.6.9}$$

$$\text{for } P^{\rho}_{\Psi\Lambda^{-1}} \leq 1 \;:\; E[\Psi(x_t) \mid \mathcal{F}_t^y] \geq \tilde{E}[\Psi(\Phi^{-1}(z_t))] - k_t^{2\rho} \tag{3.6.10}$$

$$\text{for } P^{\rho}_{\Psi\Lambda^{-1}} \geq -1 \;:\; E[\Psi(x_t) \mid \mathcal{F}_t^y] \leq \tilde{E}[\Psi(\Phi^{-1}(z_t))] + k_t^{2\rho}. \tag{3.6.11}$$

<u>Proof</u>: From the equivalence of measures ρ, $\tilde{\rho}$ the exponential processes Λ_t, Λ_t^{-1} are positive martingales a.s., which implies $\tilde{E}(\Lambda_t \mid \mathcal{F}_t^y) > 0$, a.s. $E(\Lambda_t^{-1} \mid \mathcal{F}_t^y) > 0$ a.s., so $k_t^{1\tilde{\rho}} \geq 0, k_t^{2\rho} \geq 0$ a.s.. Recall also that $\tilde{E}[\Psi(\Phi^{-1}(z_t) \mid \mathcal{F}_t^y)] = \tilde{E}[\Psi(\Phi^{-1}(z_t))]$, by the independence of $\mathcal{F}_t^y = \sigma\{y_s, 0 \leq s \leq t\}$ and $\{z_t, 0 \leq s \leq t\}$ under measure $\tilde{\rho}$. The proof then follows immediately. QED

Next, we shall concentrate on obtaining upper bounds for $k_t^{1\tilde{\rho}}$ and $k_t^{1\rho}$, which will then be used in Proposition 3.6.1.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Bound for $k_t^{1\tilde{\wp}}$

We shall find a lower bound $\tilde{E}[\Lambda_t|\mathcal{F}_t^y] \geq (k_t^1)_{min}$ and an upper bound $\sigma_\Lambda^{\tilde{\wp}} \leq (k_t^1)_{max}$ using the fact that under measure $\tilde{\wp}$ the signal process z_t satisfies the linear stochastic differential equation given by (3.4.9).

By the convexity of the exponential function Λ_t, it follows from Jensen's inequality that

$$\tilde{E}[\Lambda_t|\mathcal{F}_t^y] \geq e^{\tilde{E}\{\int_0^t (L_f^n h(\Phi^{-1}(z_s)) - \sum_{j=1}^n c_j z_{js}) dw'_s - \frac{1}{2}\int_0^t |L_f^n h(\Phi^{-1}(z_s)) ds - \int_0^t \sum_{j=1}^n c_j z_{js}|^2 ds | \mathcal{F}_t^y\}}$$

$$\times e^{\tilde{E}\{\int_0^t z_{1s} dy_s - \frac{1}{2}\int_0^t |z_{1s}|^2 ds | \mathcal{F}_t^y\}} \quad . \tag{3.6.12}$$

By Remark 2.2.2 (ii) (b), under measure $\tilde{\wp}$, $\{z_s,\ 0 \leq s \leq t\}$ and $\{w'_s,\ 0 \leq s \leq t\}$ are independent of $\{y_s,\ 0 \leq s \leq t\}$, so

$$\tilde{E}\{-\frac{1}{2}\int_0^t [|L_f^n h(\Phi^{-1}(z_s)) - \sum_{j=1}^n c_j z_{js}|^2 + |z_{1s}|^2] ds | \mathcal{F}_t^y\} =$$
$$-\frac{1}{2}\int_0^t \tilde{E}[|L_f^n h(\Phi^{-1}(z_s)) - \sum_{j=1}^n c_j z_{js}|^2 + |z_{1s}|^2] ds.$$

Also,

$$\tilde{E}\{\int_0^t [L_f h(\Phi^{-1}(z_s)) - \sum_{j=1}^n c_j z_{js}] dw'_s | \mathcal{F}_t^y\} = 0.$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Next, since

$$\tilde{P}\left(\int_0^T |z_{1t}|^2 dt < \infty\right) = 1 \text{ a.s.}, \quad \tilde{E}\,|z_{1s}| < \infty,\ 0 \le s \le T$$

and

$$\tilde{P}\left(\int_0^T |\tilde{E}(|z_{1s}|\,|\mathcal{F}_s^y)|^2 ds < \infty\right) = 1 \text{ a.s.},$$

then for all t in the interval $0 \le t \le T$,

$$\tilde{E}\left[\int_0^t (z_{1s} dy_s)\,|\,\mathcal{F}_t^y\right] = \int_0^t \tilde{E}\,\{z_{1s}\,|\,\mathcal{F}_s^y\} dy_s = \int_0^t \tilde{E}(z_{1s}) dy_s,$$

(see Liptser and Shiryayev [103, Thm. 5.14, pp. 185-186]).

Therefore the lower bound $(k^1_t)_{min}$ is given by

$$\tilde{E}[\Lambda_t\,|\,\mathcal{F}_t^y] \ge (k_t^1)_{min} = e^{-\frac{1}{2}\int_0^t \tilde{E}\{\,|L_f^n h(\Phi^{-1}(z_s)) - \sum_{j=1}^n c_j z_{js}|^2 + |z_{1s}|^2 ds\} ds} \times e^{\int_0^t \tilde{E}(z_{1s}) dy_s} \qquad (3.6.13)$$

We shall now determine an upper bound for $\sigma_\Lambda^{\tilde{P}}$. By definition,

$$\sigma_\Lambda^{\tilde{P}} = \{\tilde{E}[\Lambda_t^2\,|\,\mathcal{F}_t^y] - \tilde{E}^2[\Lambda_t\,|\,\mathcal{F}_t^y]\}^{1/2}$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

thus, using (3.6.13),

$$\sigma_\Lambda^{\tilde{P}} \leq \{\tilde{E}[\Lambda_t^2 \mid \mathscr{F}_t^y] - [(k_t^1)_{min}]^2\}^{1/2}. \tag{3.6.14}$$

From (3.6.1) we have

$$\tilde{E}[\Lambda_t^2 \mid \mathscr{F}_t^y] = \tilde{E}\{e^{2\int_0^t [L_f^n h(\Phi^{-1}(z_s)) - \sum_{j=1}^n c_j z_{js}]dw'_s}$$

$$\times e^{-\int_0^t [\,|L_f^n h(\Phi^{-1}(z_s)) - \sum_{j=1}^n c_j z_{js}|^2]ds + 2\int_0^t z_{1s}dy_s - \int_0^t |z_{1s}|^2 ds} \mid \mathscr{F}_t^y\}.$$

By the Schwartz inequality,

$$\tilde{E}[\Lambda_t^2 \mid \mathscr{F}_t^y] \leq \{\tilde{E}[e^{4\int_0^t [L_f^n h(\Phi^{-1}(z_s)) - \sum_{j=1}^n c_j z_{js}]dw'_s} \mid \mathscr{F}_t^y]\}^{1/2}$$

$$\times \{\tilde{E}[e^{-2\int_0^t |L_f^n h(\Phi^{-1}(z_s)) - \sum_{j=1}^n c_j z_{js}|^2 ds} \times e^{4\int_0^t z_{1s}dy_s - 2\int_0^t z_{1s}^2 ds} \mid \mathscr{F}_t^y]\}^{1/2}.$$

Again, using the $\tilde{P}$ independence of $\{z_s, 0 \leq s \leq t\}$ and $\{w'_s, 0 \leq s \leq t\}$, then

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$\tilde{E}\left[e^{4\int_0^t [L_f^n h(\Phi^{-1}(z_s)) - \sum_{j=1}^n c_j z_{js}]dw'_s} \,\middle|\, \mathcal{F}_t^y\right] = \tilde{E}\left[e^{4\int_0^t [L_f^n h(\Phi^{-1}(z_s)) - \sum_{j=1}^n c_j z_{js}]dw'_s}\right]$$

Furthermore, since $\tilde{P}(\int_0^T |L_f^n h(\Phi^{-1}(z_t))|^2 dt < \infty) = 1$ a.s., by Gihman and Skorohod [65, Lemma 2 pp. 86-87] (which is an application of Schwartz inequality and the fact that the expected value of an exponential supermartingale is less than one) the previous expression is upper bounded by

$$\tilde{E}\left[e^{\int_0^t |L_f^n h(\Phi^{-1}(z_s)) - \sum_{j=1}^n c_j z_{js}|dw'_s}\right] \leq \{\tilde{E}\left[e^{32t\int_0^t |L_f^n h(\Phi^{-1}(z_s))|^2 + 32t|\sum_{j=1}^n c_j z_{js}|^2 ds}\right]\}^{1/2}.$$

One can also find a Feynman-Kac PDE which would be the solution to the left side of the previous expression without having to use the bounding result of Gihman and Skorohod [65,Lemma 2,pp.86-87].

Thus, since $\tilde{E}e^{-2\int_0^t |L_f^n h(\Phi^{-1}(z_s)) - \sum_{j=1}^n c_j z_{js}|^2 ds} \leq 1$, we obtain

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$\tilde{E}[\Lambda_t^2 | \mathcal{F}_t^y] \leq \{\tilde{E}[e^{32t\int_0^t [\,|L_f^n h(\Phi^{-1}(z_s))|^2 + |\sum_{j=1}^n c_j z_{js}|^2] ds}]\}^{1/4} \tag{3.6.15}$$
$$\times \{\tilde{E}[e^{4\int_0^t z_{1s} dy_s - 2\int_0^t z_{1s}^2 ds} | \mathcal{F}_t^y]\}^{1/2}.$$

The unconditional expectation on the right side of (3.6.15) can be further upper bounded by using the property that $L_f^n h(\Phi^{-1}(z))$ satifies a linear growth condition which is due to the linear growth assumption on f, h. For the exact value of this unconditional expectation one has to solve a Feynman-Kac formula where the only nonlinearity is due to $L_f^n h(\Phi^{-1}(z))$.

Finally, the conditional expectation on the right side of (3.6.15) can be evaluated in a similar fashion as in Section 3.3.5 by solving a Feynman-Kac type PDE where

$$e^{4y_t z_1} q(z,t) \triangleq \tilde{E}[e^{4\int_0^t z_{1s} dy_s - 2\int_0^t z_{1s}^2 ds} | \mathcal{F}_t^y] \tag{3.6.16}$$

and satisfies the PDE

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$\frac{\partial}{\partial t}q(z,t) = L(t)^{*}q(z,t) - Q(t,z)q(z,t) \quad , \quad 0 \le t \le T$$
$$q(z,0) = p_0(z) \tag{3.6.17}$$

with $L(t)$ the Kolmogorov's operator associated with the diffusion process z_t given by (3.4.9) and $Q(t,z) \triangleq 4y_t z_2 + 2z_{1t}^2$.

Remark 3.6.1 The solution of the PDE (3.4.17) gives an unnormalized conditional density of the Gaussian type. It is of Gaussian type because $Q(t,z)$ is a quadratic function in the coordinate z and the forward operator $L(t)$ corresponding to the process z_t given by (3.4.9) satisfies a linear stochastic differential equation. The procedure for solving (3.4.18) is exactly the same as the one given in Appendix 7.G.

Summarizing, we get the following.

Lemma 3.6.1 Let (3.1.1), (3.1.2) be the filtering problem under consideration. Let f, h satisfy a linear growth condition, f satisfy a Lipschitz condition, σ bounded, f, σ, h of C^∞ class, and assume that the necessary and sufficient conditions of Theorem 3.3.1 are satisfied with a global diffeomorphism $\Phi = (h(x)\ ,\ ...,\ L_f^{n-1}(h(x))^T$. For any integrable function $\Psi(x)$ we have the following lower and upper bound:

$$\tilde{E}\,[\Psi(\Phi^{-1}(z_t))] - \sigma_\Psi^{\tilde{p}}\,\frac{(k_t^1)_{max}}{(k_t^1)_{min}} \le E\,[\Psi(x_t)\,|\,\mathcal{F}_t^y] \le \tilde{E}\,[\Psi(\Phi^{-1}(z_t))] + \sigma_\Psi^{\tilde{p}}\,\frac{(k_t^1)_{max}}{(k_t^1)_{min}}$$

where

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$(k_t^1)_{max} = \{ (\tilde{E}[e^{32t\int_0^t F(z,t)dt}])^{1/4} \times [e^{4z_1 y_t} \int_{R^n} q(z,t)dz]^{1/2} - (k_t^1)^2_{min} \}^{1/2}$$

$$F(z,t) = |L_f^n h(\Phi^{-1}(z))|^2 + |\sum_{j=1}^{n} c_j z_{jt}|^2$$

and $(k^1_t)_{min}$ is given by (3.6.13).

Proof: As above. QED

If we try to follow the same procedure to obtain an appropriate upper bound for $k_t^{1\rho}$ using (3.6.10), (3.6.11), we perform expectations with respect to measure ρ. This, however, becomes a very difficult task since under measure ρ we can no longer use the independence of $\mathcal{F}_t^y$ and $\{z_s, 0 \le s \le t\}$, so we cannot repeat the above procedure.

3.7 RELATION TO PREVIOUS WORK

In previous work, Snyder and Rhodes [121] and Bobrovsky and Zakai [23,24] derived lower bounds for the nonlinear filtering problem. More specifically, Snyder and Rhodes presented a lower bound in estimating Gaussian processes from nonlinear observations while Bobrovsky and Zakai gave a lower bound on the filtering error of certain nondegenerate Markov processes. The above bounds are based on the Van Trees version of the Cramér-Rao bound [128], in particular $\bar{\varepsilon}^2 \ge J_T^{-1}$ where $\bar{\varepsilon}^2$ is the mean-square estimation error of the process x_T and J_T is the Fisher information matrix.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Here we shall extend the filtering problems that admit the lower bound given by Bobrovsky and Zakai to Markov processes which satisfy degenerate stochastic differential equations (we say that a stochastic differential system is degenerate if not all components of the state process are directly affected by a noisy input). The lower bound given by Bobrovsky and Zakai is the filtering error of some suitable Gaussian system chosen in such a way that

$$J_T^{ij} = E\{\frac{\partial^2}{\partial x_i \, \partial x_j} \ln \Lambda_T\}$$

is a nonnegative definite matrix J; where $\Lambda_T \triangleq \frac{dP}{dP_g}$ is the Radon-Nikodym derivative of the measure P induced by the nonlinear system under consideration with respect to the measure P_g induced by a Gaussian system. However, if the dimension of the state noise w_t differs from that of x_t, in general, J_T would not be nonnegative definite, thus a lower bound on the estimation error cannot be found. This implication is avoided if the necessary and sufficient conditions of Theorem 3.2.1 or Theorem 3.3.1 are satisfied. The following theorem, then, is an extension of Bobrovsky and Zakai [23,24] for the case of degenerate Markov processes.

<u>Theorem 3.7.1</u>

Suppose x_t, y_t and $\tilde{z}_t, \tilde{y}_t$ satisfy the following stochastic equations on probability spaces $(\Omega, \mathcal{F}_t, P)$, $(\tilde{\Omega}, \tilde{\mathcal{F}}_t, \tilde{P})$, respectively:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$(\Omega, \mathcal{F}_t, P)\ \Sigma_1: \begin{cases} dx_t = f(x_t)dt+\sigma(x_t)dw_t \\ dy_t = h(x_t)dt+dbt \end{cases} ; (\tilde{\Omega}, \tilde{\mathcal{F}}_t, \tilde{P})\ \Sigma_2: \begin{cases} d\tilde{z}_t = A_t\tilde{z}_t\, dt+Bd\tilde{w}_t \\ d\tilde{y}_t = C_t\ \tilde{z}_t dt+db_t \end{cases}$$

with

(i) w_t, b_t of dimension one, $x_t \in R^n$, and
(ii) initial densities $p_0(x)=p_0(\tilde{z})$.

If the conditions of Theorem 3.2.1 or Theorem 3.3.1 are satisfied, then

$$\bar{\varepsilon}_i^2 = E\left[\phi_i(x_t) - E(\phi_i(x_t) \mid \mathcal{F}_t^y)\right]^2 \geq \tilde{E}\left[\tilde{z}_{it} - E(\tilde{z}_{it} \mid \mathcal{F}_t^y)\right]^2 \tag{3.7.1}$$

where

$$A_t = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & & 0 \\ \vdots & & & \ddots & \vdots \\ & & & & 1 \\ \tilde{c}_1 & \tilde{c}_2 & \dots & & \tilde{c}_n \end{bmatrix}, \quad \tilde{c}_i = E\left[\frac{\partial}{\partial z_i} L_f^n\, h(\Phi^{-1}(z))\right], \ i = 1, 2, \dots, n \tag{3.7.2}$$

$$C_t^T C_t = E\{[\frac{\partial}{\partial z}(L_f^n h(\Phi^{-1}(z)))]^T \frac{\partial}{\partial z} L_f^n h(\Phi^{-1}(z)) - \tilde{C}_t^T \tilde{C}_t + \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} [1\ 0\ \dots\ 0] \tag{3.7.3}$$

with ϕ_i the i-th component of the vector

$$\Phi(x) = (h(x), L_f\, h(x), \dots, L_f^{n-1}\, h(x))^T.$$

<u>Proof:</u> Assume conditions of Theorem 3.3.1 are satisfied, then by defining

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$\frac{d\wp}{d\tilde{\wp}} = \tilde{E}\,[\Lambda_T \mid \tilde{\mathcal{F}}_T] = e^{\int_0^T (L_f^n h(\Phi^{-1}(z_s)) - \tilde{c}_s z_s)\, d\tilde{z}_{ns}}$$

$$\times\, e^{-\frac{1}{2}\int_0^T (|L_f^n h(\Phi^{-1}(z_s))|^2 - |\tilde{C}_s z_s|^2) ds}$$

$$\times\, e^{\int_0^T (h(\Phi^{-1}(z_s)) - \tilde{C}_s \tilde{z}_s) dy_s - \frac{1}{2}\int_0^T (|h(\Phi^{-1}(z_s))|^2 - |C_s \tilde{z}_s|^2) ds}$$

we have $\wp << \tilde{\wp}$ where $\tilde{z}_t = \Phi(x_t)$. Following the proof given by Bobrovsky an Zakai [24, 25], if (3.7.2), (3.7.3) are satisfied then J_T becomes a nonnegative definite matrix, thus the bound (3.7.1) is valid. The above bound remains valid if instead of Theorem 3.3.1, Theorem 3.2.1 is considered. For this case, however, t should be replaced by τ. QED

We conclude this chapter by pointing out to the reader that when local linearization is under consideration the need for solving (3.5.5) is required. However, even thought we were not able to provide the solution of this boundary value problem we can still apply the bounding technique of Section 2.5.1 to obtain an upper bound for u(z, s).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

CHAPTER 4

OPTIMAL CONTROL OF PARTIALLY OBSERVED DIFFUSIONS

4.1 PRECISE PROBLEM STATEMENT

We study the stochastic partially observed control problem stated in Section 1.2. Consider the stochastic system.

$$dx_t = f(t, x_t, u_t)dt + \sigma(t, x_t, u_t)dw_t \tag{4.1.1}$$

$$dy_t = h(t,x_t)dt + g(t,x_t)dw_t + \tilde{g}(t)db_t. \tag{4.1.2}$$

where, assuming w., b. are uncorrelated for all t, the quadratic covariation between $\{x_s, 0 \le s \le t\}$ and $\{y_s, 0 \le s \le t\}$ is

$$\langle x,y\rangle_t = \int_0^t \sigma(s,x_s)g^T(s,x_s)ds. \tag{4.1.3}$$

The problem is then to derive the necessary conditions for

$$\min \{J(u);\ u\in U_{ad}\},$$
$$J(u) = E^u\{\int_0^T \text{II}(t,x_t,u_t)dt + \kappa(x_T)\} \tag{4.1.4}$$

subject to constraints (4.1.1), (4.1.2) with controls u taking values in a non empty compact subset of R^k.

Throughout this chapter we shall assume the following assumptions:

(A′1) $f:[0, T]xR^nxU\to R^n$, Borel measurable, bounded, continuous C^∞ in x with bounded first derivatives in x, u;

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

(A′2) $\sigma:[0, T]xR^n xU \to R^n \otimes R^m$, Borel measurable, bounded, continuous C^∞ in x with bounded first derivatives in x, u and bounded second derivatives in u;

(A′3) $h:[0, T]xR^n \to R^d$, Borel measurable, bounded, continuous C^∞ in x with bounded first derivatives in x;

(A′4) $g:[0, T]xR^n \to R^d \otimes R^m$, Borel measurable, bounded, continuous C^∞ in x;

(A′5) $\tilde{g}:[0, T] \to R^d xR^d$, Borel measurable, continuous;

(A′6) $\pi:[0, T]xR^n xU \to R^1$, Borel measurable, bounded, continuous C^∞in x, C^1 in u and

$$|\frac{\partial \pi}{\partial u}(t,x,u)| \le \gamma(x), \quad \gamma(x) \in L^2(R^n);$$

(A′7) $\kappa:R^n \to R^1$, Borel measurable, bounded, and $\kappa(x) \in L^2(R^n)$;

(A′8) there exist β_1, $\beta_2 > 0$ such that

$$g(t,x)\, g^T(t,x) + \tilde{g}(t)\, \tilde{g}^T(t) \ge \beta I_d$$

$$\tilde{g}(t)\, \tilde{g}^T(t) \ge \beta_2\, I_d$$

where I_d denotes the dxd identity matrix;

(A′9) $p_0(x) \in L^2(R^n)$, the density of x_0;

(A′10) the admissible control set U_{ad} is a non empty compact convex, subset of R^k such that u (t, .) consists of $\mathcal{F}_t^y \triangleq \sigma(y_s, 0 \le s \le t)$-adapted processes on $[0, T]xC([0, T]; R^m)$ with values in U and

$$u(t,y_t) \in L_y^2([0,T]xC([0,T];R^d)).$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

As in Section 2.1, consider the space $\Omega = R^n \otimes C([0, T]; R^m) \otimes C([0, T]; R^d)$ with coordinate functions (x,w,y). In this case, however, the problem is more difficult since the independence of $w.$, $y.$, is violated. Thus, the measure $\tilde{P}_1$ on Ω is no longer a Wiener measure, hence

$$\begin{aligned} \mathcal{F}_T &= B_T^n \otimes B_T^m \otimes B_T^d \\ \Omega &= R^n \otimes C([0,T]; R^m) \otimes C([0,T]; R^d) \\ \tilde{P}_1 &= P_1(dx,dw,dy) \end{aligned} \tag{4.1.5}$$

where B_T^k is the Borel σ-algebra on $C([0,T]; R^k)$.

Pardoux [114] considers the filtering problem (4.1.1), (4.1.2) (i.e., $u = 0$) when σ is an nxn matrix, g is independent of x, and

$$g(t)g^T(t) + \tilde{g}(t)\tilde{g}^T(t) = I_d .$$

Here, we shall first prove that even in the general case above (4.1.1), (4.1.2) has weak solutions, and later determine the equation satisfied by the unnormalized conditional density of $\{x_t,\ t \in [0,T]\}$.

Lemma 4.1.1

The stochastic system of equations (4.1.1), (4.1.2) has solutions $(x_s^u,\ y_s,\ 0 \le s \le t)$ which are unique in probability law; i.e., has weak solutions.

Proof: Consider the system

$$dx_t = f(t,x_t,u_t)dt - \sigma(t,x_t,u_t)g^T(t,x_t)\, k(t,x_t))dt + \sigma(t,x_t,u_t)dw_t \tag{4.1.6}$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$dy_t = g(t,x_t)dw_t + \tilde{g}(t)db_t \tag{4.1.7}$$

for some initial condition $x_0 \in R^n$ and control variable $u \in U_{ad}$ where $k(t, x_t)$ is a bounded Borel measurable function to be defined shortly. Due to the bounded assumptions imposed earlier (which imply Lipschitz conditions) there exists a unique strong solution $\{x_s^u, y_s, 0 \le s \le t\}$ on the probability space $(\Omega, \mathscr{F}_t, \tilde{\mathcal{P}}_1)$.

Next, consider the martingale defined as

$$m_t \triangleq \int_0^t k^T(t,x_s)(g(s,x_s)dw_s + \tilde{g}(t)db_s) \in M(\mathscr{F}_T, \tilde{\mathcal{P}}_1) \tag{4.1.8}$$

and using the exponential formula, define $\mathcal{P}^u$ by

$$\Lambda_T \triangleq \tilde{E}[\frac{d\mathcal{P}^u}{d\tilde{\mathcal{P}}_1}|\mathscr{F}_T] = e^{m_T - \frac{1}{2}\langle m, m\rangle_T}.$$

Thus,

$$\Lambda_T = e^{\int_0^T k^T(t,x_s)(g(s,x_s)dw_s + \tilde{g}(s)db_s) - \frac{1}{2}\int_0^T k^T(s,x_s)(g(s,x_s)g^T(s,x_s) + \tilde{g}(s)\tilde{g}^T(s))k(s,x_s)ds} \tag{4.1.9}$$

where $k(t,x_t)$ is defined by

$$k(t,x_t) \triangleq (g(t,x_t)g^T(t,x_t) + \tilde{g}(t)\tilde{g}^T(t))^{-1}h(t,x_t). \tag{4.1.10}$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

By assumptions (A'3), (A'8), $k(t,x_t)$ is a well-defined bounded function; therefore, by Theorem 7.E.2, it follows that $\tilde{E}[\Lambda_T] = 1$, thus P^u is absolutely continuous with respect to $\tilde{P}_1$ ($P^u << \tilde{P}_1$). By the translation theorem (Theorem 7.E.3) it follows that

$$dm_t^1 = dw_t - d\langle w.,m.\rangle_t = dw_t - g^T(t,x_t)k(t,x_t)dt \in M(\mathcal{F}_t, P^u) \tag{4.1.11a}$$

$$dm_t^2 = db_t - d\langle b.,m.\rangle_t = db_t - \tilde{g}^T(t,x_t)k(t,x_t)dt \in M(\mathcal{F}_t, P^u). \tag{4.1.11b}$$

Moreover, m_t^1, m_t^2 satisfy $\langle m.^1, m.^1\rangle_t = I_m t$, $\langle m.^2, m.^2\rangle_t = I_d t$, thus m_t^1, m_t^2 are standard Brownian motions. Substituting (4.1.11) into (4.1.6), (4.1.7) we obtain the stochastic differential equation (4.1.1) and observation equation

$$dy_t = g(t,x_t)dm_t^1 + (g(t,x_t)g^T(t,x_t)k(t,x_t) + \tilde{g}(t)\tilde{g}^T(t)k(t,x_t))dt + \tilde{g}(t)dm_t^2$$

which by (4.1.10) is equivalent to (4.1.2). Thus, we have constructed the solution $\{x_s^u, y_s, 0 \le s \le t\}$ to (4.1.1), (4.1.2) on the probability space $(\Omega, \mathcal{F}_t, P^u)$ to be a weak solution which is unique in probability law. QED

The differential of the LR Λ_t is given by

$$d\Lambda_t = \Lambda_t k^T(t,x_t)(g(t,x_t)dw_t + \tilde{g}(t)db_s). \tag{4.1.12}$$

Next, we shall derive the equation satisfied by the unnormalized conditional density ρ_t of the state process by modifying the probability space $(\Omega, \mathcal{F}_t, \tilde{P}_1)$ to be

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

equal to the product probability space $(\tilde{\Omega}\otimes\hat{\Omega},\ \tilde{\mathcal{F}}_t\otimes\hat{\mathcal{F}}_t,\ \tilde{P}\otimes\hat{P})$ as given in Chapter 2 for the case when $g(t,x_t) = 0$.

Theorem 4.1.1

For any bounded, twice continuously differentiable function $\phi(\cdot)$, the unnormalized conditional probability density $\rho_t \triangleq \rho(x,t)$, where $\rho_t(\phi) \triangleq \int \phi(z)\rho_t(dz)$, satisfies the equation

$$\begin{aligned}\rho_t(\phi)(x) &= \phi(x) + \int_s^t \rho(A^u(r))\phi)(x)dr + \sum_{k=1}^{d}\int_s^t \rho(M_k(r)\phi)(x)d\tilde{y}_t^k \\ \lim_{t\downarrow 0}\rho_t(\phi)(x) &= \phi(x)\end{aligned} \qquad (4.1.13)$$

where

$$A^u(t) \triangleq \sum_{i=1}^{n} f(t,x,u)\frac{\partial}{\partial x^i} + \frac{1}{2}\sum_{i,j}^{n} a(t,x,u)^{ij}\frac{\partial^2}{\partial x^i\partial x^j},\quad \tilde{k}(t,x_t) = g(t,x_t)g^T(t,x_t) + \tilde{g}(t)g^T(t)$$

$$\tilde{M}_k^u(t) \triangleq \sum_{i=1}^{d}(\tilde{k}^{-\frac{1}{2}}(t,x))^{ik}h_i(t,x) + \sum_{i=1}^{n}(\sigma(t,x,u)g^T(t,x)\tilde{k}^{-\frac{1}{2}}(t,x))^{ik}\frac{\partial}{\partial x^i},$$

$(\cdot)^{ij}$ is the (i,j)-th component of a matrix, $\tilde{y}_t$ is a standard Brownian motion and $\tilde{k}(t,x_t)$ is independent of x.

Proof: Consider as starting point the system (4.1.6), (4.1.7) under the measure $\tilde{P}_1$ and define the new process

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$\tilde{w}_t = \int_0^t \tilde{D}(s,x_s)(dw_s - C(s,x_s)dy_s) \qquad (4.1.14)$$

where $\tilde{w}_t \in R^m$, $\tilde{D}(t,x) \in R^m \otimes R^m$ and $C(t,x) \in R^m \otimes R^d$. We shall determine $\tilde{D}(t,x)$, $C(t,x)$ so that the quadratic covariation process $\langle \tilde{w}., y.\rangle_t = 0$, i.e., $\tilde{w}_t, y_t$ are orthogonal martingale processes. From (4.1.7) we have

$$d\langle \tilde{w}.,y.\rangle_t = d\langle \int_0^{\cdot} \tilde{D}(s,x_s)(dw_t - C(s,x_s)g(s,x_s)dw - C(s,x_s)\tilde{g}(s)db_s), \int_0^{\cdot} (g(s,x_s)dw_s + \tilde{g}(s)db_s)\rangle_t$$

$$= \tilde{D}(t,x_t)(g^T(t,x_t) - C(t,x_t)g(t,x_t)g^T(t,x_t) - C(t,x_t)\tilde{g}(t)\tilde{g}^T(t))$$

$$= \tilde{D}(t,x_t)(g^T(t,x_t) - C(t,x_t)[g(t,x_t)g^T(t,x_t) + \tilde{g}(t)\tilde{g}^T(t)]).$$

If we define

$$C(t,x_t) \triangleq g^T(t,x_t)[g(t,x_t)g^T(t,x_t) + \tilde{g}(t)\tilde{g}^T(t)]^{-1} \qquad (4.1.15)$$

then $\langle \tilde{w}., y.\rangle_t = 0$, thus $\tilde{w}_t, y_t$ are orthogonal for all t. Next, we shall choose $\tilde{D}(t,x_t)$ so that the new process $\tilde{w}_t$ is a standard Wiener process. Consider the quadratic variation process

$$d\langle \tilde{w}.,\tilde{w}.\rangle_t = d\langle \int_0^{\cdot} \tilde{D}(s,x_s)(dw_s - C(s,x_s)dy_s), \int_0^{\cdot} \tilde{D}(s,x_s)(dw_s - C(s,x_s))dy_s\rangle_t$$

$$= \tilde{D}(t,x_t)[I_m - g^T(t,x_t)C^T(t,x_t) - C(t,x_t)g(t,x_t) + C(t,x_t)g(t,x_t)g^T(t,x_t)C^T(t,x_t)$$

$$+ C(t,x_t)\tilde{g}(t)\tilde{g}(t)^T(t)C^T(t,x_t)]\tilde{D}^T(t,x_t)$$

$$= \tilde{D}(t,x_t)I_m\tilde{D}^T(t,x_t) + \tilde{D}(t,x_t)C(t,x_t)[g(t,x_t)g^T(t,x_t) + \tilde{g}(t)\tilde{g}^T(t)]C^T(t,x_t)\tilde{D}^T(t,x_t)$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$- \tilde{D}(t,x_t)g^T(t,x_t)[g(t,x_t)g^T(t,x_t) + \tilde{g}(t)\tilde{g}(t)]^{-T}g(t,x_t)\tilde{D}^T(t,x_t)$$

$$- \tilde{D}(t,x_t)g^T(t,x_t)[g(t,x_t)g^T(t,x_t) + \tilde{g}(t)\tilde{g}^T(t)]^{-1}g(t,x_t)\tilde{D}^T(t,x_t)$$

$$= \tilde{D}(t,x_t)I_m\tilde{D}^T(t,x_t) - \tilde{D}(t,x_t)g^T(t,x_t)[g(t,x_t)g^T(t,x_t) + \tilde{g}(t)\tilde{g}^T(t)]^{-T}g(t,x_t)\tilde{D}^T(t,x_t)$$

$$= \tilde{D}(t,x_t)\{I_m - g^T(t,x_t)[g(t,x_t)g^T(t,x_t) + \tilde{g}(t)\tilde{g}^T(t)]^{-1}g(t,x_t)\}\tilde{D}^T(t,x_t).$$

If we define

$$\tilde{D}(t,x_t) \triangleq I_m - \{g^T(t,x_t)[g(t,x_t)g^T(t,x_t) + \tilde{g}(t)\tilde{g}^T(t)]^{-1}g(t,x_t)\}^{-1/2} \tag{4.1.16}$$

then $\tilde{w}_t$ is a standard Wiener process. Thus we have constructed a process $\tilde{w}_t$ which is orthogonal to of y_t. Notice that from assumption (A′8,) $\tilde{D}(t,x_t)$ is well defined and positive definite. From (4.1.14) we solve for w_t and obtain

$$dw_t = \tilde{D}^{-1}(t,x_t)d\tilde{w}_t + C(t,x_t)dy_t. \tag{4.1.17}$$

Now, substituting (4.1.17) into (4.1.6) we obtain

$$\begin{aligned} dx_t = {} & [f(t,x_t,u_t) - \sigma(t,x_t,u_t)g^T(t,x_t)k(t,x_t)]dt \\ & + \sigma(t,x_tu_t)\tilde{D}^{-1}(t,x_t)d\tilde{w}_t + \sigma(t,x_t,u_t)C(t,x_t)dy_t. \end{aligned} \tag{4.1.18}$$

Let us next define a new process $\tilde{y}_t$ by

$$d\tilde{y}_t \triangleq \tilde{C}(t,x_t)dy_t = \tilde{C}(t,x_t)(g(t,x_t)dw_t + \tilde{g}(t)db_t) \tag{4.1.19}$$

where $\tilde{C}(t,x_t) \in R^d \otimes R^d$. The quadratic variation of the new process $\tilde{y}_t$ is

$$d\langle\tilde{y},\tilde{y}\rangle_t = \tilde{C}(t,x_t)(g(t,x_t)g^T(t,x_t) + \tilde{g}(t)\tilde{g}^T(t))\tilde{C}^T(t,x_t)dt.$$

If we define

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$\tilde{C}(t,x_t) \triangleq (g(t,x_t)g^T(t,x_t) + \tilde{g}(t)\tilde{g}(t))^{-1/2} \tag{4.1.20}$$

we have

$$d\langle\tilde{y},\tilde{y}\rangle_t = I_m\, dt$$

so $\tilde{y}_t$ is a standard Wiener process. Next, substituting (4.1.19) into (4.1.18) we deduce

$$\begin{aligned} dx_t = {} & [f(t,x_t,u_t) - \sigma(t,x_t,u_t)g^T(t,x_t)k(t,x_t)]dt \\ & + \sigma(t,x_t,u_t)\tilde{D}^{-1}(t,x_t)d\tilde{w}_t + \sigma(t,x_t,u_t)C(t,x_t)\tilde{C}^{-1}(t,x_t)d\tilde{y}_t. \end{aligned} \tag{4.1.21}$$

From (4.1.15), (4.1.20) we also have

$$C(t,x_t)\tilde{C}^{-1}(t,x_t) = g^T(t,x_t)[g(t,x_t)g^T(t,x_t) + \tilde{g}(t)\tilde{g}^T(t)]^{-1/2}. \tag{4.1.22}$$

If we define

$$\tilde{k}(t,x_t) \triangleq (g(t,x_t)g^T(t,x_t) + \tilde{g}(t)\tilde{g}^T(t)), \tag{4.1.23}$$

we can rewrite (4.1.18) (i.e., (4.1.21)) as

$$\begin{aligned} dx_t = {} & [f(t,x_t,u_t) - \sigma(t,x_t,u_t)g^T(t,x_t)\tilde{k}^{-1}h(t,x_t)]dt \\ & + \sigma(t,x_tu_t)[I_m - g^T(t,x_t)\tilde{k}^{-1}(t,x_t)g(t,x_t)]^{1/2}d\tilde{w}_t \\ & + \sigma(t,x_t,u_t)g^T(t,x_t)\tilde{k}^{-1/2}(t,x_t)d\tilde{y}_t. \end{aligned} \tag{4.1.24}$$

The observations y_t are related to $\tilde{y}_t$ by (4.1.19). Notice that the diffusion process (4.1.6) is now expressed in terms of $\tilde{w}.$, $\tilde{y}.$ which are independent standard Wiener processes. From (4.1.19) we have

$$dy_t = \tilde{k}^{\frac{1}{2}}(t,x_t)d\tilde{y}_t$$

which is again well defined. Moreover, the LR given by 4.1.13 is now rewritten as

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$d\Lambda_t = \Lambda_t h^T(t,x_t)\tilde{k}^{-\frac{1}{2}}(t,x_t)d\tilde{y}_t. \tag{4.1.25}$$

Since the solution $\{x_s^u, y_s, 0\le s\le t\}$ is a unique strong solution to (4.1.6), (4.1.7) it follows that $\{x_s^u, y_s, 0\le s\le t\}$ is also a unique strong solution to (4.1.24) where the input is now the Wiener processes $\tilde{w}.$, $\tilde{y}.$ related to $w.$, $y.$ through (4.1.14), (4.1.17), respectively. Furthermore, the coordinate functions $(\tilde{w},\tilde{y})$ are independent of x_0; therefore, the probability space $(\Omega, \mathcal{F}_T, \tilde{P}_1)$ can be rewritten as a product of two probability spaces $(\tilde{\Omega}\otimes\hat{\Omega},\ \tilde{\mathcal{F}}_T^{\tilde{y}}\otimes\hat{\mathcal{F}}_T^{\tilde{w}},\ \tilde{P}\otimes\hat{P})$, where

$$\Omega \triangleq \tilde{\Omega}\otimes\hat{\Omega} = C([0,T];R^d)\otimes C([0,T];R^m)\otimes R^n$$
$$\mathcal{F}_T = \tilde{\mathcal{F}}_T^{\tilde{y}}\otimes\hat{\mathcal{F}}_T^{\tilde{w}}$$
$$P_1 = \tilde{P}\otimes\hat{P}$$

and by an abuse of notation $\hat{\mathcal{F}}_T^{\tilde{w}}(\equiv\mathcal{F}_T^{x,\tilde{w}})$ denotes the complete filtration on $C([0,T];R^m)\otimes B^n$, $\hat{P}(d\tilde{w})$ denotes the measure $\hat{P}(dx,d\tilde{w})$. Note that if x is deterministic the notation is well defined with no abuse.

Since the new functions $(\tilde{y},\tilde{w})$ are independent standard Wiener processes on the probability space $(\Omega,\ \mathcal{F}_T,\ \tilde{P}_1)$, the measure $\tilde{P}(d\tilde{y},d\tilde{w})$ defined on Ω is a Wiener measure which satisfies

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$\tilde{P}_1(d\tilde{y},d\tilde{w}) = \tilde{P}(d\tilde{y}) \otimes \hat{P}(d\tilde{w})$$

where $\tilde{P}(d\tilde{y}) = \mu_w^d(d\tilde{y})$ a Wiener measure.

Hence on the product probability space $(\tilde{\Omega}\otimes\hat{\Omega},\ \tilde{\mathcal{F}}_T^{\tilde{y}}\otimes\hat{\mathcal{F}}_T^{\tilde{w}},\ \tilde{P}\otimes\hat{P})$, we consider the stochastic differential equation

$$\begin{aligned} dx_t &= (f^u(t,x_t) - \sigma^u(t,x_t)g^T(t,x_t)\tilde{k}^{-1}(t,x_t)h(t,x_t))dt \\ &\quad + \sigma^u(t,x_t)E^{1/2}(t,x_t)d\tilde{w}_t + \sigma^u(t,x_t)g^T(t,x_t)\tilde{k}^{-1/2}(t,x_t)d\tilde{y}_t \end{aligned} \tag{4.1.26}$$

where superscript u denotes dependence on the control u and, for simplicity, we have defined

$$E(t,x_t) \triangleq I_m - g^T(t,x_t)\tilde{k}^{-1}(t,x_t)g(t,x_t).$$

If we now assume $\mathcal{F}_T^y = \sigma\{\tilde{y}_s, 0 \le s \le t\}$ (which is actually not true unless $g(t,x_t)$ is independent of x_t) and apply the Ito differential rule to any $\phi \in C_b^2(R^n)$ and perform some cancellations we obtain

$$\begin{aligned} \phi(x_t) &= \phi(x_0) + \int_0^t \frac{\partial\phi(x_s)}{\partial x}dx_s + \frac{1}{2}\int_0^t \frac{\partial^2\phi(x_s)}{\partial x^2}d\langle x.,x.\rangle_s \\ &= \phi(x_0) + \int_0^t \tilde{L}^u(s)\phi(x_s)ds + \int_0^t \frac{\partial\phi(x_s)}{\partial x}\sigma^u(s,x_s)E^{1/2}(s,x_s)d\tilde{w}_s \\ &\quad + \int_0^t \frac{\partial\phi(x_s)}{\partial x}\sigma^u(s,x_s)g^T(s,x_s)\tilde{k}^{-\frac{1}{2}}(s,x_s)d\tilde{y}_s \end{aligned}$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

where

$$\tilde{L}^u(\cdot) = \frac{1}{2}\sum_{i,j}^{n} a^{ij}(t,x,u)\frac{\partial^2}{\partial x^i \partial x^j} + \sum_{i=1}^{n}\left(f^i(t,x,u) - (\sigma(t,x,u)g^T(t,x)\tilde{k}^{-1}(t,x)h(t,x))\right)^i \frac{\partial}{\partial x^i}$$

$a(t,x,u) \triangleq \sigma(t,x,u)\sigma^T(t,x,u)$, and f^i the i-th component of vector field f.

Applying the Ito differential rule to $\phi(x_t)\Lambda_t$, where Λ_t is given by (4.1.25),

$$\begin{aligned}\phi(x_t)\Lambda_t &= \phi(x_0) + \int_0^t \Lambda_s \tilde{L}^u(s)\phi(x_s)ds + \int_0^t \Lambda_s \frac{\partial\phi(x_s)}{\partial x}\sigma^u(s,x_s)E^{1/2}(s,x_s)d\tilde{w}_s \\ &+ \int_0^t \Lambda_s \frac{\partial\phi(x_s)}{\partial x}\sigma^u(s,x_s)g^T(s,x_s)\tilde{k}^{-\frac{1}{2}}(s,x_s)d\tilde{y}_s + \int_0^t \Lambda_s\phi(x_s)h^T(s,x_s)\tilde{k}^{-\frac{1}{2}}(s,x_s)d\tilde{y}_s \\ &+ \int_0^t \Lambda_s \frac{\partial\phi(x_s)}{\partial x}\sigma^u(s,x_s)g^T(s,x_s)\tilde{k}^{-1}(s,x_s)h(s,x)ds.\end{aligned}$$

Thus, after cancellation

$$\begin{aligned}\phi(x_t)\Lambda_t &= \phi(x_0) + \int_0^t \Lambda_s A^u(s)\phi(x_s)ds + \int_0^t \Lambda_s \frac{\partial\phi(x_s)}{\partial x}\sigma^u(s,x_s)E^{1/2}(s,x_s)d\tilde{w}_s \\ &+ \int_0^t \Lambda_t \tilde{M}^u(s)\phi(x_s)d\tilde{y}_s\end{aligned}$$

where

$$A^u(t) \triangleq \frac{1}{2}\sum_{i,j}^{n} a^{ij}(t,x,u)\frac{\partial^2}{\partial x^i \partial x^j} + \sum_{i=1}^{n} f^i(t,x,u)\frac{\partial}{\partial x^i}$$

$$\tilde{M}_k(t) \triangleq \sum_{i=1}^{d}(\tilde{k}^{-\frac{1}{2}}(t,x_t))^{ik}h_i(t,x_t) + \sum_{i=1}^{n}(\sigma^u(t,x)g^T(t,x)\tilde{k}^{-\frac{1}{2}}(t,x_t))^{ik}\frac{\partial}{\partial x^i}. \qquad (4.1.27)$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The next and final step to prove the validity of Theorem 4.1.1 would require the conditional expectation of the integral representation of $\phi(x_t)\Lambda_t$ with respect to $\tilde{\mathcal{F}}_t^{\tilde{y}}(\equiv \mathcal{F}_t^y)$. We therefore proceed as follows. From the definition of conditional expectation (see, e.g., Wong and Hajek [132]), for every $A \in \tilde{\mathcal{F}}_t^{\tilde{y}}$,

$$\int_A \phi(x_t)\Lambda_t d\tilde{\mathcal{P}}_1 \triangleq \int_A \tilde{E}[\phi(x_t)\Lambda_t | \tilde{\mathcal{F}}_t^{\tilde{y}}] d\tilde{\mathcal{P}}_1 .$$

Recalling that the measure $\tilde{\mathcal{P}}_1$ has been decomposed into the product of the measures $\tilde{\mathcal{P}}(d\tilde{y})$ (sometimes called a delta measure) and $\hat{\mathcal{P}}(d\tilde{w})$ above, the right side of the previous definition actually denotes the double integral

$$\int_A \int_A \tilde{E}(\phi(x_t)\Lambda_t | \tilde{\mathcal{F}}_t^{\tilde{y}}) \tilde{\mathcal{P}}(d\tilde{y}) \hat{\mathcal{P}}(d\tilde{w}) .$$

This, in turn, can be rewritten as

$$\int_A \tilde{E}[\tilde{E}(\phi(x_t)\Lambda_t | \tilde{\mathcal{F}}_t^{\tilde{y}}) | \tilde{\mathcal{F}}_t^{\tilde{y}}] \hat{\mathcal{P}}(d\tilde{w}) = \int_A \tilde{E}(\phi(x_t)\Lambda_t | \tilde{\mathcal{F}}_t^{\tilde{y}}) \hat{\mathcal{P}}(d\tilde{w})$$

where the last equality follows by the measurability of the inner conditional expectation. Then, by Liptser and Shiryayev [103, pp. 187, 188] the last expression equals its integral, while we now define

$$\tilde{E}_{\hat{\mathcal{P}}}(\phi(x_t)\Lambda_t) \triangleq \tilde{E}(\phi(x_t)\Lambda_t | \tilde{\mathcal{F}}_t^{\tilde{y}}) .$$

Summarizing, we now write for every $A \in \tilde{\mathcal{F}}_t^{\tilde{y}}$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$\int_A \phi(x_t)\Lambda_t d\tilde{\mathcal{P}}_1 = \int_A \tilde{E}[\phi(x_t)\Lambda_t | \mathcal{F}_t^{\tilde{y}}] d\tilde{\mathcal{P}}_1 = \tilde{E}_{\mathcal{P}}(\phi(x_t)\Lambda_t) \tag{4.1.28}$$

Next, since $\phi(x_t)\Lambda_t$ is an $\mathcal{F}_y^{\tilde{y}} \otimes \mathcal{F}_t^{\tilde{w}}$-adapted process it follows that

$$\begin{aligned}
\tilde{E}_{\mathcal{P}}[\phi(x_t)\Lambda_t] &= \phi(x) + \tilde{E}_{\mathcal{P}}[\int_0^t \Lambda_s A^u(s)\phi(x_s)ds] \\
&+ \tilde{E}_{\mathcal{P}}[\int_0^t \Lambda_s \frac{\partial\phi(x_s)}{\partial x}\sigma^u(s,x_s)E^{\frac{1}{2}}(s,x_s)d\tilde{w}_s] \\
&+ \tilde{E}_{\mathcal{P}}[\int_0^t \Lambda_s \tilde{M}^u(s)\phi(x_s)d\tilde{y}_s].
\end{aligned}$$

Using a version of Fubini's Theorem as given by Kunita [96, Lemma 1.2, pp. 132-133] or Liptser and Shiryayev [103, Chp. 5, Thm. 5.15, pp. 187-188] and the fact that $\tilde{w}_t$ is a Wiener process with respect to measure $\hat{\mathcal{P}}$, we obtain

$$\tilde{E}_{\mathcal{P}}(\phi(x_t)\Lambda_t) = \phi(x) + \int_0^t \tilde{E}_{\mathcal{P}}[\Lambda_s A^u(s)\phi(x_s)]ds + \int_0^t \tilde{E}_{\mathcal{P}}[\Lambda_s \tilde{M}^u(s)\phi(x_s)]d\tilde{y}_s.$$

It then follows that there exists a measure-valued process $\rho_t(\phi) = \tilde{E}_{\mathcal{P}}(\Lambda_t\, \phi(x_t))$ which is considered a weak solution to (4.1.13). This completes the proof. QED

Remark 4.1.1 Notice that if we 1) exclude dependence of g on x, 2) let σ be an nxn matrix, 3) consider no control present, and 4) assume the condition

$$g(t)g^T(t) + \tilde{g}(t)\tilde{g}^T(t) = I_d \tag{4.1.29}$$

is satisfied, then $\rho_t(\phi)$ satisfies the exact equation presented by Pardoux [114, 115]. That is, the stochastic PDE given by Pardoux [114, 115] is a special case of the one given in

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Theorem 4.1.1 when the dxd matrix $\tilde{k}^{-\frac{1}{2}}(t,x)$ is replaced by I_d. Kunita [96] treats the filtering problem (4.1.1), (4.1.2) when g is zero and $dx_t = f(t,x_t)dt + \sigma(t,x_t)dw + \hat{g}(t,x_t)db_t$. His derivation is based on the one-to-one correspondence between the evolution of Kushner's equation [98] (i.e., the equation satisfied by the normalized conditional density) and the stochastic PDE satisfied by the unnormalized conditional density. Recently Bensoussan [15] presented stochastic PDE's satisfied by unnormalized density for two cases. The first case considers the filtering problem when $\sigma(x_t)$ is an nxn matrix and $g(t,x) \rightarrow g(t)$, and the second case considers the filtering problem when $\sigma(x_t) \rightarrow \sigma(x_t, y_t)$, an nxn matrix, and $dy_t = h(x_t, y_t)dt + g(y_t)dw_t + db_t$. However, his derivation is different than ours.

For the rest of the chapter we shall assume for reasons of correctness that $g(t,x) \rightarrow g(t)$ and for reasons of simplicity that (4.1.29) is satisfied. As a consequence $\tilde{k}^{-1}(t,x) \rightarrow I_d$ thus $dy_t = d\tilde{y}_t$ and $\tilde{\mathscr{F}}_t^{\tilde{y}} = \mathscr{F}_t^y$. However, the result to be presented in this chapter can always modified to cover the cases considered by Kunita [96], Bensoussan [15], and also the case when $\tilde{k}(t,x) \neq I_d$ where $\tilde{k}(t,x_t)$ is independent of x_t. We mention that the case when $g(t,x_t)$ depends on x_t is still open unless we can show that $\mathscr{F}_t^y = \tilde{\mathscr{F}}_t^{\tilde{y}}$. If we express the stochastic integral of the stochastic PDE derived in Theorem 4.1.1 in terms of a F-S integral, by (7.D.4) we have

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$d\rho_t(f)(x) = \rho_t(A^u(t)f - \frac{1}{2}\sum_{k=1}^{d} M_k^u(t)^2 f)(x)dt + \sum_{k=1}^{d} \rho_t(M_k^u(t)f)(x)\cdot dy_t^k \qquad (4.1.30)$$
$$\lim_{t\downarrow 0} \rho_t(f)(x) = f(x).$$

where

$$M_k^k(t) \triangleq M_k^u(t) + \sum_{i=1}^{n} (\sigma^u(t,x)g^T(t))\frac{\partial}{\partial x^i}.$$

The solution to (4.1.13) can be considered as a weak solution to the generalized solution of (4.1.30) by treating ϕ as a test function.

Next we represent the differential operators $A^u(t)$, $M_k^u(t)$ using the tangent space basis as described in Appendix 7.D. Thus,

$$M_k^u(t) \triangleq h_k(t) + Y_k^u(t),\ k=1,2,\ldots,d$$

$$A^u(t) \triangleq \frac{1}{2}\sum_{j=1}^{m} X_j^u(t)^2(t) + X_0^u(t).$$

$$X_j^u(t) \triangleq \sum_{i=1}^{n} X_j^i(t,x,u)\frac{\partial}{\partial x^i},\quad Y_k^u(t) \triangleq \sum_{i=1}^{n} Y_k^i(t,x,u)\frac{\partial}{\partial x^i},\quad h_k(t) \triangleq h_k(t,x_t),\ 1 \le j \le m.$$

The operator $L^u(t)$ defined by

$$L^u(t) \triangleq A^u(t) - \frac{1}{2}\sum_{k=1}^{d} M_k^u(t)^2 \qquad (4.1.31)$$

can be represented as

$$L^u(t) = \frac{1}{2}\sum_{j=1}^{m} \tilde{X}_j^u(t)^2 + \tilde{X}_0^u(t) + \tilde{h}_0(t) \qquad (4.1.32)$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

where $\tilde{X}_j^u$, $j = 0,\dots,m$ are first order differential operators denoted in the following lemma.

Lemma 4.1.2 Suppose the correlation between the state process and observation process is denoted by

$$\langle x_\cdot, y_\cdot \rangle_T = \int_0^T \sigma(t) g^T(t, x_t, u_t) dt .$$

Then

$$Y_k^u(t) = \sum_{j=1}^{m} \gamma^{kj}(t) X_j^u(t) = \sum_{i=1}^{n} \sum_{j=1}^{m} \gamma^{kj}(t) X_j^i(t, x_t, u_t) \frac{\partial}{\partial x^i}, \quad k = 1, \dots, d \qquad (4.1.33)$$

and $I_m - g^T(t)g(t)$ is nonnegative-definite, where γ^{ij} is the (i,j) component of g(t). Also defining the mxm matrix Θ with components Θ^{ij}, where $\Theta^T\Theta(t) = I_m - g^T(t)g(t)$, then the operator L(t) of (4.1.32) can be identified as

$$\tilde{X}_j^u(t) = \sum_{k=1}^{m} \Theta^{jk}(t) X_k^u, \quad j = 1,\dots,m \qquad (4.1.34)$$

$$\tilde{X}_0^u(t) = X_0^u(t) - \sum_{k=1}^{d} h_k(t) Y_k^u(t) \qquad (4.1.35)$$

$$\tilde{h}_0(t) = -\frac{1}{2} \sum_{k=1}^{d} h_k^2(t). \qquad (4.1.36)$$

Finally, the solution to (4.1.30) can be expressed as

$$\rho_t(f)(x, \bar{\omega}) = \tilde{E}_{\hat{\omega}}[f(\xi_{s,t}(x, \bar{\omega}, \cdot))\phi_{s,t}(x, \bar{\omega}, \cdot)] \qquad (4.1.37)$$

where $\phi_{s,t}(x) \equiv \phi_{s,t}(x, \bar{\omega}, \hat{\omega})$ is defined as

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$\phi_{s,t}(x) \triangleq e^{\sum_{k=1}^{d}\int_s^t h_k(r,\xi_{s,r}(x))\cdot dy_r^k + \int_s^t \tilde{h}_0(r,\xi_{s,r}(x))dr} \tag{4.1.38}$$

and the process $\xi_{s,t}(x,\tilde{\omega},\hat{\omega})$, $\tilde{\omega}\in\tilde{\Omega}$, $\hat{\omega}\in\hat{\Omega}$ starting at $\xi_s = x$ is the solution to the stochastic differential equation

$$(\tilde{\Omega}\otimes\hat{\Omega}, \tilde{\mathscr{F}}_t^y\otimes\hat{\mathscr{F}}_t^w, \tilde{P}\otimes\hat{P}): d\xi_t = \tilde{X}_0^u(t,\xi_t) + \sum_{j=1}^{m}\tilde{X}_j^u(t,\xi_t)\cdot dw_t^j + \sum_{k=1}^{d} Y_k^u(t,\xi_t)\cdot dy_t^k \tag{4.1.39}$$

defined on R^n (where in this case $\tilde{\mathscr{F}}_t^y = \mathscr{F}_t^y$ since $dy = d\tilde{y}$ a standard Brownian Motion).

Proof: The nonnegativity of $I_m - g^T(t)g(t)$ follows from Theorem 4.1.1. Next (4.1.33) is a representation of $\langle x.,y.\rangle_t$ using the tangent space basis, which can be easily verified. Using (4.1.31) and (4.1.33).

$$\begin{aligned} L^u(t) = {} & \frac{1}{2}\sum_{i,j=1}^{m}(\delta_{ij} - \sum_{i=1}^{d}\gamma^{ki}(t)\gamma^{kj}(t))X_i^u(t)X_j^u(t) \\ & - \sum_{k=1}^{d} h_k(t)Y_k^u(t) + X_0^u(t) - \frac{1}{2}\sum_{k=1}^{d} h_k^2(t). \end{aligned}$$

Since Θ is nonnegative-definite, if we define

$$\tilde{X}_j^u(t) \triangleq \sum_{k=1}^{m}\Theta^{jk}X_i^u, \quad j = 1, \dots, m$$

we obtain the representation given by (4.1.34) - (4.1.36).

From the theory of partial differential equations we can associate with the stochastic PDE (4.1.30) a diffusion process ξ_t having a generator

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$\frac{1}{2}\sum_{j=1}^{m} \tilde{X}_j^u(t) + \tilde{X}_0^u(t) + \sum_{k=1}^{d} Y_k^u(t),$$

hence the diffusion process ξ_t should satisfy (4.1.39). Finally, if we apply the Fisk-Stratonovich differential rule to $f(\xi_t)\ \phi_t(x)$ we can show (see, Kunita [93, 96]) that (4.1.37) satisfies (4.1.30).

Remark 4.1.2 Kunita [93, 94, 96] considers the existence and uniqueness of the solution to a stochastic PDE similar to (4.1.30) when the operator L(t) is degenerate (see Remark 4.1.1). When L(t) is a hypoelliptic operator Kunita [93, 94, 96] shows that (4.1.30) has a C^∞ solution (in the weak sense).

During the same period Kunita [94] proved that when $M_k(t)$ is a zeroth-order operator, the hypoellipticity of L(t) is a necessary and sufficient condition for the existence of smooth densities. His approach however, differs from the one reported in Kunita [93] in that the decomposition of (4.1.30) into two measure-valued processes is considered. These two equations are expressed as

$$d\mu_t(\phi) = \sum_{k=1}^{d} \mu_t(M_k^u(t)\phi)\cdot dy_t^k\ , \quad \lim_{t\downarrow t_0} \mu_t = \delta_x \tag{4.1.40}$$

$$dv_t(\phi) = v_t(\mu_t L^u(t)\mu_t^{-1}\phi)dt\ , \quad \lim_{t\downarrow t_0} \mu_t = \delta_x \tag{4.1.41}$$

where δ_x denotes the delta function concentrated at x. Once the above decomposition is established, then existence of a C^∞ density $\rho_t(x,\tilde{\omega},dz)$ is shown by first showing that the solution to (4.1.41) represented as $v_t(x, dz)$ has a C^∞ density. The composition $\rho_t = v_t\mu_t$ is then a solution of (4.1.30).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

In this chapter we shall adapt the result established in Lemma 4.1.2 together with the decomposition method above to present two approaches in obtaining necessary conditions for the stochastic minimum principle of partially observed diffusions having strict sense admissibility. Within each approach we shall treat two cases,

Case (i)

when $M_k^u(t) \to M_k(t)$ (no control dependence) which implies that $\sigma(t,x,u) \to \sigma(t,x)$

Case (ii)

When $M_k^u(t) \to h_k(t)$ (no correlation) but σ depends on control u.

Both approaches correspond to the Pontryagin's minimum principle in the case of deterministic systems. Approach one is easier and less involved due to the direct use of the results presented by Bismut [16] for state-valued processes and their extension by Kwakernaak [101] to measure-valued processes. Approach two appears to be more complicated since no previously known general stochastic minimum principle result is utilized. Both methods have in common the same exact representation for the variational cost. However, their adjoint (Lagrange multiplier)-processes have a different representation. This is of no surprise since there may be several forms of stochastic Lagrange multiplier processes as pointed out by Bensoussan [9, Chp. VI, pp. 221]. As mentioned in Chapter 1, Sections 1.1.4 - 1.1.5 our method differs from the currently existing approaches taken by Bensoussan [10, 15], Haussmann [68] and Elliott and Yang [50] in that, a separation principle for nonlinear systems is established similar to the one given by Wonham [130], whose results are only applicable to state-valued processes described by linear systems.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Moreover, to our knowledge our problem formulation is the first that considers correlation between state and observation processes. The approach we proposed in Section 1.2. requires the proof of certain theorems which we shall introduce in the next section. They are in fact the primary tools for our analysis.

4.2 DECOMPOSITION OF THE BACKWARD (SDE) AND ADJOINT PROCESSES

In this section we shall derive the decomposed measure-valued processes associated with the backward SDE, i.e., the adjoint of the forward SDE (4.1.30). Next, we shall derive a pair forward and backward stochastic differential equations satisfied by the process $\mu_t^{-1}(f)$ and a pair of parabolic PDE's satisfied by $\nu_t^{-1}(f)$ in both forward and backward variable. It is then evident that the inverse maps $\mu_t^{-1}(f)$, $\nu_t^{-1}(f)$ (in both forward and backward variables) have properties analogous to those shared by state-valued processes defined in Euclidean space. In fact, $\mu_t^{-1}(f)$, $\nu_t^{-1}(f)$ are solutions of stochastic differential equations and PDE's, respectively, which are the adjoint equations to the ones satisfied by $\mu_t(f)$, $\nu_t(f)$. Similar results for state-valued processes are given in Kunita [95, Thm. 3]. Indeed, using the properties of stochastic flows for processes in Euclidean spaces Bensoussan [11] derives necessary conditions for the stochastic control problem with complete observations.

The results established in this section enable us to derive in an explicit way the change in the cost due to weak variations of the optimal control and a stochastic PDE

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

satisfied by the stochastic Lagrange multiplier for the problem stated in Section 4.1. In other words the above processes constitute the tools for determining the stochastic minimum principle and the measure-valued adjoint process for the partially observed stochastic control problem.

We start by proving the following lemma stated without proof in Kunita [96, Thm. 4.2, p. 144]. For notational simplicity, the superscript u used to denote dependence on the control variable will be dropped and re-introduced in the next section.

Lemma 4.2.1

The solution $\rho_{s,T}(f)$ of (4.1.30) can be represented by

$$\begin{aligned} d\rho_{t,T}(f)(x,\hat{\omega}) &= -L(t)\rho_{t,T}(f)(x,\hat{\omega})dt - \sum_{k=1}^{d} M_k(t)\, \rho_{t,T}(f)(x,\hat{\omega})\cdot \hat{d}y_t^k \\ \lim_{t\uparrow T} \rho_{t,T}(f) &= f(x). \end{aligned} \tag{4.2.1}$$

Proof: Applying the stochastic differential rule of Theorem 7.D.1 to $f(\xi_{s,t}(x))\, \phi_{s,t}(x)$ where $\xi_{s,t}$ satisfies the stochastic differential equation (4.1.39) and $\phi_{s,t}$ is defined by (4.1.38), we obtain

$$\begin{aligned} f(\xi_{s,t}(x))\phi_{s,t}(x) &= f(x) + \int_s^t \tilde{X}_0(r) f(\xi_{s,r})\phi_{s,r} dr + \sum_{j=1}^{m} \int_s^t \tilde{X}_j(r) f(\xi_{s,r})\phi_{s,r}\cdot dw_r^j \\ &+ \sum_{k=1}^{d} \int_s^t Y_j(r) f(\xi_{s,r})\phi_{s,r}\cdot dy_r^k + \sum_{k=1}^{d} \int_s^t h_k(r) f(\xi_{s,r})\phi_{s,r}\cdot dw_r^j + \int_s^t \tilde{h}_0(r) f(\xi_{s,r})\phi_{s,r} dr. \end{aligned}$$

Using (7.D.4), the $(\hat{\Omega}, \hat{\mathcal{F}}_t^w, \hat{P})$ martingales are written in terms of Ito integrals, thus

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$f(\xi_{s,t}(x))\phi_{s,t}(x) = f(x) + \int_s^t \Big(\frac{1}{2}\sum_{j=1}^m \tilde{X}_j(r)^2 + \tilde{X}_0(r) + \tilde{h}_0(r)\Big) f(\xi_{s,r})\phi_{s,r} dr$$
$$+ \sum_{j=1}^m \int_s^t \tilde{X}_j(r) f(\xi_{s,r})\phi_{s,r} dw_r^j + \sum_{k=1}^d \int_s^t M_k(r) f(\xi_{s,r})\phi_{s,r} \cdot dy_r^k.$$

By Theorem 7.D.1, with t fixed

$$f(\xi_{s,t}(x))\phi_{s,t}(x) = f(x) + \int_s^t \Big(\frac{1}{2}\sum_{j=1}^m \tilde{X}_j(r)^2 + \tilde{X}_0(r) + \tilde{h}_0(r)\Big)(f \cdot \xi_{r,t}\phi_{r,t})(r,x) dr$$
$$+ \sum_{j=1}^m \int_s^t \tilde{X}_j(r)(f \cdot \xi_{r,t}\phi_{r,t})(r,x) dw_r^j + \sum_{k=1}^d \int_s^t M_k(r)(f \cdot \xi_{r,t}\phi_{r,t})(r,x) \cdot \hat{d}y_r^k.$$

Taking expectation with respect to measure $\hat{\mathcal{P}}$ the $\mathcal{F}_t^w$-martingales are zero. Thus,

$$\tilde{E}_{\hat{\mathcal{P}}}(f(\xi_{s,t}(x))\phi_{s,t}(x) = f(x) + \int_s^t L(t)\, \tilde{E}_{\hat{\mathcal{P}}}[(f \cdot \xi_{r,t}\phi_{r,t})(r,x)] dr$$
$$+ \sum_{k=1}^d \int_s^t M_k(r) \tilde{E}_{\hat{\mathcal{P}}}[(f \cdot \xi_{r,t}\phi_{r,t})] \cdot \hat{d}y_r^k.$$

setting t = T, s = t and differentiating with respect to the backward variable t and using (4.1.27) we obtain (4.2.1). QED

Theorem 4.2.1

The measure-valued process $V_t(x) \triangleq \mu_t(v_t(f))(x)$ is the solution of the backward SDE

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$dV_t(x) + L(t)V_t(x)dt + \sum_{k=1}^{d} M_k(t)\, V_t(x)\cdot\hat{d}y_t^k = 0 \qquad (4.2.3)$$
$$\lim_{t\uparrow T} V_t(x) = f(x)$$

where

$$d\mu_t(x,\bar{\omega}) = -\sum_{k=1}^{d} M_k(t)\, \mu_t(x,\bar{\omega})\cdot\hat{d}y_t^k \qquad (4.2.4)$$
$$\lim_{t\uparrow T} \mu_t(x,\bar{\omega}) = f(x)$$

$$dv_t(x) = -(\mu_t^{-1}L(t)\mu_t)v_t(x,\bar{\omega})dt \qquad (4.2.5)$$
$$\lim_{t\uparrow T} v_t(x,\bar{\omega}) = f(x).$$

Proof: Let us first verify that $\mu_t(v_t(f))$ satisfies (4.2.3). Using the F-S backward differential rule given by Theorem 7.B.6,

$$\begin{aligned} d(\mu_t(v_t(f)) &= d\mu_t(v_t(f)) + \mu_t(dv_t(f)) \\ &= -\sum_{k=1}^{d} M_k(t)\mu_t(v_t(f))\cdot\hat{d}y_t^k - \mu_t((\mu_t^{-1}L(t)\mu_t)v_t(f))dt \\ &= -\sum_{k=1}^{d} M_k(t)\mu_t(v_t(f))\cdot\hat{d}y_t^k - L(t)\mu_t(v_t(f))dt \end{aligned}$$

which verifies (4.2.3).

Let us first prove (4.2.4). If we set $L(t) = 0$ then $\mu_t(x,\bar{\omega})$ is a special case of (4.2.1) or (4.2.3). Let $\eta_{s,t}$ be the solution of (4.1.39) with $\tilde{X}_0, \ldots, \tilde{X}_m$ equal to zero; then

$$d\eta_t = \sum_{k=1}^{d} Y_k(t,\eta_t)\cdot dy_t^k$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

and $\eta_{s,t}(x,\hat{\omega})$ does not depend on $\hat{\omega}$. From (4.1.37), the solution of (4.2.3) (i.e., 4.2.1) is written as

$$\mu_t(x,\hat{\omega}) = f(\eta_{s,t}(x))\, e^{\sum_{k=1}^{d} \int_s^t h_k(r,\eta_{s,r}(x,\hat{\omega}))\cdot dy_r^k}$$

which can be shown as follows.

Define

$$\phi_{s,t}(x) \triangleq e^{\sum_{k=1}^{d} \int_s^t h_k(r,\eta_{s,r}(x,\hat{\omega}))\cdot dy_r^k}$$

and apply the F-S differential rule to $f(\eta_{s,t}(x))\phi_{s,t}(x)$. Thus,

$$\begin{aligned} f(\eta_{s,t}(x)\phi_{s,t}(x) &= f(x) + \sum_{k=1}^{d} \int_s^t Y_k(r) f(\eta_{s,r}(x))\phi_{s,r}\cdot dy_r^k \\ &\quad + \sum_{k=1}^{d} \int_s^t f(\eta_{s,r}(x)) h_k(r,\eta_{s,r}(x))\phi_{s,r}\cdot dy_r^k \\ &= f(x) + \sum_{k=1}^{d} \int_s^t M_k(r) f(\eta_{s,r}(x))\phi_{s,r}\cdot dy_r^k. \end{aligned}$$

Interchanging the forward variable t and backward variable s, we deduce

$$f(\eta_{s,t}(x)\phi_{s,t}(x) = f(x) + \sum_{k=1}^{d} \int_s^t M_k(r)(f\cdot\eta_{r,t}\phi_{r,t})(r,x)\cdot \hat{d}y_r^k.$$

If we fix time t and differentiate with respect to the backward variable s (4.2.4) is satisfied. Denote the solution $\mu_t(x,\hat{\omega})$ by $\mu_t(f)(x,\hat{\omega})$. From (A′3) and the bounded

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

assumption on f, it can be shown (see Kunita [94, 96]) that (4.2.4) has a unique solution in the L^2 sense. Moreover, the map μ_t is one-to-one and onto with inverse μ_t^{-1} given by

$$\mu_t^{-1}(f)(x,\hat{\omega}) = f(\eta_{s,t}^{-1}(x))\, e^{-\sum_{k=1}^{d}\int_s^t h_k(r,\eta_{r,t}^{-1}(x,\hat{\omega}))(r)\cdot \hat{d}y_r^k}.$$

It remains to show (4.2.5). Consider the second-order operator $\mu_t^{-1}L(t)\mu_t$ which is well-defined since $\mu_t^{-1}: C_b^\infty(R^n) \to C_b^\infty(R^n)$. Then,

$$\begin{aligned}\mu_t^{-1}L(t)\mu_t &= \mu_t^{-1}\Big(\frac{1}{2}\sum_{j=1}^{m}\tilde{X}_j(t)^2 + \tilde{X}_0(t) + \tilde{h}_0(t)\Big)\mu_t \\ &= \frac{1}{2}\sum_{j=1}^{m}(\mu_t^{-1}\tilde{X}_j(t)\mu_t)^2 + \mu_t^{-1}\tilde{X}_0(t)\mu_t + \tilde{h}_0(t).\end{aligned}$$

But

$$\begin{aligned}\mu_t^{-1}\tilde{X}_j(t)\mu_t &= \eta_{s,t}^{-1}(x)\exp\{-\sum_{k=1}^{d}\int_s^t h_k(r,\eta_{r,t}^{-1}(x,\hat{\omega}))\cdot\hat{d}y_r^k\} \\ &\quad x\,\tilde{X}_j(t)(\eta_{s,r}(x)\exp\{\sum_{k=1}^{d}\int_s^t h_k(r,\eta_{s,r}(x,\hat{\omega}))\cdot dy_r^k\}) \\ &= \eta_{s,t}^{-1}(x)\exp\{-\sum_{k=1}^{d}\int_s^t h_k(r,\eta_{r,t}^{-1}(x,\hat{\omega}))\cdot\hat{d}y_r^k\}[\tilde{X}_j(t)(\eta_{s,t}(x))\exp\{\sum_{k=1}^{d}\int_s^t h_k(r,\eta_{s,r}(x,\hat{\omega}))\cdot dy_r^k\} \\ &\quad + \tilde{X}_j(t)(\exp\{\sum_{k=1}^{d}\int_s^t h_k(r,\eta_{s,r}(x,\hat{\omega}))\cdot dy_r^k\})\eta_{s,t}(x)].\end{aligned}$$

Since

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$\int_s^t h_k(r,\eta_{r,t}^{-1}(y,\hat{\omega}))|_{y=\eta_{s,t}(x)} \cdot \hat{d}y_r^k = \int_s^t h_k(r,\eta_{s,r}(x,\hat{\omega})) \cdot dy_r^k$$

which is an application of the backward F-S integral (see Kunita [94, Lemma 1]) and

$$\tilde{X}_j(\eta_{s,t}(x))(\eta_{s,t}^{-1}(x)) = \eta_{s,t*}(\tilde{X}_j)(x)$$

which is an application of (7.D.12) (with f=1 where $\eta_{s,t*}(\tilde{X}_j)$ is a stochastic vector field and $\eta_{s,t*}$ is the differential map of $\eta_{s,t}$), we deduce

$$\begin{aligned}
\mu_t^{-1}\tilde{X}_j(t)\mu_t &= \eta_{s,t*}(\tilde{X}_j) + \tilde{X}_j(t)(\sum_{k=1}^{d}\int_s^t h_k(r,\eta_{s,r}(x,\hat{\omega}))\cdot dy_r^k) \\
&= \eta_{s,t*}(\tilde{X}_j) + \sum_{k=1}^{d}\int_s^t \tilde{X}_j(r)h_k(r,\eta_{s,r}(x,\hat{\omega}))\cdot dy_r^k) \\
&= \eta_{s,t*}(\tilde{X}_j) + g_{s,t}^j(x).
\end{aligned}$$

The function $g_{s,t}^j$ is now defined as

$$g_{s,t}^j(x) \triangleq \sum_{k=1}^{d}\int_s^t \tilde{X}_j(r)h_k(r,\eta_{s,r}(x,\hat{\omega}))\cdot dy_r^k.$$

Thus,

$$\mu_{s,t}^{-1}L(t)\mu_{s,t} = \frac{1}{2}\sum_{j=1}^{m}(\eta_{s,t*}(\tilde{X}_j) + g_{s,t}^j(x))^2 + \eta_{s,t*}(\tilde{X}_0(t)) + \tilde{h}_0(t) + g_{s,t}^0(x)$$

Following Kunita [92, 95], we construct the solution $\xi_{s,t}(x,\hat{\omega},\hat{\omega})$ of (4.1.39) as follows: Let

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$d\zeta_t = \sum_{j=1}^{m} \eta_{s,t*}(\tilde{X}_j)(\zeta_t)\cdot dw_t^j + \eta_{s,t*}(\tilde{X}_0)(\zeta_t)dt,\ \zeta_s = x$$
$$= \sum_{j=1}^{m} \tilde{X}_j(\zeta_t)\cdot dw_t^j + \tilde{X}_0(\zeta_t)dt.$$

By Kunita [92, Sect. 4, Proposition 4.2], the solution $\xi_{s,t}$ can be represented by the composition $\eta_{s,t}\,\zeta_{s,t}$ (the proof requires the extended Ito's formula). The solution to (4.2.5) can now be represented as

$$v_t(x,\tilde{\omega}) = \tilde{E}_{\hat{\omega}}\{f(\zeta_{s,t}(x))e^{\sum_{j=1}^{m}\int_s^t g_{s,r}^j(\zeta_{s,r}(x))\cdot dw_r^j + \int_s^t g_{s,r}^0(\zeta_{s,r}(x))dr + \int_s^t \tilde{h}_0(r)dr}\}$$

which can be shown by defining

$$\phi_{s,t}(x) \triangleq e^{\sum_{j=1}^{m}\int_s^t g_{s,r}^j(\zeta_{s,r}(x))\cdot dw_r^j + \int_s^t g_{s,r}^0(\zeta_{s,r}(x))dr}\ e^{\int_s^t \tilde{h}_0(r)dr}$$

and applying the extended Ito's formula to $f(\zeta_{s,t}(x))\,\phi_{s,t}(x)$, thus

$$= f(x)) + \int_s^t \eta_{s,r*}(\tilde{X}_0(r) + g_{s,r}^0(\zeta_{s,r}) + h_0(r))f(\zeta_{s,r})\phi_{s,r}dr$$
$$+ \frac{1}{2}\sum_{j=1}^{m}\int_s^t (\eta_{s,r*}(\tilde{X}_j(r)) + g_{s,r}^j(\zeta_{s,r}))^2 f(\zeta_{s,r})\phi_{s,r}dr$$
$$+ \sum_{j=1}^{m}\int_s^t (\eta_{s,r*}(\tilde{X}_j(r)) + g_{s,r}^j)f(\zeta_{s,r})\phi_{s,r}dw_r^j.$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Writing the first two components of the above equation in terms of $\mu_{s,t}^{-1}L(t)\mu_{s,t}$ we obtain

$$f(\zeta_{s,t})\phi_{s,t}(x) = \int_s^t(\mu_{s,r}^{-1}L(t)\mu_{s,r})f(\zeta_{s,r})\phi_{s,r}dr + \sum_{j=1}^m\int_s^t(\eta_{s,r*}(\tilde{X}_j(r) + g_{s,r}^j)f(\zeta_{s,r})\phi_{s,r}dw_r^j.$$

Finally, interchanging the forward variable t and backward variable s

$$\begin{aligned} f(\zeta_{s,t})\phi_{s,t}(x) = &\int_s^t(\mu_{r,t}^{-1}L(t)\mu_{r,t})f\cdot\zeta_{r,t}\phi_{r,t}dr \\ &+ \sum_{j=1}^m\int_s^t(\eta_{r,t*}(\tilde{X}_j(r) + g_{s,r}^j)f\cdot\zeta_{r,t}\phi_{r,t}dw_r^j, \end{aligned}$$

taking expectation with respect to measure $\hat{P}$, by interchanging the operator $E_{\hat{P}}$ with the second order operator (obviously we can do this since $\mu_{s,t}^{-1}L(t)\mu_{s,t}$ is bounded),

$$E_{\hat{P}}[f(\zeta_{s,t})\phi_{s,t}(x)] = \int_s^t(\mu_{r,t}^{-1}L(t)\mu_{r,t})E_{\hat{P}}[f\cdot\zeta_{r,t}\phi_{r,t}(x)]dr$$

and differentiating with respect to variable s we deduce (4.2.5). QED

Theorem 4.2.2

Suppose we define $\bar{\mu}_t \triangleq \mu_t^{-1}(f)$, $\bar{v}_t \triangleq v_t^{-1}(f)$ then $\bar{\mu}_t, \bar{v}_t$ are the inverse maps of (4.1.40), (4.1.41) respectively. Furthermore, $\bar{\mu}_t$, $\bar{v}_t$ satisfy the backward measure-valued processes

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$d\bar{\mu}_t(x,\hat{\omega}) = \sum_{k=1}^{d} \bar{\mu}_t(M_k(t)f)\ (x,\hat{\omega})\cdot\hat{d}y_t^k \qquad (4.2.6)$$
$$\lim_{t\uparrow T} \bar{\mu}_t(f) = f(x)$$

$$d\bar{v}_t(f)(x,\hat{\omega}) = \bar{v}_t(\bar{\mu}_t L(t)\mu_t^{-1}f)(x,\hat{\omega})dt \qquad (4.2.7)$$
$$\lim_{t\uparrow T} \bar{v}_t(f) = f(x)$$

Proof: The proof of the first part of the theorem is easily established by applying stochastic differential rule to $\mu_t(\bar{\mu}_t(f)), v_t(\bar{v}_t(f))$. The proof of (4.2.6), (4.2.7) can be shown in a similar fashion as the proof of Theorem 4.2.1. Consider the solution to stochastic differential equation (4.1.39) having solution $\eta_{s,t}(x,\hat{\omega})$ when $L(t) = 0$. Then, $\eta_{s,t}(x,\hat{\omega})$ would satisfy

$$d\eta_{s,t} = \sum_{j=1}^{d} Y_j(t,\eta_{s,t})\cdot dy_t^k$$

and the inverse operator $\mu_{s,t}^{-1}(f)$ is defined as before by

$$\mu_{s,t}^{-1}(f)(x,\hat{\omega}) = f(\eta_{s,t}^{-1}(x,\hat{\omega}))e^{-\sum_{k=1}^{d}\int_s^t h_k(r,\eta_{r,t}^{-1}(x,\hat{\omega}))\cdot\hat{d}y_r^k} \qquad (4.2.8)$$

As shown by Kunita [95, Thm. 2] the inverse map $\eta_{s,t}^{-1}(x,\hat{\omega})$ satisfies the backward stochastic differential equation

$$d\hat{\eta}_{s,t} = \sum_{j=1}^{d} Y_j(t,\hat{\eta}_{s,t})\cdot\hat{d}y_t^k,\ \hat{\eta}_{t,t}(x) = x \qquad (4.2.9)$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

(see also Theorem 7.D.2). Since in our case Y_j^i, $j=1, \ldots, m$, $i=1, \ldots, d$ are bounded and $\eta_{s,t} \in R^n$ the inverse map $\eta_{s,t}^{-1}(x,\bar{\omega}) \in C^\infty(R^n)$. If we define $\phi_{s,t}$ to be the exponential component of (4.2.8) and apply the stochastic differential rule to $f(\hat{\eta}_{s,t}(t))\phi_{s,t}(x)$ where $\phi_{s,t}(x)$ is expressed in terms of $\hat{\eta}_{s,t}(x,\omega)$ we deduce

$$
\begin{aligned}
f(\hat{\eta}_{s,t}(x)\phi_{s,t}(x) &= x - \sum_{k=1}^{d} \int_s^t Y_k(r) f(\hat{\eta}_{r,t}(x))\phi_{r,t}(x) \cdot \hat{d}y_r^k \\
&\quad - \sum_{k=1}^{d} \int_s^t h_k(r) f(\hat{\eta}_{r,t}(x))\phi_{r,t}(x) \cdot \hat{d}y_r^k \\
&= x - \sum_{k=1}^{d} \int_s^t M_k(r) f(\hat{\eta}_{r,t}(x))\phi_{r,t}(x) \cdot \hat{d}y_r^k .
\end{aligned}
$$

The above equation satisfies (4.2.6) which can be shown by fixing t and differentiating with respect to s. The proof of (4.2.7) is similar to the one given in Theorem 4.2.1.

Next we shall present analogues results for measure-valued processes integrated forward in time.

Theorem 4.2.3

Let $\mu_{s,t}^{-1}(f)$, $\nu_{s,t}^{-1}(f)$ are the inverse operators to the measure-valued processes (4.1.40), (4.1.41) respectively. Then, the measure-valued processes equations $\mu_{s,t}^{-1}(f)$, $\nu_{s,t}^{-1}(f)$ satisfy the forward equations

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$d\mu_{s,t}^{-1}(f) = -\sum_{k=1}^{d} M_k(t)\mu_{s,t}^{-1}(f)\cdot dy_t^k$$
$$\lim_{t\downarrow s}\mu_{s,t}^{-1}(f) = f \tag{4.2.10}$$

$$dv_{s,t}^{-1}(f) = -(\mu_{s,t}L(t)\mu_{s,t}^{-1})v_{s,t}^{-1}(f)dt$$
$$\lim_{t\downarrow s}\mu_{s,t}^{-1}(f) = f \tag{4.2.11}$$

Proof: Before we provide the proof let us verify that

$$\mu_{s,t}^{-1}\mu_{s,t}(f) = f$$
$$v_{s,t}^{-1}v_{s,t}(f) = f.$$

by direct use of (4.2.8), (4.2.11). Applying Ito differential rule

$$\mu_{s,t}^{-1}\mu_{s,t}(f) = f + \int_s^t d\mu_{s,r}^{-1}(d\mu_{s,r}(f)) + \int_s^t \mu_{s,r}^{-1}(d\mu_{s,r}(f))$$
$$= f - \sum_{k=1}^{d}\int_s^t M_k(r)\mu_{s,r}^{-1}(\mu_{s,r}(f))\cdot dy_r^k + \sum_{k=1}^{d}\int_s^t \mu_{s,r}^{-1}(\mu_{s,r}(M_k(r)f))\cdot dy_r^k = f.$$

Repeating for $v_{s,t}^{-1}(v_{s,t}(f))$

$$v_{s,t}^{-1}(v_{s,t}(f)) = f - \int_s^t ((\mu_{s,r}L(r)\mu_{s,r}^{-1})v_{s,r}^{-1}v_{s,r}(f))dr + \int_s^t v_{s,r}^{-1}(v_{s,r}(\mu_{s,r}\, L(r)\mu_{s,r}^{-1}f))dr = f.$$

We start by considering the stochastic differential equation (4.2.10). Setting $L(t) = 0$, the solution to (4.1.39) on the product probability space $(\tilde{\Omega}\otimes\tilde{\Omega}, \mathcal{F}_t^y\otimes\mathcal{F}_t^w, \tilde{P}\otimes\tilde{P})$ is represented by

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$d\eta_t = \sum_{j=1}^{d} Y_j(t,\eta_t)\cdot dy_t^k. \tag{4.2.12}$$

Again the operator $\mu_{s,t}^{-1}$ (f) is well defined and is expressed as

$$\mu_{s,t}^{-1}(f) = f(\eta_{s,t}^{-1}(x,\bar{\omega}))\ \exp\{\ -\sum_{j=1}^{d}\int_s^t h_k(r,\eta_{r,t}^{-1}(x,\bar{w}))\cdot\hat{d}y_r^k\}. \tag{4.2.13}$$

We shall show that (4.2.13) satisfies (4.2.10). The key step in proving that (4.2.13) satisfies (4.2.10) lies in a theorem given by Kunita [95, Thm. 2]. This theorem states that if $\hat{\eta}_{s,t} = \eta_{s,t}^{-1}$ then

$$\hat{\eta}_{s,t} = x - \sum_{j=1}^{d}\int_s^t Y_j(r,\hat{\eta}_{r,t}(x))\cdot\hat{d}y_r^k\ \ ,\ \hat{\eta}_{t,t} = x. \tag{4.2.14}$$

where $\hat{\eta}_{s,t} = \eta_{s,t}^{-1}$ are onto maps as presented in Theorem 4.2.2. Thus (4.2.13) can be expressed as

$$\mu_{s,t}^{-1}(f) = f(\hat{\eta}_{s,t}(x,\bar{\omega}))\ \exp\{\ -\sum_{j=1}^{d}\int_s^t h_k(r,\hat{\eta}_{r,t}(x))\cdot\hat{d}y_r^k\}. \tag{4.2.15}$$

Setting $\tilde{\phi}_{s,t}(x,z) = \exp\{-\sum_{j=1}^{d}\int_s^t h_k(r,\hat{\eta}_{r,t}(x))\cdot\hat{d}y_r^k\}z$ then by applying stochastic differential formula to $f(\hat{\eta}_{s,t}(x))\,\phi_{s,t}(x,1)$ we have

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$f(\eta_{s,t}(x))\,\tilde{\phi}_{s,t}(x,1) = f - \sum_{j=1}^{d}\int_s^t Y_j(r) f(\eta_{r,t}(x))\,\tilde{\phi}_{r,t}\circ \hat{d}y_r^k$$

$$- \sum_{j=1}^{d}\int_s^t h_k(r) f(\eta_{r,t}(x))\,\tilde{\phi}_{r,t}\circ \hat{d}y_r^k$$

$$= f - \sum_{j=1}^{d}\int_s^t M_k(r)\, f(\eta_{r,t}(x))\,\tilde{\phi}_{r,t}\circ \hat{d}y_r^k.$$

Interchanging the forward variable s and backward variable t using (Theorem 7.D.2) we write

$$f(\eta_{s,t}(x))\,\tilde{\phi}_{s,t}(x,1) = f - \sum_{j=1}^{d}\int_s^t M_k(r)(f\cdot\eta_{s,r}\,\tilde{\phi}_{s,r})(x)\circ dy_r^k.$$

Differentiating with respect to t by fixing the initial variable s, we see that (4.2.10) is verified.

Next, we consider the inverse map to (4.2.15). The inverse map is expressed as

$$\mu_{s,t}(f)(x) = f(\eta_{s,t}^{-1}(x))\,\exp\sum_{j=1}^{d}\int_s^t h_k(r,\eta_{s,r}^{-1}(x))\circ dy_r^k$$

where the justification is given by Kunita [94, Lemma 1]. Moreover, the operator $\mu_{s,t}\,L(t)\,\mu_{s,t}^{-1}$ is well-defined (see Theorem 4.2.1) and is given by

$$\mu_{s,t}\,L(t)\,\mu_{s,t}^{-1} = \frac{1}{2}\sum_{j=1}^{m}(\mu_{s,t}\,\tilde{X}_j(t)\,\mu_{s,t}^{-1})^2 + \mu_{s,t}\,\tilde{X}_0(t)\mu_{s,t}^{-1} + \tilde{h}_0(t).$$

Next, we need to find an expression for $\mu_{s,t}\,\tilde{X}_j(t)\,\mu_{s,t}^{-1}$. Since

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$\mu_{s,t}\tilde{X}_j(t)\mu_{s,t}^{-1} = \eta_{s,t}^{-1}(x)\exp\{\sum_{k=1}^{d}\int_s^t h_k(r,\eta_{s,r}^{-1}(x))\cdot dy_r^k\}\tilde{X}_j(t)[\eta_{s,t}(x)\exp\{-\sum_{k=1}^{d}\int_s^t h_k(r,\eta_{r,t}(x))\cdot \hat{d}y_r^k\}]$$

$$= \eta_{s,t}^{-1}(x)\tilde{X}_j(t)(\eta_{s,t}(x))$$

$$+ \eta_{s,t}^{-1}(x)\ \{\exp\sum_{k=1}^{d}\int_s^t h_k(r,\eta_{s,r}^{-1}(x)\cdot dy_r^k\}\ X_j(t)(\exp\{\sum_{k=1}^{d} - \int_s^t h_k(r,\eta_{r,t}(x))\cdot \hat{d}y_r^k)\}\,\eta_{s,t}(x).$$

Using the identity $\xi_{s,t*}(X)f(x) = X(f\cdot\xi_{s,t})(\xi_{s,t}^{-1}(x))$ given by Kunita [4], where $(\xi_{s,t})_*$ is

by definition a linear map from $T_x\,(R^n)$ into $T_{\xi_{s,\,t}\,(x)}\,(R^{\,n})$ (see, Appendix D) it follows

that

$$\eta_{s,t}^{-1}(x)X_j(t)(\eta_{s,t}(x)) = \eta_{s,t*}(X_j(t))$$

and by Kunita [94, Lemma 1].

$$\exp\{\sum_{k=1}^{d}\int_s^t h_k(r,\eta_{s,r}^{-1}(x))\cdot dy_r^k\}X_j(t)(\exp\{-\sum_{k=1}^{d}\int_s^t h_k(r,\eta_{r,t}(x))\cdot \hat{d}y_r^k\}$$

$$= -\sum_{k=1}^{d}\int_s^t X_j(r)(h_k(r)\cdot\eta_{s,r})(z)\cdot dy_r^k\Big|_{z=\eta_{s,t}^{-1}(x)}.$$

Therefore, for almost all $\tilde{\omega}$, $\mu_{s,t}L(t)\mu_{s,t}^{-1}$ is a second-order operator written as

$$\mu_{s,t}L(t)\mu_{s,t}^{-1} = \frac{1}{2}\sum_{j=1}^{m}(\eta_{s,t*}(\tilde{X}_j(t)) - g_{s,t}^{j}(x))^2 + \eta_{s,t*}(\tilde{X}_0(t)) - g_{s,t}^{0}(x) + \tilde{h}_0(t)$$

where

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$g^j_{s,t}(x) = \sum_{k=1}^{d} \int_s^t \tilde{X}_j(r)(h_k(r)\cdot\eta_{s,r})(r,z)\cdot dy_r^k \Big|_{z=\eta_{s,t}^{-1}(x)}.$$

Next, we shall construct the solution to (4.2.11) similarly but instead of considering the solution $\xi_{s,t}$ of (4.1.39) in the forward direction this solution is constructed using the composition $\eta_{s,t}\cdot\zeta_{s,t}$ where $\zeta_{s,t}$ satisfies a backward SDE. Then, proceeding as in Theorem 4.2.1 one can obtain (4.2.11) by interchanging backward and forward variables to deduce an equation evolve in the forward variable. QED

Summarizing, we have established now that the solution maps $\bar{\mu}_t, \mu_t^{-1}$ satisfying (4.2.6), (4.2.10), respectively, are the adjoint stochastic differential equations adjoint to (4.2.4), (4.1.40), respectively. Likewise $\bar{v}_t, v_t^{-1}$ are the adjoint processes to (4.2.5), (4.1.41), respectively. Similar results for state-valued processes are established by Kunita in [95, Theorem 3].

In this thesis we assume the bounded conditions stated earlier. However, one is often concerned with control problems satisfying a weaker linear growth condition. We shall show that the decomposition established by (4.2.10), (4.1.41) holds in a weaker sense by proving existence of their solutions.

Lemma 4.2.2 The decomposition results of the current section can be extended to systems having Lipschitz and linear growth conditions as long as the coefficients of operators L(t), M(t) are of C^k class ($k > 2$), their first and second derivatives are bounded, and $\{M_k\}$, $\{\tilde{X}_j\}_{j=1}^m$ have bounded coefficients.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Proof: We shall only prove existence results. Let us consider the decomposition $\mu_{s,t}^{-1}(f)$ of Theorem 4.2.3. First we shall prove that the solution to (4.2.10) exists in L^2 sense.

Consider the operator $\mu_{s,t}^{-1}(f)$ given by (4.2.15) and backward SDE (4.2.14). Since all the coefficients of $L(t)$, $M_j(t)$ are of C^k class, $k > 2$ and their first derivatives are bounded, by Kunita [92, Thm. 2.3] $\hat{\eta}_{s,t}(.,\hat{\omega})$ is a diffeomorphism from R^n onto R^n of C^{k-1} class for any t a.s and there exists a unique strong solution $\hat{\eta}_{s,t}(x,\hat{\omega})$ satisfying (4.2.14). If we assume that f is a bounded function, then by applying the Schwartz inequality twice, we obtain

$$\tilde{E}_{\hat{\rho}}|\mu_{s,t}^{-1}(f)|^2 = \tilde{E}_{\hat{\rho}}|f(\hat{\eta}_{s,t}(x,\hat{\omega}))\exp\{-\sum_{k=1}^{d}\int_s^t h_k(r,\hat{\eta}_{r,t}(x))\cdot\hat{d}y_r^k\}|^2$$

$$\leq \tilde{E}_{\hat{\rho}}|f(\hat{\eta}_{s,t}(x,\hat{\omega}))|^2\tilde{E}_{\hat{\rho}}\exp\{-2\sum_{k=1}^{d}\int_s^t h_k(r,\hat{\eta}_{r,t}(x))\cdot\hat{d}y_r^k\}$$

$$\leq K(t,x)\,\tilde{E}_{\hat{\rho}}[\exp\{-2\sum_{k=1}^{d}\int_s^t h_k(r,\hat{\eta}_{r,t}(x))\cdot\hat{d}y_r^k\}]$$

$$= K(t,x)\,\tilde{E}_{\hat{\rho}}[\exp\{-2\sum_{k=1}^{d}\int_s^t h_k(r,\hat{\eta}_{r,t}(x))\hat{d}y_r^k-\sum_{k=1}^{d}\int_s^t Y_k(r)(h_k(r,\hat{\eta}_{r,t}(x))dr\}]$$

$$\leq K(t,x)\tilde{E}_{\hat{\rho}}[\exp\{-4\sum_{k=1}^{d}\int_s^t h_k(r,\hat{\eta}_{r,t}(x))\hat{d}y_r^k - 8\sum_{k=1}^{d}\int_s^t h_k^2(r,\hat{\eta}_{r,t}(x))dr\}]$$

$$\times\tilde{E}_{\hat{\rho}}[\exp\{-2\sum_{k=1}^{d}\int_s^d Y_k(r)(h_k(r),\hat{\eta}_{r,t}(x))dr+8\sum_{k=1}^{d}\int_s^t h_k^2(r,\hat{\eta}_{r,t}(x))dr\}]$$

$$\leq K(t,x)<\infty.$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

For fixed t, the first exponential function is an $\mathcal{F}_t^y \otimes \hat{\mathcal{F}}_t^w$-backward martingale whose expectation is 1. By Remark 7.E.2 the second expectation satisfies $\tilde{E}_{\hat{\rho}}(\cdot)<\infty$. Thus, $\mu_{s,t}^{-1}(f)(x)$ is an L^2 solution. Let us next, consider the measure-valued process $v_{s,t}(f)$ of (4.1.41). It can be shown (see Kunita [94]) that the solution $v_{s,t}$ of (4.1.41) is represented by

$$v_{s,t}(f)(x) = \tilde{E}_{\hat{\rho}}[f(\zeta_{s,t}(x))e^{-\sum_{j=1}^{m}\int_s^t g_{s,r}^j(\zeta_{s,r}(x))\cdot dw_r^j - \int_s^t g_{s,r}^0(\zeta_{s,r}(x))dr}] \tag{4.2.16}$$

where

$$\mu_{s,t}L(t)\mu_{s,t}^{-1} = \frac{1}{2}\sum_{j=1}^{d}(\eta_{s,t*}^{-1}(\tilde{X}_j) - g_{s,t}^j)^2 + (\eta_{s,t*}^{-1}(\tilde{X}_0) - g_{s,t}^0)$$

$$g_{s,t}^j(x) = g_{s,t}(\tilde{X}_j)(x) = \sum_{k=1}^{d}\int_s^t (\tilde{X}_j(r)h_k(r,n_{r,t}^{-1}(\cdot))(z)\,|_{z=\eta_{s,t}(x)}\cdot\hat{d}y_r^k \tag{4.2.17}$$

$$g_{s,t}^0(x) = g_{s,t}(\tilde{X}_0) - h_0(t,\eta_{s,t})$$

and $\zeta_{s,t}(x,\hat{\omega},\hat{\omega})$ is the solution to

$$d\zeta_t = \eta_{s,t*}^{-1}(\tilde{X}_0(t))(\zeta_t)dt + \sum_{j=1}^{m}\eta_{s,t*}^{-1}(\tilde{X}_j(t))(\zeta_t)\cdot dw_t^j,\ \zeta_s = x. \tag{4.1.18}$$

The composition $\eta_{s,t}\zeta_{s,t}$ provides a representation for the solution $\xi_{s,t}$ of (4.1.39). To prove existence of $v_{s,t}(f)$ we shall show that

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$\tilde{E}_{\tilde{\wp}}|v_{s,t}(f)(x,\tilde{\omega})|^2 < \infty \text{ a.s..}$$

But using Jensen's inequality,

$$\tilde{E}_{\tilde{\wp}}|v_{s,t}(f)(x,\tilde{\omega})|^2 \le \tilde{E}_{\tilde{\wp}\otimes\wp}[\,|f(\zeta_{s,t}(x))|^2\phi_{s,t}^2]$$

where,

$$\phi_{s,t} = e^{-\sum_{j=1}^{m}\int_s^t g_{s,r}^j(\zeta_{s,r}(x)\cdot dw_r^j - \int_s^t g_{s,r}^0(\zeta_{s,r}(x))dr}.$$

Thus, since f is bounded, it is sufficient to show that $\phi_{s,t}^2$ is integrable for all t and x.

Since the solutions $\eta_{s,t}$, $\xi_{s,t}$, $\zeta_{s,t}$ exist and f is a bounded function, writing the exponent of $\phi_{s,t}$ using Ito integral representation, we have

$$\tilde{E}_{\tilde{\wp}\otimes\wp}[\phi_{s,t}^2] = \tilde{E}_{\tilde{\wp}\otimes\wp}\left(e^{-2\sum_{j=1}^{t} g_{s,r}^j(\zeta_{s,r}(x))dw_r^j - 2\int_s^t g_{s,r}^0(\zeta_{s,r}(x))dr - \sum_{j=1}^{m}\int_s^t \eta_{s,r*}^{-1}(\tilde{X}_j(r))g_{s,r}^j(\zeta_{s,r}(x))dr_r^j}\right)$$

$$\le \tilde{E}_{\tilde{\wp}\otimes\wp}\left(e^{-4\sum_{j=1}^{m}\int_s^t g_{s,r}^j(\zeta_{s,r}(x))dw_r^j - 8\sum_{j=1}^{m}\int_s^t |g_{s,r}^j(\zeta_{s,r}(x))|^2 dr}\right)$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$\times \tilde{E}_{\tilde{\rho}\otimes\tilde{\rho}}\left(e^{-2\sum_{j=1}^{t} \eta_{s,r*}^{-1}(\tilde{X}_j(r))g_{s,r}^{j}(\zeta_{s,r}(x))dr - 4\int_s^t g_{s,r}^{0}(\zeta_{s,r}(x))dr + 8\sum_{j=1}^{m}\int_s^t |g_{s,r}^{j}(\zeta_{s,r}(x))|^2 dr}\right).$$

The first exponential is an $\mathcal{F}_t^y \otimes \mathcal{F}_t^w$-adapted supermartingale thus, a $\tilde{E}_{\tilde{\rho}\otimes\tilde{\rho}}(\cdot) \leq 1$. Since the remaining exponential term has an exponent which is of quadratic growth, then, by

$$- \sum_{j=1}^{m} \int_s^t \eta_{s,r*}^{-1}(\tilde{X}_j(r))g_{s,r}^{j}(\zeta_{s,r}(x))dr - 4\int_s^t g_{s,r}^{0}(\zeta_{s,r}(x))dr$$

$$\leq (t-s)\sum_{j=1}^{m}\int_s^t (\eta_{s,r}^{-1}(\tilde{X}_j(r)))^2 (g_{s,r}^{j}(\zeta_{s,r}(x)))^2 dr + 4(t-s)\int_s^t |g_{s,r}^{0}(\zeta_{s,r}(x))|^2 dr$$

it follows that $\phi_{s,t}^2(t)$ is integrable by a version of Remark 7.E.2. Uniqueness would also follows since the solutions $\zeta_{s,t}$, $\xi_{s,t}$, $\eta_{s,t}$ are unique. QED

4.3 APPROACH 1: STOCHASTIC MINIMUM PRINCIPLE

Consider the stochastic control problem (4.1.1) - (4.1.4). Using the decomposition (4.1.40), (4.1.41) the performance cost can be written in a separated form as

$$\begin{aligned} J(u(\cdot)) &= \tilde{E}_{\tilde{\rho}}\{\rho_{s,T}(\kappa(x)) + \int_s^T \rho_{s,r}(\mathrm{II}(r,x,u(r))dr\} \\ &= \tilde{E}_{\tilde{\rho}}\{\nu_{s,T}\mu_{s,T}(\kappa(x)) + \int_s^T \nu_{s,r}\mu_{s,r}(\mathrm{II}(r,x,u(r))dr\} \end{aligned} \tag{4.3.1}$$

where $\rho_{s,t} = \nu_{s,t}\, \mu_{s,t}$.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

CASE I

Consider the case when σ is independent of the control variable u. As stated in Section 1.2.2 $L^u(t)$ defined by (4.1.32) is the only operator which contains any explicit dependence on control u which is also evident from (4.1.33)-(4.1.35). Therefore, for this problem we have

$$\tilde{X}_j^u(t) \triangleq \sum_{k=1}^{m} \Theta^{jk} \sum_{i=1}^{n} \sigma_k^i(t,x)\frac{\partial}{\partial x^i} \tag{4.3.2}$$

$$\tilde{X}_0^u(t) \triangleq \sum_{i=1}^{n} f^i(t,x,u)\frac{\partial}{\partial x^i} - \sum_{k=1}^{d} h_k(t)Y_k(t),\ Y_k \triangleq \sum_{j=1}^{m}\sum_{i=1}^{n}\gamma^{kj}(t)\sigma_j^i(t,x)\frac{\partial}{\partial x^i}. \tag{4.3.3}$$

Suppose that $u(t)\in U_{ad}$ is an optimal control. Then since U_{ad} is a convex set (by assumption (A'10)) for any other control $\tilde{u}(t)\in U_{ad}$, $u(t) + \varepsilon\tilde{u}(t)$ is also in U_{ad} for each $t\in[0,T]$ and $\varepsilon\in[0, 1]$. Therefore we may construct a mapping ζ: $0 \le \varepsilon \le 1 \to U$ given by $\zeta(\varepsilon) \triangleq u + \varepsilon\tilde{u}$. This map $\zeta(\varepsilon)$ is called a weak variation of the control and is the map we shall consider for the current and next sections.

As a result of the decomposition of $\rho_{s,t}$ into (4.1.40), (4.1.41) any control variation would only affect the measure-valued process $v_{s,t}(f)$. This is easily seen since $\mu_{s,t}$ does not depend on the control explicitly but only $v_{s,t}$ does ($\mu_{s,t}(x,\hat{\omega})$ is independent of $\hat{\omega}\in\hat{\Omega}$). That is, $\mu_{s,t}$ depends on the initial state x and the observations $\mathcal{F}^y_{s,t}$. Therefore if we consider $u(t)$, $\tilde{u}(t)\in U_{ad}$ and denote by $v^{\varepsilon}_{s,t}$ the measure-valued process corresponding to control $u(\bullet) + \varepsilon\tilde{u}(\bullet)$,with $v^{\varepsilon}_{s,t}$ defined by (4.1.41), and with $u(\bullet)$ replaced by $u(\bullet) + \varepsilon\tilde{u}(\bullet)$, we have

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$dz_{s,t}(f) = z_{s,t}\,(\mu_{s,t}L^{u}(t)\mu_{s,t}^{-1}f)dt + v_{s,t}(\mu_{s,t}\frac{\partial L^{u}(t)}{\partial u}\tilde{u}(t)\mu_{s,t}^{-1}f)dt$$
$$\lim_{t\downarrow s} z_{s,t} = 0. \tag{4.3.4}$$

where

$$z_{s,t}(f) = \frac{\partial}{\partial\varepsilon} v_{s,t}^{\varepsilon}(f)\,|_{\varepsilon=0},$$

which can be verified by considering instead of $z_{s,t}$ the perturbed process $\rho_{s,t}^{B}$

defined by $\rho_{s,t}^{B}(f) \triangleq \frac{\partial}{\partial\varepsilon}\rho_{s,t}^{\varepsilon}(f)\,|_{\varepsilon=0}$.

Theorem 4.3.1

Let $u(\bullet)$, $\tilde{u}(\bullet) \in U_{ad}$. For $\varepsilon \in [0, 1]$ let $v_{s,t}^{\varepsilon}$ correspond to the solution of (4.1.41)

with control function $u(\bullet) + \varepsilon\tilde{u}(\bullet)$ and having initial condition δ_x. Then

$$v_{s,t}^{\varepsilon}(f) = v_{s,t}(f) + \varepsilon z_{s,t}(f) + o(t,\varepsilon) \tag{4.3.5}$$

where $z_{s,t}(f)$ is the solution to (4.3.4) and, as $\varepsilon \to 0$,

(i) $v_{s,t}^{\varepsilon}(f) \to v_{s,t}(f)$ in $L^2(\tilde{\Omega}\otimes\hat{\Omega}, \tilde{\mathcal{F}}_t^y\otimes\hat{\mathcal{F}}_t^w, \tilde{P}\otimes\hat{P})$,

(ii) $\dfrac{v_{s,t}^{\varepsilon}(f) - v_{s,t}(f)}{\varepsilon} \to z_{s,t}(f)$ in $L^2(\tilde{\Omega}\otimes\hat{\Omega}, \tilde{\mathcal{F}}_t^y\otimes\hat{\mathcal{F}}^w, \tilde{P}\otimes\hat{P})$,

(iii) the solution $z_{s,t}(f)$ of (4.3.4) exists and is unique.

Proof:

(i) We have

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$dv^{\varepsilon}_{s,t}(f) = v^{\varepsilon}_{s,t}(\mu_{s,t}L^{u+\varepsilon\tilde{u}}(t)\mu^{-1}_{s,t}f)dt$$

and by Kunita [94, 96]

$$v^{\varepsilon}_{s,t}(f) = \tilde{E}_{\hat{\rho}}[f(\zeta^{\varepsilon}_{s,t})e^{-\sum_{j=1}^{m}\int_s^t g^j_{s,r}(\zeta^{\varepsilon}_{s,r})\cdot dw^j_r - \int_s^t g^0_{s,r}(\zeta^{\varepsilon}_{s,r})}].$$

By Jensen's inequality

$$\tilde{E}_{\hat{\rho}}[|v^{\varepsilon}_{s,t}(f) - v_{s,t}(f)|^2] \le \tilde{E}_{\rho\otimes\hat{\rho}}[|f(\zeta^{\varepsilon}_{s,t})\phi^{\varepsilon}_{s,t} - f(\zeta_{s,t})\phi_{s,t}|^2]$$

where $\phi_{s,t}, \phi^{\varepsilon}_{s,t}$ are the exponential terms of $v^{\varepsilon}_{s,t}(f), v_{s,t}(f)$, respectively, defined by (4.2.16). By first expressing the F-S integrals in terms of Ito integrals and then using the Ito differential rule, we obtain

$$\begin{aligned} d\phi^{\varepsilon}_{s,t} = & -\sum_{j=1}^{m}\phi^{\varepsilon}_{s,t}g^j_{s,t}(\zeta^{\varepsilon}_{s,t})dw^j_t + \frac{1}{2}\sum_{j=1}^{m}\phi^{\varepsilon}_{s,t}|g^j_{s,t}(\zeta^{\varepsilon}_{s,t})|^2 dt \\ & -\frac{1}{2}\sum_{j=1}^{m}\phi_{s,r}\eta^{-1}_{s,r*}(\tilde{X}^{\varepsilon}_j(t))g^j_{s,t}(\zeta^{\varepsilon}_{s,t})dt - \phi^{\varepsilon}_{s,t}g^0_{s,t}(\zeta^{\varepsilon}_{s,t})dt, \end{aligned}$$

where $\zeta_{s,t}{}^{\varepsilon}$, $\zeta_{s,t}$ satisfy the stochastic differential equation (4.2.18). Since

$$\begin{aligned} f(\zeta^{\varepsilon}_{s,t})\phi^{\varepsilon}_{s,t} - f(\zeta_{s,t})\phi_{s,t} &= f(\zeta^{\varepsilon}_{s,t})\phi^{\varepsilon}_{s,t} + f(\zeta_{s,t})\phi^{\varepsilon}_{s,t} - f(\zeta_{s,t})\phi_{s,t} - f(\zeta_{s,t})\phi^{\varepsilon}_{s,t} \\ &= -[f(\zeta_{s,t}) - f(\zeta^{\varepsilon}_{s,t})]\phi^{\varepsilon}_{s,t} + f(\zeta_{s,t})[\phi^{\varepsilon}_{s,t} - \phi_{s,t}] \end{aligned}$$

then

$$|f(\zeta^{\varepsilon}_{s,t})\phi^{\varepsilon}_{s,t} - f(\zeta_{s,t})\phi_{s,t}| = |f(\zeta_{s,t})[\phi^{\varepsilon}_{s,t} - \phi_{s,t}] - [f(\zeta_{s,t}) - f(\zeta^{\varepsilon}_{s,t})]\phi^{\varepsilon}_{s,t}|.$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Thus,

$$\tilde{E}_{\hat{P}}[|v^{\varepsilon}_{s,t}(f) - v_{s,t}(f)|^2] \le \tilde{E}_{\tilde{P}\otimes\hat{P}}[|f(\zeta_{s,t})(\phi^{\varepsilon}_{s,t} - \phi_{s,t})|^2]$$
$$+ \tilde{E}_{\tilde{P}\otimes\hat{P}}[|\phi^{\varepsilon}_{s,t}[f(\zeta_{s,t}) - f(\zeta^{\varepsilon}_{s,t})]|^2].$$

By the bounded assumption on f, the above expression is further upper bounded by

$$\le \tilde{E}_{\tilde{P}\otimes\hat{P}}|\phi^{\varepsilon}_{s,t} - \phi_{s,t}|^2 + \tilde{E}_{\tilde{P}\otimes\hat{P}}|\phi_{s,t}|^4\tilde{E}_{\tilde{P}\otimes\hat{P}}|f(\zeta_{s,t}) - f(\zeta^{\varepsilon}_{s,t})|^4$$

But

$$|f(\zeta_{s,t} - f(\zeta_{s,t})^{\varepsilon})|^q \rightarrow 0 \ \tilde{P}\otimes\hat{P} \text{ a.s.}$$
$$|\phi^{\varepsilon}_{s,t} - \phi_{s,t}|^q \qquad \rightarrow 0 \ \tilde{P}\otimes\hat{P} \text{ a.s.}$$

(i.e., in some L^q space, $q \ge 1$) due to the bounded assumption on $\tilde{X}^i_j$, $\tilde{h}_0$, h_i.

Therefore, we conclude

$$\tilde{E}_{\tilde{P}\otimes\hat{P}}|v^{\varepsilon}_{s,t}(f) - v_{s,t}(f)|^2 \rightarrow 0.$$

(ii) Let $v^{\varepsilon}_{s,t}(f)$ be the solution to (4.1.40) corresponding to the control $u(\cdot) + \varepsilon\tilde{u}(\cdot)$ and set

$$\tilde{v}^{\varepsilon}_{s,t}(f) = \frac{v^{\varepsilon}_{s,t}(f) - v_{s,t}(f)}{\varepsilon} - z_{s,t}(f),$$
$$\tilde{v}^{\varepsilon}_{s,s}(f) = 0.$$

Then,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$d\tilde{v}^{\varepsilon}_{s,t}(f) = \frac{dv^{\varepsilon}_{s,t}(f) - dv_{s,t}(f)}{\varepsilon} - dz_{s,t}(f)$$

$$= \frac{v^{\varepsilon}_{s,t}(\mu_{s,t}L^{u+\varepsilon\tilde{u}}(t)\mu^{-1}_{s,t}f) - v_{s,t}(\mu_{s,t}L^{u}(t)\mu^{-1}_{s,t}f)}{\varepsilon}$$

$$- z_{s,t}(\mu_{s,t}L^{u}(t)\mu^{-1}_{s,t}f)dt - v_{s,t}(\mu_{s,t}\frac{\partial L^{u}(t)}{\partial u}\tilde{u}(t)\mu^{-1}_{s,t}f)dt.$$

If we replace $v^{\varepsilon}_{s,t}(f)$ by $\varepsilon\tilde{v}^{\varepsilon}_{s,t}(f) + \varepsilon z_{s,t}(f) + v_{s,t}(f)$, we get

$$\tilde{v}^{\varepsilon}_{s,t}(f) = \int_s^T \{\tilde{v}^{\varepsilon}_{s,t}(\mu_{s,t}L^{u+\varepsilon\tilde{u}}(t)\,\mu^{-1}_{s,t}f) + z_{s,t}(\mu_{s,t}[L^{u+\varepsilon\tilde{u}}(t) - L^{u}(t)]\mu^{-1}_{s,t}f)\}dt$$

$$+ \int_s^T \frac{v_{s,t}(\mu_{s,t}[L^{u+\varepsilon\tilde{u}}(t) - L^{u}(t) - \varepsilon\frac{\partial L^{u}(t)}{\partial u}\tilde{u}(t)]\mu^{-1}_{s,t}f)}{\varepsilon} dt.$$

By the bounded assumption on the coefficients of the operator $L^u(t)$ it follows that $L^u(t)$ satisfies a uniform Lipschitz condition, thus by (4.1.32) and (4.3.2), (4.3.3),

Then using the Gronwall inequality (see (Fleming and Rishel [56 , pp. 198]), $\tilde{v}^{\varepsilon}_{s,t}(f) \to 0$ in L^2 as $\varepsilon \to 0$.

(iii) Consider the composition $z_{s,t}\mu_{s,t}$. Then

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$d(z_{s,t}\mu_{s,t}(f)) = z_{s,t}\Big(\sum_{k=1}^{d}\mu_{s,t}(M_k(t)f)\cdot dw_t^k\Big)$$
$$+\ z_{s,t}(\mu_{s,t}L^{u}(t)\mu_{s,t}^{-1}\mu_{s,t}f) + v_{s,t}(\mu_{s,t}\frac{\partial L^{u}(t)}{\partial u}\tilde{u}(t)\mu_{s,t}^{-1}\mu_{s,t}f)dt$$
$$= \sum_{k=1}^{d} z_{s,t}\mu_{s,t}(M_k(t)f)\cdot dw_t^k + z_{s,t}(\mu_{s,t}L^{u}(t)f)dt + v_{s,t}\mu_{s,t}(\frac{\partial L^{u}}{\partial u}\tilde{u}(t)f)dt.$$

Let $P_t = z_{s,t}\mu_{s,t}$; then

$$dP_t(f) = P_t(L^{u}(t)f)dt + \rho_t(\frac{\partial u}{\partial u}\tilde{u}(t)f)dt + \sum_{k=1}^{d} P_t(M_k(t)f)\cdot dw_t^k$$

and by treating f as a test function we deduce

$$dP_t = L^{u}(t)^{*}P_t dt + \frac{\partial L^{u}(t)^{*}}{\partial u}\rho_t\tilde{u}(t)dt + \sum_{k=1}^{d} M_k(t)^{*}P_t\cdot dw_t^k$$

where ρ_t is the unnormalized conditional density. Since f is bounded and $\tilde{u}\in L^2(\tilde{\Omega}, \tilde{\mathcal{F}}_t^y, \tilde{\mathcal{P}})$, it follows that $\rho_t(\frac{\partial L^{u}(t)}{\partial u}\tilde{u}(t)f)\in L^2(\tilde{\Omega}, \tilde{\mathcal{F}}_t^y, \tilde{\mathcal{P}})$. If we set

$$\bar{P}_t = P_t - \int_s^t \rho_r(\frac{\partial L^{u}(r)}{\partial u}\tilde{u}(r)f)dr$$

then

$$d\bar{P}_t(f) = P_t(L^{u}(t)f)dt + \sum_{k=1}^{d} P_t(M_k(t)f)\cdot dw_t^k$$

and $\bar{P}_t(f)$ has a unique L^2 solution (see Kunita [93, 94, 96], Pardoux [112, 114]). Then, it also follows that $P_t(f)$ has a unique solution in L^2 sense. But since $P_t(f)$ is considered as a weak generalized solution of P_t, by the composition above we have $z_{s,t} = P_t\mu_{s,t}^{-1}$ which also has a solution in L^2 sense since $\mu_{s,t}^{-1}$ does (see Lemma 4.2.2). QED

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

If the Gateaux derivative of J(•) as a function on the Hilbert space $L_y^2(\tilde{\Omega}\otimes\tilde{\Omega}, \tilde{\mathcal{F}}_t^y\otimes\tilde{\mathcal{F}}_t^y, \tilde{\wp}\otimes\tilde{\wp})$ is $\mathcal{F}_{s,t}^y$-adapted and well-defined, we have the following lemma.

<u>Lemma 4.3.1</u> The cost function J is Gateaux differentiable and satisfies

$$\begin{aligned}\frac{d}{d\varepsilon} J(u(\cdot) + \varepsilon\tilde{u}(\cdot))\big|_{\varepsilon=0} &= \tilde{E}_{\tilde{\wp}}\{z_{s,T}\mu_{s,T}(\kappa(x)) \\ &+ \int_s^T [z_{s,r}\mu_{s,r}(\pi^u(r)) + v_{s,r}\mu_{s,r}(\frac{\partial\pi^u}{\partial u}(r)\tilde{u}(r))]dr.\}\end{aligned} \tag{4.3.6}$$

<u>Proof:</u> Denote by $\tilde{J}$ the right side of (4.3.6). We then have

$$\begin{aligned}\frac{J(u(\cdot)+\varepsilon\tilde{u}) - J(u(\cdot))}{\varepsilon} - \tilde{J} &= \frac{1}{\varepsilon}\tilde{E}_{\tilde{\wp}}\{v^\varepsilon_{s,T}\mu_{s,T}(\kappa(x)) + \int_s^T v^\varepsilon_{s,r}\mu_{s,r}\pi^{u+\varepsilon\tilde{u}}(r)dr\} \\ &-\frac{1}{\varepsilon}\tilde{E}_{\tilde{\wp}}\{v_{s,T}\mu_{s,T}(\kappa(x)) + \int_s^T v_{s,r}\mu_{s,r}\pi^u(r)dr\} \\ &- \tilde{E}_{\tilde{\wp}}\{z_{s,T}\mu_{s,T}(\kappa(x)) + \int_s^T [z_{s,r}\mu_{s,r}(\pi^u(r)) + v_{s,r}\mu_{s,r}\frac{\partial\pi^u(r)}{\partial u^\varepsilon}\tilde{u}(r))]dr\}.\end{aligned}$$

Replacing $v^\varepsilon_{s,t}$ by $\varepsilon\tilde{v}^\varepsilon_{s,t} + \varepsilon z_{s,t} + v_{s,t}$, the right side of the preceding equation becomes

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$\frac{1}{\varepsilon}\tilde{E}_{\tilde{\rho}}\{(\varepsilon\tilde{v}^{\varepsilon}_{s,T}+\varepsilon z_{s,T}+v_{s,T})\mu_{s,T}(\kappa(x))+\int_s^T[(\varepsilon\tilde{v}^{\varepsilon}_{s,r}+\varepsilon z_{s,r}+v_{s,r})\mu_{s,t}(\pi^{u+\varepsilon\tilde{u}}(r))]dr\}$$

$$-\frac{1}{\varepsilon}\tilde{E}_{\tilde{\rho}}\{v_{s,T}\mu_{s,T}(\kappa(x))+\int_s^T v_{s,r}\mu_{s,r}\pi^{u}(r)dr\}$$

$$-\tilde{E}_{\tilde{\rho}}\{z_{s,T}\mu_{s,T}(\kappa(x))+\int_s^T[z_{s,t}\mu_{s,t}(\kappa^{u}(r))+v_{s,r}\mu_{s,r}(\frac{\partial\pi^{u}(r)}{\partial u^{\varepsilon}}\tilde{u}(r))]dr\}$$

$$=\tilde{E}_{\tilde{\rho}}\int_s^T\tilde{v}_{s,r}\mu_{s,r}(\pi^{u+\varepsilon\tilde{u}}(r)-\pi^{u}(r)-\varepsilon\frac{\partial\pi^{u}(r)}{\partial u^{\varepsilon}}\tilde{u}(r)dr$$

$$+\tilde{E}_{\tilde{\rho}}\int_s^T z_{s,r}\mu_{s,r}(\pi^{u+\varepsilon\tilde{u}}(r)-\pi^{u}(r))dr+\tilde{E}_{\tilde{\rho}}\int_s^T\tilde{v}^{\varepsilon}_{s,r}\mu_{s,r}(\pi^{u+\varepsilon\tilde{u}}(r))dr.$$

Letting ε tend to zero we notice that the first and second terms of the right side of the previous expression tend to zero due to the bounded assumption on $\frac{\partial\pi^{u}}{\partial u}$. The third term also tends to zero due to Theorem 4.3.1 (ii). QED

Remark 4.3.2 The cost function (4.3.6) and the perturbed process $z_{s,t}$ are related to the cost function and perturbed process given by Bensoussan [7, 10, 15] through the composition:

$$\rho^{B}_{s,t}(f)=z_{s,r}\mu_{s,T}(f)$$

which satisfies

$$d\rho^{B}_{s,t}(f)=\rho_{s,t}(\frac{\partial L^{u}(t)}{\partial u}\tilde{u}(t)f)+\rho^{B}_{s,t}(L^{u}(t)f)+\sum_{k=1}^{d}\rho^{B}_{s,t}(M_k(t)f)\cdot dy_t^{k}. \tag{4.3.7}$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Indeed if we set $M_k(t) = h_k(t)$, and consider the stochastic control problem treated by Bensoussan [10] the measure-valued process $\rho^B_{s,t}(x,\omega)$ corresponds to the perturbed process considered by Bensoussan [10, eqn. 2.7]. The Gauteaux differential of $J(\cdot)$ is exactly the one considered by Bensoussan [10, eqn. 2.9] since, by definition, the composition $z_{s,t}\,\mu_{s,t}$ is equal to $\rho^B_{s,t}$.

MINIMUM PRINCIPLE

At this point we introduce the measure-valued process $P_{s,t}(x,\omega)$ defined by

$$\begin{aligned} dP_{s,t}(x,\omega) &= -(\mu_{s,t}L^u(t)\mu_{s,t}^{-1})P_{s,t}(x,\omega)dt - \mu_{s,t}\pi^u(t)(x,\omega)dt \\ \lim_{t\uparrow T} P_{s,t} &= \mu_{s,T}(\kappa(x)) \end{aligned} \tag{4.3.8}$$

The choice of the measure-valued process satisfying (4.3.8) is obtained by viewing the perturbed process satisfying (4.3.4) and performance cost (4.3.1) as the deterministic analog of the control problem given by Fleming and Rishel [56] for state-valued processes. In fact the homogeneous part of (4.3.8) should be the adjoint operator to the homogeneous part of (4.3.4). Note the striking similarity that exists between the Lagrange multiplier as defined for deterministic systems and that of (4.3.8); for example, the term $\mu_{s,t}\,\pi^u(t)$ corresponds to the integral cost and the final condition corresponds to the terminal cost.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Lemma 4.3.2 The measure-valued process $P_{s,t}(x,\hat{\omega})$ satisfying (4.3.8) is the adjoint process to the measure-valued process $z_{s,t}(f)(x,\hat{\omega})$ satisfying (4.3.4). Moreover, there exists a unique solution $P_{s,t}(x,\hat{\omega})$ to (4.3.8) in L^2 sense.

Proof: The first part follows by the composition of (4.3.4) and (4.3.8) which will be shown shortly. The second part follows by defining

$$\bar{P}_t = P_{s,t}(x,\hat{\omega}) + \int_s^t \mu_{s,r}\Pi^u(r)dr$$

and differentiating to deduce

$$\begin{aligned} d\bar{P}_t &= -(\mu_{s,t}L^u(t)\mu_{s,t}^{-1})P_{s,t}(x,\hat{\omega})dt \\ P_T &= \mu_{s,T}(\kappa(x)) \end{aligned}$$

which is a parabolic type partial differential equation. It is well-defined for almost all $\hat{\omega}$ and has a unique solution (see Kunita [96]) or Theorem 4.3.1 (iii). QED

The first part of Lemma 4.3.2 will be evident in the next lemma (i.e., the homogeneous parts of (4.3.4), (4.3.8) are adjoint).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Lemma 4.3.3 The variational cost (4.3.6) can be expressed as

$$\frac{d}{d\varepsilon}J(u(\cdot)+\varepsilon\tilde{u}(\cdot))\big|_{\varepsilon=0} = \tilde{E}_{\tilde{\rho}}\{\int_s^T [v_{s,r}\mu_{s,r}(\frac{\partial \Pi^{u}(r)}{\partial u}\tilde{u}(r))$$

$$+ v_{s,r}(\mu_{s,r}\frac{\partial L^{u}(r)}{\partial u}\tilde{u}(r)\mu_{s,r}^{-1}P_{s,r}(x))dr\}$$

$$= \tilde{E}_{\tilde{\rho}}\{\int_s^T \langle \frac{\partial \Pi^{u}(r)}{\partial u},\rho_t\rangle \tilde{u}(r)dr + \int_s^T \langle \frac{\partial L^{u}(r)}{\partial u}\hat{P}_r,\rho_r\rangle \tilde{u}(r)dr\} \tag{4.3.9}$$

where

$$\hat{P}_t(x,\bar{\omega}) \triangleq \tilde{E}_{\tilde{\rho}}[\tilde{P}_t(x,\bar{\omega})\,|\,\mathcal{F}^{y}_{s,t}] \quad ,\tilde{P}_t(x,\bar{\omega}) \triangleq \mu_t^{-1}P_t(x,\bar{\omega}).$$

Proof: We start by applying the Ito differential rule to the composition $z_{s,t}(P_{s,t}(x,\bar{\omega}))$, thus,

$$z_{s,T}(P_{s,T}(x,\bar{\omega}) = z_{s,s}(P_{s,s}(x,\bar{\omega})) + \int_s^T z_{s,r}(dP_{s,r}(x,\bar{\omega})) + \int_s^T dz_{s,r}(P_{s,r}(x,\bar{\omega})).$$

Using (4.3.4) and (4.3.8) we have

$$z_{s,T}\mu_{s,T}(\kappa(x)) = -\int_s^T z_{s,r}((\mu_{s,r}L^{u}(r)\mu_{s,r}^{-1})P_{s,r}(x,\bar{\omega}))dr - \int_s^T z_{s,r}(\mu_{s,r}\Pi^{u}(r))dr$$

$$+\int_s^T z_{s,r}(\mu_{s,r}L^{u}(r)\mu_{s,r}^{-1}P_{s,r}(x,\bar{\omega})) + \int_s^T v_{s,r}(\mu_{s,r}\frac{\partial L^{u}(r)}{\partial u}\tilde{u}(r)\mu_{s,r}^{-1}P_{s,r}(x,\bar{\omega}))dr$$

which is equivalent to

$$z_{s,T}\mu_{s,T}(\kappa(x)) = \int_s^T v_{s,r}(\mu_{s,r}\frac{\partial L^{u}(r)}{\partial u}\tilde{u}(r)\mu_{s,r}^{-1}P_{s,r}(x,\bar{\omega}))dr - \int_s^T z_{s,r}(\mu_{s,r}\Pi^{u}(r))dr.$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Substituting the last equation into (4.3.6), the variational cost is expressed as

$$\frac{d}{d\varepsilon}J(u(\cdot)+\varepsilon\tilde{u}(\cdot))\big|_{\varepsilon=0}=E_{\tilde{\wp}}\{\int_s^T v_{s,r}(\mu_{s,r}\frac{\partial L^{u(r)}}{\partial u}\tilde{u}(r)\mu_{s,r}^{-1}P_{s,r}(x,\tilde{\omega}))dr+\int_s^T v_{s,r}\mu_{s,r}(\frac{\partial\pi^{u(r)}}{\partial u}\tilde{u}(r))dr.$$

Notice that $\rho_{s,t} = v_{s,t}\mu_{s,t}$ which implies

$$\frac{d}{d\varepsilon}J(u(\cdot)+\varepsilon\tilde{u}(\cdot))\big|_{\varepsilon=0}=E_{\tilde{\wp}}\{\int_s^T \rho_{s,t}(\frac{\partial L^{u(r)}}{\partial u}\tilde{u}(r)\mu_{s,r}^{-1}P_{s,r}(x,\tilde{\omega}))dr+\int_s^T \rho_{s,r}(\frac{\partial\pi^{u(r)}}{\partial u}\tilde{u}(r))dr.$$

Next, suppose we define $\tilde{P}_t \triangleq \mu_{s,t}^{-1}P_{s,t}(x,\tilde{\omega})$ where, as shown in Theorem 4.2.3 $\mu_t^{-1}(f)$ satisfies (4.2.10) and is the inverse map to $\mu_t(f)$ (also the adjoint map to $\mu_t(f)$). By applying the F-S differential rule to the composition $\mu_t^{-1}P_t$ (dropping the time index s) and using Theorem 4.2.3 we have,

$$\mu_T^{-1}(P_T(x,\tilde{\omega}))=\mu_T^{-1}(\mu_T(\kappa(x)))=\mu_s^{-1}(P_s(x,\tilde{\omega}))+\int_s^T d\mu_r^{-1}(P_r(x,\tilde{\omega}))+\int_s^T \mu_r^{-1}(dP_r(x,\tilde{\omega}))$$

$$\begin{aligned}\kappa(x)-\mu_s^{-1}(P_s(x,\tilde{\omega})) &= -\sum_{k=1}^{d}\int_s^T M_k(r)\mu_r^{-1}(P_r(x,\tilde{\omega}))\cdot dy_r^k-\int_s^T \mu_r^{-1}((\mu_r L^{u(r)}\mu_u^{-1})P_r(x,\tilde{\omega}))dr\\ &\quad -\int_s^T \mu_r^{-1}(\mu_r\pi^{u(r)})dr.\end{aligned}$$

If we differentiate with respect to the variable s we obtain

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$d\tilde{P}_t(x,\tilde{\omega}) + L^u(t)\tilde{P}_t(x,\tilde{\omega})dt + \pi^u(t)dt + \sum_{k=1}^{d} M_k(t)\tilde{P}_t(x,\tilde{\omega}) \cdot \hat{d}y_t^k = 0$$
$$\lim_{t\uparrow T} \tilde{P}_t(x,\tilde{\omega}) = \kappa(x) \qquad (4.3.11)$$

which is the backward SDE given in Theorem 2.4.1. Therefore, we can rewrite the variational cost as

$$\begin{aligned}\frac{d}{d\varepsilon}J(u(\cdot)+\varepsilon\tilde{u}(\cdot))\,|_{\varepsilon=0} &= \tilde{E}_{\tilde{\wp}} \int_s^T \langle \frac{\partial \pi^u}{\partial u}\tilde{u}(r), \rho_r \rangle dr \\ &+ \int_s^T \frac{\partial}{\partial u}\langle L^u(r)\tilde{u}(r)\tilde{P}_r, \rho_r\rangle dr\} \\ &= \tilde{E}_{\tilde{\wp}}\{\int_s^T [\langle \frac{\partial \pi^u}{\partial u}, \rho_r\rangle + \frac{\partial}{\partial u}\langle L^u(r)\tilde{P}_r, \rho_r\rangle]\tilde{u}(r)dr \qquad (4.3.12)\end{aligned}$$

where the last equality follows since $\tilde{u}(\cdot)$ is $\tilde{\mathcal{F}}^y_{s,t}$-adapted. Next we make use of the conditional optimality given by Striebel [123 , Chp. 4] which states that whenever $u \in U_{ad}$ is conditionally optimal, it is also optimal. Thus, reconditioning (4.3.12) on the filtration $\mathcal{F}^y_{s,T}$ we can write

$$\frac{d}{d\varepsilon}J(u(\cdot)+\varepsilon\tilde{u}(\cdot))\,|_{\varepsilon=0} = \tilde{E}_{\tilde{\wp}}\{\int_s^T [\langle \frac{\partial \pi^u(r)}{\partial u}, \rho_r\rangle + \frac{\partial}{\partial u}\langle L^u(r)\hat{P}_r\rho_r\rangle]\,\tilde{u}(r)dr\}. \qquad (4.3.13)$$

where

$$\hat{P}_t(x,\tilde{\omega}) = \tilde{E}_{\tilde{\wp}}[(\tilde{P}_t(x,\tilde{\omega})\,|\,\mathcal{F}^y_{s,t}]. \quad \text{QED}$$

The form of the variational cost established by (4.3.12) (or (4.3.13)) is exactly the one obtained by Bensoussan [10, Sect. 2.3,] and [7, 15]. The backward SDE satisfying

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

(4.3.11) and the process $\hat{P}_t(x,\tilde{\omega})$ are exactly the ones obtained by the above author in [10, Lemma 2.2] and also in [15].

The following theorem presents the necessary condition for choosing the control to optimize the cost function.

Theorem 4.3.2

Let U_{ad} be a convex set. Suppose $u(\bullet)$ is an optimal control for the problem

$$J(v(\cdot)) = \tilde{E}_{\tilde{\wp}}\{\rho_{s,T}(\kappa(x)) + \int_s^T \rho_{s,r}(\Pi^{u}(r))dr\} \tag{4.3.14}$$

$$d\rho_{s,t} = L^u(t)^*\rho_{s,t}dt + \sum_{k=1}^{d} M_k(t)^*\rho_{s,t}\bullet dy_t^k,\ \rho_{s,s} = p_0. \tag{4.3.15}$$

Then there exists a unique $\hat{P}_t(x,\tilde{\omega})\in L_y^2$ such that the following condition is satisfied:

$$\sum_{j=1}^{k}(\tilde{u}_j(t)-u_j(t))[\int_{R^n}\{\frac{\partial\Pi}{\partial\tilde{u}_j}(t,x,u(t))+\frac{\partial}{\partial\tilde{u}_j}L^{u}\hat{P}_t(x,\tilde{\omega})\}\rho_t(x,\tilde{\omega})]dx\geq 0 \tag{4.3.16}$$

for all $\tilde{u}\in U_{ad}$ a.e. in t, a.s.. The Hamiltonian function $H_t(\rho_t, \hat{P}_t, u_t)$ is given by

$$\frac{\partial}{\partial\tilde{u}}H(\rho_t,\hat{P}_t,\tilde{u}_t) = \frac{\partial}{\partial\tilde{u}}\int_{R^n}[\Pi^{\tilde{u}}(t)+L^{\tilde{u}}\hat{P}_t]\rho_t dx = \frac{\partial}{\partial\tilde{u}}\{\langle\rho_t,\Pi^{\tilde{u}}(t)\rangle+\langle\hat{P}_t,L^{\tilde{u}}(t)^*p_t\rangle\} \tag{4.3.17}$$

Proof: First we start by showing existence of $\hat{P}_t(x,\tilde{\omega})$. Thus, by Jensen's inequality

$$\tilde{E}_{\tilde{\wp}}|\hat{P}_t(x,\tilde{\omega})|^2 = \tilde{E}_{\tilde{\wp}}|\tilde{E}_{\tilde{\wp}}(\tilde{P}_t(x,\tilde{\omega})|\mathcal{F}_{s,t}^y)|^2 \leq \tilde{E}_{\tilde{\wp}}|\tilde{P}_t(x,\tilde{\omega})|^2.$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

But $\tilde{P}_t(x,\tilde{\omega})$ is the solution to (4.3.11) which can be shown to have a unique solution by defining $\bar{P}_t(x,\tilde{\omega}) = \tilde{P}_t(x,\tilde{\omega}) - \int_t^T \Pi^u(r)dr$ and proceeding as in Lemma 4.3.2. Next we shall prove (4.3.16). Since U_{ad} is convex and $\tilde{u}(\bullet) \in U_{ad}$ it follows that $u(\bullet) + \varepsilon(\tilde{u}(\bullet) - u(\bullet))$ is also admissible and

$$J(u(\cdot) + \varepsilon(\tilde{u}(\cdot) - u(\cdot))) \geq J(u(\cdot)).$$

Thus, we can write

$$\frac{d}{d\varepsilon} J(u(\cdot) + \varepsilon(\tilde{u}(\cdot) - u(\cdot)))\,|_{\varepsilon=0} \geq 0.$$

Therefore using the same procedure as before if follows that

$$\tilde{E}_{\tilde{\wp}} \int_s^T \sum_{j=1}^{k} (\tilde{u}_j(t) - u_j(t)) [\int_{R^n} \{ \frac{\partial \Pi}{\partial \tilde{u}_j}(t,x,u(t)) + \frac{\partial}{\partial \tilde{u}_j} L^u \tilde{P}_t(x,\tilde{\omega}) \} \rho_t(x,\tilde{\omega}) \} dx \geq 0$$

for all $\tilde{u}(\bullet) \in U_{ad}$, as shown by Fleming and Rishel [56, Them. 11.2, pp. 41] for deterministic systems. The rest of the proof is a consequence of the proof presented by Bensoussan [9, Chp. VI, Them. 1.2, pp. 232-234]. Suppose $t_0 \in (0, T)$ and take $\tilde{u}$ small enough so that $t_0 + \theta < T$. Choose $\tilde{u}_0 \in L^2$ taking values in $C((0, T); R^d)$ where $\tilde{u}_0 \in U_{ad}$ a.s. Take

$$\tilde{u}(t) = u(t) \text{ for } t \notin (t_0, t_0 + \theta)$$
$$\tilde{u}(t) = v_0 \text{ for } t \in (t_0, t_0 + \theta)$$

which is again admissible. Define the stochastic process $\lambda_j(t)$ by

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$\lambda_j(t) \triangleq \int_{R^n} \{\frac{\partial\Pi}{\partial\tilde{u}_j}(t,x,u(t)) + \frac{\partial}{\partial\tilde{u}_j} L^u \hat{P}_t(x,\tilde{\omega})\}\rho_t(x,\tilde{\omega})dx.$$

$$\lambda_j(t) \triangleq \int_{R^n} \{\frac{\sigma\Pi}{\partial\tilde{u}_j}(t,x,u(t)) + \frac{\sigma}{\partial\tilde{u}_j} L^u \hat{P}_t(x,\tilde{\omega})\}\rho_t(x,\tilde{\omega})dx.$$

Then we deduce that

$$\tilde{E}_{\tilde{\rho}} \sum_{j=1}^{k} v_{\theta j} \int_{t_0}^{t_0+\theta} \lambda_j(t)dt - \tilde{E}_{\tilde{\rho}} \int_{t_0}^{t_0+\theta} \sum_{j=1}^{k} u_j(t)\lambda_j(t)dt \geq 0.$$

If we consider $\lambda_j(t)$ as a process in L^2 space we have

$$\frac{1}{\theta} \int_{t_o}^{t_o+\theta} \lambda_j(t)dt \to \lambda_j(t_o) \text{ in } L^2 \text{ sense a.e. } t_0.$$

Therefore, we obtain

$$\tilde{E}_{\tilde{\rho}} \sum_{j=1}^{k} (\tilde{u}_{0j} - u_j(t_0))\, \lambda_j(t_0) \geq 0. \quad \text{a.e. } t_0 \in (0,T).$$

Next, choose $\tilde{u}$ to be deterministic in U_{ad} and set

$$\gamma(t_0) = (\tilde{u} - u(t_0))\, \lambda_j(t_0) \text{ which is } \tilde{\mathcal{F}}^y_{t_0}\text{-measurable.}$$

Also set $A = \{\tilde{\omega};\, \gamma(t_0) < 0\}$ and take $\tilde{u}_0 = \tilde{u}$ in A, $\tilde{u}_0 = u(t_0)$ outside A. Then we deduce that $E_{\tilde{\rho}}(I_{\{\tilde{\omega};\gamma(t_0)\}}) \geq 0$ which is a contradiction, unless A has measure 0. QED

Remark 4.3.3 The variational cost (4.3.13) and Theorem 4.3.2 which provide the necessary conditions of optimality for deciding what control is optimal are the exact necessary condition obtained by Bensoussan [10, Thm. 2.1]. His treatment, however, is based on energy inequalities and linear contraction maps. Our justification of the existence of $\hat{P}$ as given Theorem 4.3.2 is based on the Pontryagin's minimum principle

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

when extended to measure-valued processes. Bensoussan [10, 15] considers the case when no control enters the diffusion coefficient and the correlation between the observation process and state process is zero. His early result on this subject [10] is based on the robust version of (4.3.11) defined by $u_{s,t}(x,\omega) \triangleq \tilde{P}_{s,t}(x,\omega)\, e^{y_t h(t,x)}$ and introduced in Chapter 2. Therefore, his formulation cannot be extended to the general case considered here since no such robust version exists when correlation between state process and observation process is allowed (due to h^i, h^j being non commutative). Very recently Bensoussan [15] provided a different approach to the existence and representation of the process $\hat{P}(x,\omega)$ which is based on Galerkin's approximations for Sobolev spaces to approximate the process $\hat{P}(x,\omega)$ using a finite-dimensional basis. Using this approximation he then derives a stochastic PDE satisfied by $\hat{P}(x,\omega)$ without having to introduce the robust version $u_{s,t}$ as done in Bensoussan [15].

Remark 4.3.4 We shall now present a formal derivation of the equation satisfied by the adjoint process $\hat{P}_t$ using (4.3.17). Since (4.3.17) provides the partial of H with respect to $\tilde{u}$, it follows that the Hamiltonian should be represented by

$$H \triangleq \langle \rho_t, \pi^{u}(t)\rangle + \langle \hat{P}_t, L^{\tilde{u}}(t)^* \rho_t \rangle + \tilde{r}_t$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

where $\tilde{r}_t$ is some process independent of $\tilde{u}$. From deterministic optimization theory we know that the dynamic constraint (i.e., the state derivative $\dot{x}$) is related to H by $\dot{x} = \frac{\partial H}{\partial q}$

where q is the adjoint process (Lagrange multiplier). Applying the last comment to optimization in L^2 space, the state of our system (in this case ρ_t) should satisfy $d\rho_t = \frac{\partial H}{\partial \hat{P}}$

where $\hat{P}$ is the Lagrange multiplier. But ρ_t satisfies the stochastic PDE

$$d\rho_t = A^{\tilde{u}}(t)^* \rho_t dt + \sum_{k=1}^{d} M_k(t)^* \rho_t dy_t^k,$$

Therefore the unknown process in the L^2 norm set up should be given by

$$\sum_{k=1}^{d} \langle r_t, M_k(t)^* \rho_t \rangle dy_t^k$$

where r_t is some $\mathcal{F}_t^y$-adapted process and r_t is interpreted as the second adjoint process. This statement will be made clearer shortly. In fact this is indeed the formulation we shall explore in the next section to give a formal derivation of the stochastic PDE satisfied by $\hat{P}_t$.

ADJOINT-PROCESS REPRESENTATION

Having obtained the result of Theorem 4.3.2, we shall determine a representation for the adjoint process using Bismut's [16] minimum principle for a system with complete

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

information and its extension by Kwakernaak [101] to systems with partial information having output feedback.

Bismut's Minimum Principle: Bismut's version of the stochastic minimum principle applies to minimization of the criterion

$$J = \tilde{E}\{\kappa(x_T) + \int_0^T \Pi^{u}(t)dt\} \tag{4.3.18}$$

for the system

$$dx_t = f(t,x_t,u_t)dt + \sigma(t,x_t,u_t)dw_t \tag{4.3.19}$$

where $\mathscr{F}_t \triangleq \sigma\{x_0, w_s; 0 \leq s \leq t\}$. The state process x_t, $t \in [0,T]$ is assumed to be measurable with respect to the filtration $\mathscr{F}_t$, $t \in [0, T]$, which results in an open loop control (i.e., not of feedback form) and system (4.3.19) is assumed to have Ito solutions. Bismut defines the Hamiltonian function by

$$H(t,x,u,q,r) = q_t f(t,x_t,u_t) + r_t \sigma(t,x_t,u_t) + \Pi^{u}(t). \tag{4.3.20}$$

The necessary conditions of optimality are then given by

$$dx_t = \frac{\partial H(t,x,q,r)}{\partial q}dt + \frac{\partial H(t,x,q,r)}{\partial r}dw_t \quad , x_0 \tag{4.3.21}$$

$$\begin{aligned} dq_t &= -\frac{\partial H(t,x,q,r)}{\partial x}dt - \frac{\partial H(t,x,q,r)}{\partial r}dw_t \\ q_T &= \frac{\partial \kappa}{\partial x_T} \end{aligned} \tag{4.3.22}$$

where q_t, r_t are $\mathscr{F}_t$-adapted processes. If the control minimizes (4.3.18) and $\{x_s, 0 \leq s \leq t\}$ is the solution to (4.3.19), then u_t minimizes $H(t,x,v,q,r)$ with respect to

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

ν for each $t \in [0, T]$. The above minimum principle is extended formally by Kwakernaak [101] to systems that are partially observed by defining (4.3.18) - (4.3.22) in an L^2 setting. His treatment however was completely formal, and he assumed that q_t, r_t are $\mathcal{F}_t^y$-adapted processes (output feedback).

Here we shall make use of the minimum principle given by Bismut [16] and Kwakernaak [101] to obtain the representation of the adjoint process $\hat{P}_t(x,\tilde{\omega})$ without assuming output feedback since by Theorem 4.3.2 $\hat{P}_t(x,\tilde{w})$ is an $\tilde{\mathcal{F}}_t^y$-adapted process. If we form the analog of the Hamiltonian function (4.3.20) using the Hamiltonian function given by (4.3.17) (Theorem 4.3.2), we deduce that

$$H_t(\rho_t, \hat{P}_t, r_t) = \langle \hat{P}_t, A^u(t)^* \rho_t \rangle + \sum_{k=1}^{d} \langle r_t^k, M_k(t)^* \rho_t \rangle + \langle \pi_t^u, \rho_t \rangle \qquad (4.3.23)$$

where $A^u(t)^*$ is the adjoint operator of $A^u(t)$ given by

$$A^u(t) = L^u(t) + \frac{1}{2} \sum_{k=1}^{d} M_k(t)^2. \qquad (4.3.24)$$

The operator $A^u(t)$ is obtained by writing (4.3.15) in terms of the Ito integral representation. Therefore, by casting (4.3.22) in an L^2 setting we deduce that the adjoint process $\hat{P}_t(x,\tilde{\omega})$ takes the form

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$- \hat{P}_t(x,\bar{\omega}) = \frac{\delta H_t}{\delta \rho}(\rho_t, \hat{P}_t, r_t) + \sum_{k=1}^{d} \frac{\delta H_t}{\delta(M_k^* \rho_t)}(\rho_t, \hat{P}_t, r_t) dy_t \tag{4.3.26}$$

$$\hat{P}_T(x,\bar{\omega}) = \kappa(x).$$

where the above derivative denotes a Fréchet derivative.

Thus,

$$d\hat{P}_t(x,\bar{\omega}) = -A^{u}(t)\hat{P}_t(x,\bar{\omega})dt - \mathrm{II}^{u}(t)dt - \sum_{k=1}^{d} M_k(t) r_t^k(\bar{\omega})dt - \sum_{k=1}^{d} r_t^k(\bar{\omega}) dy_t^k$$

which is the equation satisfied by the adjoint process defined by

$$\hat{P}_t \triangleq E_{\hat{P}}(\hat{P}_t \mid \mathcal{F}_t^y),$$

where $r_t^k(\bar{\omega})$ an $\mathcal{F}_t^y$-adapted process.

<u>Theorem 4.3.3</u> Assume the minimum principle of Theorem 4.3.2 is satisfied where the differential of the Hamiltonian function with respect to control $\tilde{u}(\bullet)$ is given by (4.3.17). Then the adjoint process $\hat{P}_t(x,\bar{\omega})$ is represented by

$$d\hat{P}_t(x,\bar{\omega}) + A^{u}(t)\hat{P}_t(x,\bar{\omega})dt + \mathrm{II}^{u}(t)dt + \sum_{k=1}^{d} M_k(t) r_t^k(\bar{\omega})dt + \sum_{k=1}^{d} r_t^k(\bar{\omega}) dy_t^k = 0 \tag{4.3.27}$$

$$\lim_{t \uparrow T} \tilde{P}_t(x,\bar{\omega}) = \kappa(x)$$

where $\hat{P}_t(x,\omega)$ has an L^2 solution, and $\{r_t^k(\bar{\omega})\}_{k=1}^{d}$ is an $\mathcal{F}_t^y$-adapted process. Moreover,the conditions of the minimum principle of Bismut (i.e., (4.2.20) - (4.2.22)) when represented in an L^2 space are satisfied.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Proof: It follows from Theorem 4.3.2, the definition of the Hamiltonian function (4.3.17), the definition of the adjoint process $\hat{P}_t(x, \hat{\omega})$, and the minimum principle conditions of Bismut. QED

CASE II

Consider the case when σ depends on the control variable u and the correlation between the state process and observation process is zero. This is indeed the case introduced in Section 1.2.2. That is, $M_k(t) \to h_k(t,x)$, $\sigma \to \sigma(t, x, u) = \sigma^u(t, x)$. It then follows from (4.3.2), (4.3.3) that

$$\tilde{X}_j^u(t) \to \sum_{i=1}^{n} \sigma_k^i(t, x, u)\frac{\partial}{\partial x^i} = X_j^u(t) \tag{4.3.28}$$

$$\tilde{X}_0^u(t) \to \sum_{i=1}^{n} f^i(t, x, u)\frac{\partial}{\partial x^i} = X_0^u(t) \tag{4.3.29}$$

$$Y_k(t) \to 0. \tag{4.3.30}$$

From the analysis of CASE I it follows that if we can show a similar theorem to Theorem 4.3.1 and a similar lemma to the Lemma 4.3.1 under the above conditions, the minimum principle and adjoint-process representation given by Theorem 4.3.2 and Theorem 4.3.3, respectively, remain valid under the current conditions.

Indeed if we choose the set of admissible control functions U_{ad} to consist of the $\mathcal{F}_t^y$, $t \in [0, T]$-adapted functions

$$u: [0, T] x C([0, T]; R_d) \to U$$

such that

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$|u(t, y_t)| \leq K(1 + ||y||_t), \quad ||y||_t = \sup\{|y_s|, s \in [0, T]\}$$

and assume that

$$\sup \int_0^T |\Pi(t, x, u)| dt < \infty, \quad |\Pi(t,x,u)| \in L^2,$$

then Theorem 4.3.1 and Theorem 4.3.2 remain valid by considering the sup norm $||\cdot||_t$.

Background information in proving Theorems 4.3.1, 4.3.2 under the above conditions are found in Haussmann [68] and Elliott and Yang [50].

Next, we shall derive the minimum principle of the partially observed system using a different approach which does not require the formal definition of the adjoint process, as presented by Bismut [16] and Kwakernaak [101], to obtain the backward equation (4.3.27).

4.4 APPROACH 2: STOCHASTIC MINIMUM PRINCIPLE

CASE I

In this section we shall derive the necessary conditions of optimality when the cost function is given by (4.1.3) subject to constraints (4.1.1), (4.1.2). That is, we treat the same problem stated in Section 4.1 but instead of using the approach given in Section 4.3 (specifically the formal derivation of 4.3.27), we adapt some results of deterministic control optimization found in Fleming and Rishel [56, Chp. 2, Thms. 11.1, 11.2] and of completely observed stochastic control optimization found in Bensoussan [9, Sect. 4, pp. 25-42]. However, our problem is described in terms of measure-valued processes rather

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

than state-valued processes considered by the above authors. The approach we consider is based on inverse maps and stochastic flows for measure-valued processes. Furthermore, we shall see that our method of solution will require the martingale representation theorem given Liptser and Shiryayev [103, Chp. 5, Thm. 5.7, pp. 167-170] when extended to processes adapted to the filtration generated by Wiener processes given also by Liptser and Shiryayev [103, Chp. 5, Thm. 5.8, pp. 171].

The important conclusion that one obtains from the result of this section in comparison to the results of the previous section is that the minimum principle given by Theorem 4.3.2 remains the same, but the adjoint process has a different representation.

We start by calling upon the procedure outlined in Section (1.1.2), Approach 2, Case (I). Assuming the cost function is described by (4.3.1), the state of the system is given by the two measure-valued processes $\nu_{s,t}$, $\mu_{s,t}$ satisfying (4.1.41), (4.1.40) respectively. The perturbed measure-valued process $z_{s,t}$ is the same as the one defined in Section 4.3, and given by (4.3.4) where the convergence conditions of Theorem 4.3.1 remain valid. We start with the following proposition.

<u>Proposition 4.4.1</u> The measure-valued process $z_{s,t}$ described by (4.3.4) can be expressed as

$$z_{s,t} = \psi_{s,t}\nu_{s,t} \tag{4.4.1}$$

where

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$\psi_{s,t}(f) = \int_s^t v_{s,r}(\mu_{s,r}\frac{\partial L^{u}(r)}{\partial u}\tilde{u}(r)\mu_{s,r}^{-1}v_{s,r}^{-1}(f))dr \qquad (4.4.2)$$
$$\lim_{t\downarrow s} \psi_{s,t} = 0.$$

Proof: Applying the F-S differential rule to the composition $z_{s,t}v_{s,t}^{-1}$ using (4.3.4) and (4.2.11) we deduce

$$z_{s,t}v_{s,t}^{-1}(f) = z_{s,s}\, v_{s,s}^{-1}(f) + \int_s^t z_{s,r}(dv_{s,r}^{-1}(f)) + \int_s^t dz_{s,r}(v_{s,r}^{-1}(f))$$

$$= -\int_s^t z_{s,r}((\mu_{s,r}L^{u}(r)\mu_{s,r}^{-1})v_{s,r}^{-1}(f))dr + \int_s^t z_{s,r}(\mu_{s,r}L^{u}(r)\mu_{s,r}^{-1}v_{s,r}^{-1}(f))dr + \int_s^t v_{s,r}(\mu_{s,r}\frac{\partial L^{u}}{\partial u}(r)\tilde{u}(r)\mu_{s,r}^{-1}v_{s,r}^{-1}(f))dr.$$

Since $\psi_{s,t} \triangleq z_{s,t}v_{s,t}^{-1}$ then

$$\psi_{s,t}(f) = \int_s^t v_{s,r}\,(\mu_{s,r}\,\frac{\partial L^{u}(r)}{\partial u}\tilde{u}(r)\mu_{s,r}^{-1}v_{s,r}^{-1}(f))dr$$

which is (4.4.2). Using $\rho_{s,t} = v_{s,t}\mu_{s,t}$ then (4.4.2) can be written as

$$\psi_{s,t}(f) = \int_s^t \rho_{s,r}(\frac{\partial L^{u}(r)}{\partial u}\tilde{u}(r)\mu_{s,r}^{-1}v_{s,r}^{-1}(f))dr. \quad \text{QED} \qquad (4.4.3)$$

Next, we express the variational cost function (4.3.6) of Lemma 4.3.1 in terms of the new measure-valued process $\psi_{s,t}$ which is well-defined and bounded since $\rho_{s,t}$, $\mu_{s,t}^{-1}$, $v_{s,t}^{-1}$ are all well-defined and bounded for any $f \in C_b^{\infty}(R^n)$. Thus, from (4.3.6)

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$\frac{dJ}{d\varepsilon}(u(\cdot) + \varepsilon\tilde{u}(\cdot))|_{\varepsilon=0} = \tilde{E}_{\mathfrak{P}}\{\psi_{s,T}\nu_{s,T}\mu_{s,T}(\kappa(x)) + \int_s^T [\psi_{s,r}\nu_{s,r}\mu_{s,r}(\Pi^u(r)) + \nu_{s,r}\mu_{s,r}(\frac{\partial\Pi^u(r)}{\partial u}\tilde{u}(r))]dr\}.$$

But $\rho_{s,t} = \nu_{s,t}\mu_{s,t}$, hence

$$\frac{dJ}{d\varepsilon}(u(\cdot) + \varepsilon\tilde{u}(\cdot)) = \tilde{E}_{\mathfrak{P}}\{\psi_{s,T}\rho_{s,T}(\kappa(x)) + \int_s^T [\psi_{s,r}\rho_{s,r}(\Pi^u(r)) + \rho_{s,r}(\frac{\partial\Pi^u(r)}{\partial u}\tilde{u}(r))]dr. \tag{4.4.4}$$

At this point we recognize that the process $\lambda_{s,t}$ defined by

$$\lambda_{s,T}(\tilde{\omega}) \triangleq \nu_{s,T}\mu_{s,T}(\kappa(x)) + \int_s^T \nu_{s,r}\mu_{s,r}(\Pi^u(r))dr \tag{4.4.5}$$

is an $\tilde{\mathcal{F}}^y_{s,T}$-adapted process having solution in $L^2(C([0,T];R^d))$ due to κ,Π^u being bounded functions.

Before we present the martingale representation theorem we shall need the following estimate.

Proposition 4.4.2 The function $\lambda_{s,T}(\tilde{\omega})$ satisfies $\tilde{E}_{\mathfrak{P}}|\lambda_{s,T}(\tilde{\omega})|^2<\infty$ a.s..

Proof: We recognize that

$$\lambda_{s,T}(\tilde{\omega}) = \tilde{E}_{\mathfrak{P}}\{\kappa(\xi_{s,T}(x))\phi_{s,T}(x,1) + \phi_{s,T}(x,1)\int_s^T \Pi(t,\xi_{s,r}(x),u(r))dr\} \tag{4.4.6}$$

where

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$\phi_{s,T}(x,1) \triangleq \exp\{\sum_{k=1}^{d}\int_s^T h_k(r,\xi_{s,r}(x))\cdot dy_r^k + \int_s^t \tilde{h}_0(r,\xi_{s,r}(x))^2 dr\}, \tag{4.4.7}$$

and $\xi_{s,t}(x)$ is the solution to (4.1.39). By Jensen's inequality

$$\tilde{E}_{\hat{\rho}}|\lambda_{s,T}(\tilde{\omega})|^2 \leq \tilde{E}_{\hat{\rho}\otimes\hat{\rho}}[\,|\kappa(\xi_{s,T}(x)) + \int_s^T \pi(t,\xi_{s,t}(x),u(t))dt|^2\,|\phi_{s,T}(x)|^2]$$

and the bounded assumptions on κ, π it is sufficient to show the integrability of $\phi^2_{s,t}(x,1)$

for all t and x a.s.. First write $\phi_{s,t}(x,1)$ using the Ito integral representation, thus $\phi^2_{s,t}(x,1)$ is represented by

$$\begin{aligned}\phi^2_{s,t}(x,1) &= \exp\{2\sum_{k=1}^{d}[\int_s^T h_k(r,\xi_{s,r}(x)dy_r^k - 2\int_s^T h_k^2(r,\xi_{s,r}(x))dr]\}\\ &\quad x\ \exp\{\int_s^T\{\sum_{k=1}^{d}[Y_k(r)h_k(r,\xi_{s,r}(x)) + 2h_k(r,\xi_{s,r}(x))]dr + 2\tilde{h}_0(r,\xi_{s,r}(x)\}dr\}.\end{aligned} \tag{4.4.8}$$

Since the integrand of the exponent of the second part of (4.4.8) is a bounded function, the second part is dominated by a constant depending on s and t only. The expectation of the first part is equal to one since it is a martingale with respect to $\tilde{\mathcal{F}}^y_{s,T}\otimes\hat{\mathcal{F}}^w_{s,T}$, hence $\phi^2_{s,T}(x,1)$

is integrable. QED

We could also show that if instead of assumptions (A'1) - (A'11) the weaker assumptions of Section 4.1 are satisfied which include the linear-quadratic problem, the following proposition holds.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Proposition 4.4.3 When the stochastic control problem under consideration satisfies assumptions (A1) - (A7) of Section 2.2 and $\sup_{s\le t\le T}\int_s^T |\pi(t,\xi_{s,t}(x),u(t,y_t)|^2 dt<\infty\ \rho_1$. a.s. then $\lambda_{s,T}(\tilde{\omega})$ is an integrable process.

Proof: Here we need to show that $\tilde{E}_{\tilde{\rho}}|\lambda_{s,T}(\tilde{\omega})|<\infty$. Using Jensen's inequality we have (by assuming $\sqrt{a}\le a$)

$$\tilde{E}_{\tilde{\rho}}|\lambda_{s,T}(\tilde{\omega})|\le \tilde{E}_{\tilde{\rho}\otimes\tilde{\rho}}|\kappa(\xi_{s,T}(x))\ \phi_{s,T}(x,1)+\phi_{s,T}(x,1)\int_s^T \pi(t,\xi_{s,t}(x),u(t)dt|$$

$$\le \tilde{E}_{\tilde{\rho}\otimes\tilde{\rho}}\{\ |\kappa(\xi_{s,T}(x))\ \phi_{s,T}(x,1)|+|\phi_{s,T}(x,1)\int_s^T \pi(t,\xi_{s,t}(x),u(t))dt|\}$$

$$\le \tilde{E}_{\tilde{\rho}\otimes\tilde{\rho}}|\kappa(\xi_{s,T}(x))|^2\tilde{E}_{\tilde{\rho}\otimes\tilde{\rho}}\ [\phi^2_{s,T}(x,1)]$$

$$+\tilde{E}_{\tilde{\rho}\otimes\tilde{\rho}}[\,|\int_s^T \pi(t,\xi_{s,t}(x),u(t)dt|^2]\tilde{E}_{\tilde{\rho}\otimes\tilde{\rho}}\ [\phi^2_{s,T}(x,1)].$$

Since $\xi_{s,T}(x)$ is a unique strong solution to (4.1.39) and the inequality (see assumption A6, Section 2.2)

$$|\kappa(\xi)|\le K(1+|\xi|^q)<q<\infty$$

is satisfied, it follows by Liptser and Shiryayev [103, Chp. 4, Thm. 4.6, pp. 128-130) that $\tilde{E}_{\tilde{\rho}\otimes\tilde{\rho}}|\kappa(\xi)|^{2m}<\infty$ for $m\ge 1$.

Therefore, as in Proposition 4.4.2 it is sufficient to show that $\tilde{E}_{\tilde{\rho}\otimes\tilde{\rho}}\ \phi^2_{s,T}\ (x,1)$ is integrable for all x, t. Using the same approach as in Lemma 4.2.2, by Jensen's inequality,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$\exp \int_s^T 2\{\sum_{k=1}^d (Y_k(r)h_k(r) + 4h_k(r)) + 2\tilde{h}_0(r)\}dr$$

$$= \exp \frac{1}{T-s}\int_s^T 2(T-s)\{\sum_{k=1}^d (Y_k(r)h_k(r) + 4h_k(r)) + 2\tilde{h}_0(r)\}dr$$

$$\leq \frac{1}{T-s}\int_s^T \exp 2(T-s)\{\sum_{k=1}^d (Y_k(t)h_k(t) + 4h_k(t)) + 2\tilde{h}_0(t)\}dt$$

Hence, if $(T - s) \leq 2\delta$, for some $\delta > 0$, then

$$\tilde{E}_{\hat{P}\otimes\hat{P}} \{\exp \int_s^T 2\{\sum_{k=1}^d (Y_k(r)\, h_k(r) + 4h_k(r)) + 2\tilde{h}_0(r)\}dr\}$$

$$\leq \sup_{0 \leq t \leq T} \tilde{E}_{\hat{P}\otimes\hat{P}} \exp\{4\delta\{\sum_{k=1}^d (Y_k(t)h_k(t) + rh_k(r)) + 2\tilde{h}_0(t)\}\} \leq \infty .$$

The last inequality follows from Liptser and Shiryayev [103, Chp. 4, Thm. 4.7, pp. 220] since by assumptions (A2), (A3) of Section 2.2, f, h satisfy a linear growth condition, σ is bounded, and the correlation between state process and observation process is bounded. QED

Returning to the process $\lambda_{s,T}(\tilde{\omega})$ defined by (4.4.5) let us define another process

$\chi_{s,T}(\tilde{\omega}) \sim \tilde{\mathcal{F}}^y_{s,T}$ by

$$\chi_{s,t}(\tilde{\omega}) \triangleq -\int_s^T v_{s,r}\mu_{s,r}\, \Pi^{u}(r)dr + \tilde{E}_{\hat{P}} [\lambda_{s,T}(\tilde{\omega}) \,|\, \tilde{\mathcal{F}}^y_{s,t}]. \tag{4.4.9}$$

where by (4.4.5)

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$\chi_{s,T}(\tilde{\omega}) = \nu_{s,T}\mu_{s,T}(\kappa(x)) = \rho_{s,T}(\kappa(x)). \tag{4.4.10}$$

Theorem 4.4.1 Suppose Proposition 4.4.2 or Proposition 4.4.3 holds. Let $\tilde{E}_{\tilde{\rho}}(\lambda_{s,T}(\tilde{\omega})\,|\,\tilde{\mathcal{F}}^{y}_{s,t})$, $t\in[s,T]$ (a right continuous modification of conditional expectations).

Then there exists a process $G^{k}_{s,t}(\tilde{\omega}) \sim \tilde{\mathcal{F}}^{y}_{s,t}$, $t\in[s,T]$, such that

$$\tilde{\rho}\{\int_s^T |G^{k}_{r,t}(\tilde{\omega})|^2 dr < \infty\} = 1, \text{ a.s.},$$

for all $1 \leq k \leq d$, $t\in[s,T]$, and

$$\tilde{E}_{\tilde{\rho}}(\lambda_{s,T}(\tilde{\omega})\,|\,\tilde{\mathcal{F}}^{y}_{s,t}) = \tilde{E}_{\tilde{\rho}}\lambda_{s,T}(\tilde{\omega}) + \sum_{k=1}^{d}\int_s^t G^{k}_{s,r}(\tilde{\omega})dy^{k}_{r} \quad \text{, a.s..} \tag{4.4.11}$$

Moreover,

$$\lambda_{s,T}(\tilde{\omega}) = \tilde{E}_{\tilde{\rho}}[\lambda_{s,T}(\tilde{\omega})] + \sum_{k=1}^{d}\int_s^T G^{k}_{s,r}(\tilde{\omega})dy^{k}_{r}. \tag{4.4.12}$$

Proof: See in Liptser and Shiryayev [103, Chp. 5, Thm. 5.8, pp. 171]. QED

By an application Theorem 4.4.1, we deduce that

$$\begin{aligned} d\chi_{s,t}(\tilde{\omega}) &= -\nu_{s,t}\mu_{s,t}\,\mathbb{I}^{u}(t)dt + \sum_{k=1}^{d} G^{k}_{s,t}(\tilde{\omega})dy^{k}_{t} \\ \lim_{t\uparrow T}\chi_{s,t}(\tilde{\omega}) &= \rho_{s,T}(\kappa(x)). \end{aligned} \tag{4.4.13}$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

MINIMUM PRINCIPLE

Next, we shall find the differential equation satisfied by the composition $\psi_{s,t}\chi_{s,t}(\tilde{\omega})$ so that we can express the terminal variational cost of (4.4.4) in terms of an integral process as in the previous section. Since $\psi_{s,t}$ is given by (4.4.3) and $\chi_{s,t}$ satisfies (4.4.13), by the application of Ito differential rule we have

$$\psi_{s,T}\chi_{s,T}(\tilde{\omega}) = \psi_{s,T}\, v_{s,T}\mu_{s,T}(\kappa(x)) = \int_s^T d\psi_{s,r}(\chi_{s,r}(\tilde{\omega})) + \int_s^T \psi_{s,r}(d\chi_{s,r}(\tilde{\omega}))$$

$$= \int_s^T v_{s,r}(\mu_{s,r}\frac{\partial L^{u(r)}}{\partial u}\tilde{u}(r)\mu_{s,r}^{-1}v_{s,r}^{-1}\chi_{s,r}(\tilde{\omega}))dr - \int_s^T \psi_{s,r}(v_{s,r}\mu_{s,r}\Pi^{u(r)})dr$$

$$+ \sum_{k=1}^{d}\int_s^T \psi_{s,r}(G_{s,r}^{k}(\tilde{\omega})dy_r^k. \tag{4.4.14}$$

Substituting (4.4.14) into (4.4.4) we deduce

$$\frac{dJ}{d\varepsilon}(u(\cdot) + \varepsilon\tilde{u}(\cdot))\,|_{\varepsilon=0} = E_{\hat{P}}\{\int_s^T [\psi_{s,r}\rho_{s,r}(\Pi^{u(r)}) + \rho_{s,r}(\frac{\partial \Pi^{u(r)}}{\partial u}\tilde{u}(r))]dr\}$$

$$+ \int_s^T [v_{s,r}(\mu_{s,r}\frac{\partial L^{u(r)}}{\partial u}\tilde{u}(r)\mu_{s,r}^{-1}v_{s,r}^{-1}\chi_{s,r}(\tilde{\omega})) - \psi_{s,r}(v_{s,r}\mu_{s,r}\Pi^{u(r)})]dr$$

$$\sum_{k=1}^{d}\int_s^T \psi_{s,r}(G_{s,r}^{k}(\tilde{\omega}))dy_r^k\}. \tag{4.4.15}$$

Thus, we have the following lemma.

Lemma 4.4.1 The cost functional (4.3.6) or, equivalently, (4.4.4) can be expressed as

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$\frac{dJ}{d\varepsilon}(u(\cdot) + \varepsilon\tilde{u}(\cdot))\,|_{\varepsilon=0} = \tilde{E}_{\tilde{\wp}}\{\int_s^T [v_{s,r}\mu_{s,r}(\frac{\partial \pi^{u(r)}}{\partial u}\tilde{u}(r))dr$$

$$+ v_{s,r}(\mu_{s,r}\frac{\partial L^{u(r)}}{\partial u}\tilde{u}(r)\mu_{s,r}^{-1}v_{s,r}^{-1}\chi_{s,r}(\tilde{\omega}))\}$$

$$= \tilde{E}_{\tilde{\wp}}\{\int_s^T \langle\frac{\partial \pi^{u(t)}}{\partial u}, \rho_t\rangle\tilde{u}(t)dt + \int_s^T \langle\frac{\partial L^{u(t)}}{\partial u}\bar{P}_t(\tilde{\omega}), \rho_t\rangle\tilde{u}(t)dt\} \tag{4.4.16}$$

where

$$\bar{P}(x,\tilde{\omega}) \triangleq \mu_t^{-1}v_t^{-1}\chi_t(x,\tilde{\omega}). \tag{4.4.17}$$

Proof: It follows from (4.4.15) and the fact that processes $\{y_t^k;\ s \le t \le T\}$, $1 \le k \le d$ are $\tilde{\mathcal{F}}_{s,T}^y$-adapted, $\tilde{\wp}$-measure Brownian motions. QED

Theorem 4.4.2

Let U_{ad} be a convex set. Suppose $u(\bullet)$ is an optimal control for the problem

$$J(\tilde{u}(\cdot) = E_{\tilde{\wp}}\{\rho_{s,T}(s(x)) + \int_s^T \rho_{s,r}(\pi^{u(r)})dr\}$$

$$d\rho_{s,t} = L^{u(t)*}\rho_{s,t}dt + \sum_{k=1}^{d} M_k(t)^*(t)\rho_{s,t}\bullet dy_t^k \quad , \rho_{s,s} = \rho_0 \tag{4.4.18}$$

Then there exists a measure-valued process $\bar{P}_t(x,\tilde{\omega}) \in (\tilde{\Omega}, \mathcal{F}_t^y, \tilde{\wp})$ such that the minimum principle is given by

$$\sum_{j=1}^{k}(\tilde{u}_j(t) - u_j(t)[\int_{R^n}\{\frac{\partial \pi}{\partial \tilde{u}_j}(t,x,u(t)) + \frac{\partial}{\partial \tilde{u}_j}L^u\,\bar{P}_t(x,\tilde{\omega})\}\ \rho_t(x,\tilde{\omega})dx] \ge 0 \tag{4.4.19}$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

for all $\tilde{u} \in U_{ad}$ a.e t, a.s.. Suppressing the time variable s the Hamiltonian function $H_t(\rho_t, \bar{P}_t, u_t)$ is expressed as

$$\frac{\partial}{\partial \tilde{u}} H_t(\rho_t, \bar{P}_t, \tilde{u}_t) = \frac{\partial}{\partial \tilde{u}} \int_{R^n} [\pi^{\tilde{u}^v}(t) + L^u(t)\bar{P}_t]\rho_t dx = \frac{\partial}{\partial \tilde{u}} \{ \langle \pi^{\tilde{u}}(t), \rho_t \rangle + \langle \bar{P}_t, L^u(t)^* \rho_t \rangle \}. \tag{4.4.20}$$

Furthermore, the measure-valued process defined by $\bar{P}_{s,t} \triangleq \mu_{s,t}^{-1} \, v_{s,t}^{-1} \, \chi_{s,t}(x,\bar{\omega})$ satisfies the backward stochastic PDE

$$d\bar{P}_t(x,\bar{\omega}) + A^u(t)\, \bar{P}_t(x,\bar{\omega})dt + \pi^u(t)dt + \sum_{k=1}^{d} M_k(t) r_t^k(\bar{\omega})dt + \sum_{k=1}^{d} (M_k(t)\, \bar{P}_t(x,\bar{\omega}) - r_t^k(\bar{\omega}))dy_t^k = 0$$

$$\lim_{t \uparrow T} \bar{P}_t(x,\bar{\omega}) = \kappa(x). \tag{4.4.21}$$

Proof: The minimum principle (4.4.19) can be shown as in Theorem 4.3.2. The definition of the Hamiltonian process follows from (4.4.19).

Next, we shall prove (4.4.21). Consider the composition $\mu_{s,t}^{-1} \, v_{s,t}^{-1}(f)$. It follows from Theorem 4.2.3 that

$$\mu_{s,r}^{-1} \, v_{s,t}^{-1}(f) = f + \int_s^t d\mu_{s,t}^{-1}(v_{s,t}^{-1}(f)) + \int_s^t \mu_{s,t}^{-1}(dv_{s,t}^{-1}(f))$$

$$= f - \sum_{k=1}^{d} \int_s^t M_k(r)\, \mu_{s,r}^{-1}(v_{s,r}^{-1}(f)) \cdot dy_r^k - \int_s^t \mu_{s,r}^{-1}((\mu_{s,r}\, L^u(r)\, \mu_{s,r}^{-1})\, v_{s,r}^{-1}(f))dr$$

$$= f - \sum_{k=1}^{d} \int_s^t M_k(r)\, \mu_{s,r}^{-1} \, v_{s,r}^{-1}(f) \cdot dy_r^k - \int_s^t L^u(r)\, \mu_{s,r}^{-1} v_{s,r}^{-1}(f)dr.$$

Differentiating with respect to t and defining $P_t(x,\bar{\omega}) \triangleq \mu_{s,t}^{-1} \, v_{s,t}^{-1}(f)$ we obtain

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$\begin{aligned} &dP_t(x,\hat{\omega}) + L^u(t)P_t(x,\hat{\omega}) + \sum_{k=1}^{d} M_k(t)P_t(x,\hat{\omega})\cdot dy_t^k = 0 \\ &\lim_{t\downarrow s} P_t(x,\hat{\omega}) = f(x) \end{aligned} \tag{4.4.22}$$

which has a unique solution in the L^2 sense (see, Kunita [96]) and is adjoint to the forward equation satisfied by the unnormalized conditional density. Results of this nature for state-valued processes are found in Kunita [95]. We now apply the Ito differential rule to $P_t\ \chi_t(x,\hat{\omega})$ by first writing the last term of (4.4.22) in terms of the Ito integral.

Thus,

$$\begin{aligned} d(P_t\ (\chi_t(x,\hat{\omega}))) &= dP_t(\chi_t(x,\hat{\omega})) + P_t(d\chi_t(x,\hat{\omega})) \\ &= -A^u(t)\ P_t(\chi_t(x,\hat{\omega}))dt - \sum_{k=1}^{d} M_k(t)\ P_t(\chi_t(x,\hat{\omega}))dy_t^k \\ &- P_t\ (v_t\mu_t(\pi^u(t))) + \sum_{k=1}^{d} P_t(G^k(t)(\hat{\omega}))dy_t^k - \sum_{k=1}^{d} M_k(t)P_t(G^k(t)(\hat{\omega}))dt \end{aligned} \tag{4.4.23}$$

where the last term is due to the quadratic variation of the composition $P_t\ \chi_t(x,\hat{\omega})$ and the operator $A^u(t)$ is given by

$$A^u(t) = L^u(t) + \frac{1}{2}\sum_{k=1}^{d} M_k^2(t)$$

which is also defined in Section 4.1.

Since, by the definition of P_t,

$$P_t\ (v_t\mu_t(\pi^u(t))) = \mu_t^{-1}v_t^{-1}(v_t\mu_t(\pi^u(t)) = \pi^u(t) \tag{4.4.23}$$

and, at the final time

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$P_T\chi_T(\tilde{\omega}) = \mu_T^{-1} v_T^{-1} v_T \mu_T \kappa(x) = \kappa(x),$$

if we define $\bar{P}_t(x,\tilde{\omega}) \triangleq P_t(\chi_t(x,\tilde{\omega}))$ then

$$d\bar{P}_t(x,\tilde{\omega}) = -A^u(t)\bar{P}_t(x,\tilde{\omega})dt - \pi^u(t)dt - \sum_{k=1}^{d} M_k(t)\bar{P}_t(x,\tilde{\omega})dy_t^k - \sum_{k=1}^{d} M_k(t)P_t(G_k(\tilde{\omega}))dt$$
$$+ \sum_{k=1}^{d} P_t(G_k(\tilde{\omega}))dy_t^k.$$

Define $r_t^k(\tilde{\omega}) \triangleq P_t(G_t^k(\tilde{\omega})) = \mu_t^{-1} v_t^{-1}(G_t^k(\tilde{\omega}))$; then $r_t^k(\tilde{\omega})$ is an $\tilde{\mathcal{F}}_t^y$-adapted process which can be verified since P_t satisfies (4.4.22), thus (4.4.21) follows. The existence and uniqueness of (4.4.21) can be shown as in Kunita [96]. QED

<u>Remark 4.4.2</u> The principle of optimality of Theorem 4.4.2 given by (4.4.10) is of the same exact form as the one given in Section 4.3 under Theorem 4.3.2. The adjoint processes $\hat{P}_t$ of Theorem 4.3.2 and $\bar{P}_t$ of Theorem 4.4.2 have different representations. This is based on the different approach taken to arrive at the processes $\hat{P}_t$, $\bar{P}_t$. There is, however a great deal of similarity between the adjoint process given by Bensoussan [9, Sect. 4, p. 31, equation 4.17] for the case of completely observed state-valued control problem and the measure-valued process (4.4.1).

<u>Remark 4.4.3</u> We believe that the method used to derive the result of this section can be extended to the case when the assumption (A′1) through (A′10) are weakened to cover the case of stochastic differential equations with linear growth at least for the case when $h(t, x_t)$ is a bounded function.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

CASE II

The result established in this section can be modified to cover the case introduced in Section 1.2.2. The justification is already given in Section 4.3 under Case II.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

CHAPTER 5

NONLINEAR DECISION PROBLEM

5.1 PRECISE PROBLEM STATEMENT

The objective of this chapter is to study the detection problem described in Section 1.3 of the introduction from the point of view of stochastic PDE's rather than optimal filtering estimates. We believe that the theory presented in Chapter 2 along with the developments of this chapter allow for evaluating the likelihood function as well as the performance bounds of certain nonlinear systems that cannot be evaluated using traditional methods.

We shall study the following decesion problem: given a measurable space $(\Omega, \mathscr{F}_t)$, suppose two probability measures $\wp_0$, $\wp_1$ are defined on it such that they describe the decesion problem

$$(\Omega, \mathscr{F}_t, \wp)\ H_1: \begin{cases} dx_t^1 = f(t, x_t^1)dt + \sigma^1(t, x_t^1)dw_t^1 & , x_{t_0}^1 = x_0^1 \\ dy_t = h^1(t, x_t^1)dt + db_t^1 & , y_{t_0} = 0 \end{cases} \tag{5.1.1}$$

$$(\Omega, \mathscr{F}_t, \wp_0)\ H_0: \begin{cases} dx_t^0 = f^0(t, x_t^0)dt + \sigma^0(t, x_t^0)dw_t^0 & , x_{t_0}^0 = x_0^0 \\ dy_t = h^0(t, x_t^0)dt + db_t^0 & , y_{t_0} = 0. \end{cases} \tag{5.1.2}$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

We shall make the following assumptions:

(A1) $\{w_u^i;\ 0 \le u \le t\}$, $\{b_u^i;\ 0 \le u \le t\}$ are independent Brownian motion processes in R^m, R^d, respectively, under measure $\mathcal{P}_i$ which are also independent of x_0^i for $i = 0, 1$;

(A2) $\mathcal{F}_t \triangleq \sigma\{x_u^1, x_u^0, y_u, 0 \le u \le t\}$ is an increasing family of sub-σ-fields on the measurable space $(\Omega, \mathcal{F}_t)$ satisfying the usual conditions for all $t \in [0,T]$;

(A3) $E_i(x_0)^{2m} < \infty$, $m \ge 1$ for $i = 0, 1$, where E_i denotes expectation with respect to measure $\mathcal{P}_i$;

(A4) f^i: $[0,T] \times R^n \to R^n$ is Borel measurable; continuous, continuously differentiable in x^i satisfying the Lipschitz and linear growth conditions of Theorem 7.C.1 for $i=0,1$;

(A5) σ^i: $[0,T]\times R^n \to R^n \otimes R^m$ is Borel measurable, matrix-valued function, continuous, continuously differentiable in x^i and K^i is a constant such that

$$||\sigma^i(t, x^i)|| + ||\sigma_x^i(t, x^i)|| \le K_1^i \quad \text{for} \quad i = 0, 1;$$

(A6) h^i: $[0,T]\times R^n \to R^d$ is Borel measurable, continuous, continuously differentiable in x^i and there exist a constant K_2^i such that

$$|h(t, x^i)|^2 \le K_2^i(1 + |x^i|^2) \quad \text{for} \quad i = 0, 1.$$

Remark 5.1.1 The conditions stated would be sufficient to establish existence and uniqueness of a weak solution $\{x_t^i;\ 0 \le t \le T\}$ for $i=0, 1$ as defined in Appendix 7.C. Moreover, the process x_t^i, $i = 0, 1$ as well as the process y_t are semi-martingales for all $t \in [0,T]$.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Remark 5.1.2 By Remark 5.1.1 it follows that $w_t^i,\ b_t^i \in M(\mathcal{F}_t,\ P_i)$, i = 0, 1.

The likelihood-ratio LR for the preceding problem has been shown by Duncan [45] and Kailath [80,81,82] to be

$$\Lambda_T = e^{\int_0^T [\hat{h}^1(t, x_t^1) - \hat{h}^0(t, x_t^1)]^T dy_t - \frac{1}{2}\int_0^T [\,|\hat{h}^1(t, x_t^1)|^2 - |\hat{h}^0(t, x_t^0)|^2]dt} \tag{5.1.3}$$

where the estimate $\hat{h}^i(t,\ x_t^i)$ is the conditional mean defined by

$$\hat{h}^i(t,\ x_t^i) \triangleq E_i[h^i(t,\ x_t^i)\,|\,y(u),\ 0 \le u \le t]. \tag{5.1.4}$$

The noisy observations of (5.1.4) generate the filtration $\mathcal{F}_t^y$. Having formulated the LR above, the decesion problem is then the following:

1) Determine a system that provides the least-squares optimal estimates of the signal $h^i(t,\ x_t^i)$ so that a decision is made on which hypothesis to accept;

2) once the decision strategy of 1) is established, obtain the error performance of the LR test.

For both decision and error performance one chooses a threshold γ and performs the following test:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$\Lambda_T \underset{H_0}{\overset{H_1}{\gtrless}} \gamma. \tag{5.1.5}$$

We wish to emphasize at this point that, in general, if the filtering equations of (5.1.1), (5.1.2) are used to obtain the estimates $\hat{h}^i(t, x_t^i)$, the filtering equations represent infinite-dimensional filters. Therefore one is usually led to a consideration of an approximation to $\hat{h}^i(t, x_t^i)$, called a sub-optimal estimate. This kind of approach was one of the problems treated by Evans [ ] and Hibey [ ].

Before we proceed in solving the detection problem as stated above, using the modern approach to nonlinear filtering we shall prove that the LR (5.1.3) can be expressed as a conditional expectation with respect to the filtration $\mathcal{F}_t^y$ of some likelihood-ratio Ψ_T which is restricted to the σ-algebra $\mathcal{F}_t$. This is the key element in our development since the LR Ψ_T provides the connection between the estimation problem of Chapter 2 and the detection problem studied by Evans [ ] and Hibey [ ]. Furthermore, we shall see that our proposed method of solution will not require the knowledge of the filtering equations for $\hat{h}^i(t, x_t^i)$ but rather the solution of a certain stochastic PDE. We now present the theorem that provides the connection between the LR (5.1.3) to the LR Ψ_T defined on the σ-algebra $\mathcal{F}_t$.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Theorem 5.1.1

The likelihood-ratio Λ_T of (5.1.3) can be expressed as

$$\Lambda_T = E_0[\Psi_T \,|\, \mathcal{F}_T^y]$$

where Ψ_T is given by

$$\Psi_T = e^{\int_0^T [h^1(t, x_t^1) - h^0(t, x_t^0)]^T dy_t - \frac{1}{2}\int_0^T [\,|h^1(t, x_t^1)|^2 - |h^0(t, x_t^0)|^2]dt} \tag{5.1.6}$$

Proof: We start by establishing (5.1.6). First notice that

$$y_t - \int_0^t h^0(s,\, x_s^0)ds \in M_{loc}(\mathcal{F}_t,\, P_0).$$

Since $h^i(t,\, x_t^i)$, $i = 0,\, 1$, are predictable, we can define

$$m_t \triangleq \int_0^t [h^1(s,\, x_s^1) - h^0(s,\, x_s^0)]^T [dy_s - h^0(s,\, x_s^0)ds] \in M_{loc}(\mathcal{F}_t,\, P_0). \tag{5.1.7}$$

Therefore, defining Ψ_T as a likelihood-ratio and using the exponential formula (Theorem 7.E.1), we obtain

$$\Psi_T \triangleq E_0[\frac{dP_1}{dP_0} \,|\, \mathcal{F}_T] = e^{m_T - \frac{1}{2}\langle m.,\, m.\rangle_T} \tag{5.1.8}$$

and, by (A4) - (A6), $E_0[\Psi_T] = 1$ so that $P_1 << P_0$. The absolute continuity of P_1 with respect to P_0 also follows from Remark 7.E.2 since, for some $\delta > 0$,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$\sup_{t\in[0,T]} E_0[e^{\delta\int_0^T \langle m, m\rangle_t dt}] < \infty. \tag{5.1.9}$$

Next substituting (5.1.7) into (5.1.8), we get

$$\Psi_T = e^{\int_0^T [h^1(s,x_s^1) - h^0(s,x_s^0)]^T [dy_s - h^0(s,x_s^0)ds] - \frac{1}{2}\int_0^T |h^1(s,x_s^1) - h^0(s,x_s^0)|^2 ds}$$

which can be rewritten as

$$\Psi_T = e^{\int_0^T [h^1(s,x_s) - h^0(s,x_s^0)]^T dy_s - \frac{1}{2}\int_0^T [|h^1(s,x_s^1)|^2 - |h^0(s,x_s^0)|^2] ds} \tag{5.1.10}$$

where $\Psi_T \in M_{loc}(\mathcal{F}_T, P_0)$.

Next, using the fact that $y_t - \int_0^t h^0(s, x_s^0)ds$ and m_t of (5.1.7) are in the class $M_{loc}(\mathcal{F}_t, P_0)$, the martingale translation theorem gives

$$\tilde{m}_t = \int_0^{\cdot} (dy_s - h^0(s,x_s^0)ds) - d\langle (y_\cdot - \int_0^{\cdot} h^0(s,x_s^0)ds), \int_0^{\cdot} (h^1(s,x_s^1) - h^0(s,x_s^0))^T (dy_s - h^0(s,x_s^0)ds)\rangle_t. \tag{5.1.11}$$

Therefore

$$\tilde{m}_t = \int_0^t (dy_s - h^0(s, x_s^0)ds) - \int_0^t (h^1(s, x_s^1) - h^0(s, x_s^0))ds.$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

But $\langle \tilde{m}, \tilde{m} \rangle_t = t$ and $\tilde{m}_t \in M_{loc}(\mathscr{F}_t, P_1)$ are sufficient conditions for $\tilde{m}_t$ to be a Brownian motion as stated in Remark 7.B.4. Thus, (5.1.11) can be written as

$$dy_t = h^1(t, x_t^1)dt + db_t^1$$

with respect to $(\mathscr{F}_t, P_1)$, where the Brownian motion $b_t^1 = \tilde{m}_t$ a.s..

Thus, under $(\mathscr{F}_t, P_0)$, the processes x_t^1, x_t^2, y_t satisfy

$$(\Omega, \mathscr{F}_t, P_0)\ H_0: \begin{cases} dx_t^1 = f^1(t, x_t^1)dt + \sigma^1(t, x_t^1)dw_t^1 \\ dx_t^0 = f^0(t, x_t^0)dt + \sigma^0(t, x_t^0)dw_t^0 \\ dy_t = h^0(t, x_t^0)dt + db_t^0. \end{cases} \tag{5.1.12}$$

Now, to establish the relationship $\Lambda_T = E_0(\Psi_T | \mathscr{F}_T^y)$, we apply the Ito differential rule (Thm.7.B.3) to Ψ_T in (5.1.10) and get

$$d\Psi_t = \Psi_t(h^1(t, x_t^1) - h^0(t, x_t^0))^T (dy_t - h^0(t, x_t^0)\ dt)$$

which is a local martingale under $(\mathscr{F}_t, P_0)$. Therefore, by a general martingale representation theorem as given by Lipster and Shiryayev [103], we deduce that

$$\hat{\Psi}_t \triangleq E_0[\Psi_t | \mathscr{F}_t^y] = E_0[\Psi_0] + \int_0^t \Sigma_s dv_s \tag{5.1.13}$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

where v_t is the innovations process under $(\mathcal{F}_t^y, P_0)$. Similar results in a filtering context are found in Lipster and Shiryayev [103, Chp. 8, Lemma 8.2, p. 300] and van Schuppen [127, Thm. 5.2.1, pp. 87]. The problem now remains to identify the adapted process Σ_t. Using $dv_t = (h^0(t, x_t^0) - \hat{h}^0(t, x_t^0))dt + db_t^0$, we find

$$\begin{aligned} d(\Psi_t - \hat{\Psi}_t) &= \Psi_t(h^1(t, x_t^1) - h^0(t, x_t^0))^T db_t^0 - \Sigma_t dv_t \\ &= \Psi_t(h^1(t, x_t^1) - h^0(t, x_t^0))^T dv_t - \Psi_t(h^1(t, x_t^1) - h^0(t, x_t^0))^T (h^0(t, x_t^0) - \hat{h}^0(t, x_t^0))dt - \Sigma_t dv_t \end{aligned}$$

and, by the Ito differential rule, obtain

$$\begin{aligned} (\Psi_t - \hat{\Psi}_t)v_t &= (\Psi_s - \hat{\Psi}_s)v_s + \int_s^t (\Psi_\tau - \hat{\Psi}_\tau)dv_\tau + \int_s^t d(\Psi_\tau - \hat{\Psi}_\tau)v_\tau + \int_s^t d\langle \Psi_\cdot - \hat{\Psi}_\cdot, v_\cdot \rangle_\tau \\ &= (\Psi_s - \hat{\Psi}_s)v_s + \int_s^t (\Psi_\tau - \hat{\Psi}_\tau)dv_\tau + \int_s^t d(\Psi_\tau - \hat{\Psi}_\tau)v_\tau \\ &\quad + \int_s^t \Psi_\tau(h^1(\tau, x_\tau^1) - h^0(\tau, x_\tau^0))d\tau - \int_s^t \Sigma_\tau \, d\tau. \end{aligned}$$

Continuing,

$$\begin{aligned} (\Psi_t - \hat{\Psi}_t)v_t &= (\Psi_s - \hat{\Psi}_s)v_s + \int_s^t (\Psi_\tau - \hat{\Psi}_\tau)(h^0(\tau, x_\tau^0) - \hat{h}^0(\tau, x_\tau^0))d\tau + \int_s^t (\Psi_\tau - \hat{\Psi}_\tau)db_\tau^0 \\ &\quad + \int_s^t v_\tau \Psi_\tau(h^1(\tau, x_\tau^1) - h^0(\tau, x_\tau^0))db_\tau^0 - \int_s^t v_\tau \Sigma_\tau dv_\tau \\ &\quad + \int_s^t \Psi_\tau(h^1(\tau, x_\tau^1) - h^0(\tau, x_\tau^0)) \, d\tau - \int_s^t \Sigma_\tau d\tau. \end{aligned} \tag{5.1.14}$$

Now, $\hat{\Psi}_t$ is locally integrable which can be shown by Jensen's inequality as follows.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$E_0(|\hat{\Psi}_t|) = E_0|E_0(\Psi_t|\mathcal{F}_t^y)| \le E_0(E_0[|\Psi_t||\mathcal{F}_t^y]) = E_0|\Psi_t|.$$

Thus, since Ψ_t is locally integrable $\hat{\Psi}_t$ is also locally integrable. By the local integrability of v_t,

$$E_0[(\Psi_t - \hat{\Psi}_t)\, v_t|\mathcal{F}_s^y] = 0 \quad \text{for all } s \le t. \tag{5.1.15}$$

Substituting (5.1.14) into (5.1.15), the first term is zero. The third term is also zero since

$$E_0[\int_s^t (\Psi_\tau - \hat{\Psi}_\tau)db_\tau^0|\mathcal{F}_s^y] = E_0[E_0(\int_s^t (\Psi_\tau - \hat{\Psi}_\tau)db_\tau^0|\mathcal{F}_s)|\mathcal{F}_s^y] = 0.$$

The fourth and fifth terms cancel each other due to

$$\begin{aligned}
&E_0[\int_s^t v_\tau\, \Psi_\tau(h^1(\tau, x_\tau) - h^0(\tau, x_\tau^0))db_\tau^0|\mathcal{F}_s^y] \\
&= E_0[E_0(\int_s^t v_\tau\, \Psi_\tau(h(\tau, x_\tau^1) - h^0(\tau, x_\tau^0))db_\tau^0|\mathcal{F}_\tau^y)|\mathcal{F}_s^y] \\
&= E_0[\int_s^t v_\tau\, E_0[\Psi_\tau(h^1(\tau, x_\tau^1) - h^0(\tau, x_\tau^0))db_\tau^0|\mathcal{F}_\tau^y]|\mathcal{F}_s^y] \\
&= E_0[\int_s^t v_\tau\, \Sigma_\tau dv_\tau|\mathcal{F}_s^y].
\end{aligned}$$

Therefore we are left with the terms

$$\begin{aligned}
0 = E_0[&\int_s^t (\Psi_\tau - \hat{\Psi}_\tau)(h^0(\tau, x_\tau^0) - \hat{h}^0(\tau, x_\tau^0))]d\tau \\
&+ \int_s^t \Psi_\tau(h^1(\tau, x_\tau^1) - h^0(\tau, x_\tau^0))d\tau - \int_s^t \Sigma_\tau d\tau|\mathcal{F}_s^y].
\end{aligned}$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Due to the integrability of the first two terms and the fact that $\hat{h}^0(t, x_t^0)$ is $\mathcal{F}_t^y$-adapted we obtain

$$\begin{aligned}\Sigma_t &= E_0[(\Psi_t - \hat{\Psi}_t)(h^0(t, x_t^0) - \hat{h}^0(t, x_t^0)) | \mathcal{F}_\tau^y + E_0[\Psi_t(h^1(t, x_t^1) - h^0(t, x_t^0)) | \mathcal{F}_\tau^y] \\ &= E_0[\Psi_t h^0(t, x_t^0) | \mathcal{F}_t^y] - \hat{\Psi}_t \hat{h}^0(t, x_t^0) + \hat{\Psi}_t \hat{h}^0(t, x_t^0) - \hat{\Psi}_t \hat{h}^0(t, x_t^0) \\ &+ E_0[\Psi_t h^1(t, x_t^1) | \mathcal{F}_t^y] - E_0[\Psi_t h^0(t, x_t^0) | \mathcal{F}_t^y].\end{aligned}$$

Therefore the representation of Σ_t is

$$\Sigma_t = E_0[(\Psi_t h^1(t, x_t^1) | \mathcal{F}_t^y] - E_0[\Psi_t | \mathcal{F}_t^y] E_0[h^0(t, x_t^0) | \mathcal{F}_t^y]. \tag{5.1.16}$$

Since $h^1(t, x_t^1)$ is an integrable function, using (7.E.6),

$$E_1[h^1(t, x_t^1) | \mathcal{F}_t^y] = \frac{E_0[h^1(t, x_t^1) \Psi_t | \mathcal{F}_t^y]}{E_0[\Psi_t | \mathcal{F}_t^y]}. \tag{5.1.17}$$

Thus we can write (5.1.13) using (5.1.16), (5.1.17) as

$$\begin{aligned}E_0[\Psi_t | \mathcal{F}_t^y] &= 1 + \int_0^t [E_0(\Psi_\tau | \mathcal{F}_\tau^y) E_1(h^1(\tau, x_\tau^1) | \mathcal{F}_\tau^y) \\ &\quad - E_0(\Psi_\tau | \mathcal{F}_\tau^y) E_0(h^0(\tau, x_\tau^0) | \mathcal{F}_\tau^y)] dv_\tau \\ &= 1 + \int_0^t E_0(\Psi_\tau | \mathcal{F}_\tau^y)[E_1(h^1(\tau, x_\tau^1) | \mathcal{F}_\tau^y) - E_0(h_0(\tau, x_\tau^0) | \mathcal{F}_\tau^y)] dv_\tau.\end{aligned} \tag{5.1.18}$$

Applying the Ito differential rule to $\ln E_0(\Psi_t | \mathcal{F}_t^y)$,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$d(\ln \hat{\Psi}_t) = \frac{d\hat{\Psi}_t}{\hat{\Psi}_t} - \frac{d\langle\hat{\Psi}, \hat{\Psi}\rangle_t}{2\hat{\Psi}_t^2}$$

where $d\hat{\Psi}_t$ is the differential of (5.1.18). It then follows that

$$d(\ln \hat{\Psi}_t) = [E_1(h^1(t, x_t^1) \mid \mathcal{F}_t^y) - E_0(h^0(t, x_t^0) \mid \mathcal{F}_t^y)]dv_t$$
$$- \frac{1}{2} \mid E_1(h^1(t, x_t^1) \mid \mathcal{F}_t^y) - E_0(h^0(t, x_t^0) \mid \mathcal{F}_t^y) \mid^2 dt$$

so that

$$\hat{\Psi}_t = E_0[\Psi_t \mid \mathcal{F}_t^y] = e^{\int_s^t [\hat{h}^1(\tau, x_\tau^1) - \hat{h}^0(\tau, x_\tau^0)]dv_\tau - \frac{1}{2}\int_0^t |\hat{h}^1(\tau, x_\tau^1) - \hat{h}^0(\tau, x_\tau^0)|^2 d\tau} \tag{5.1.19}$$

Finally, substituting the innovations process $dv_t = dy_t - \hat{h}^0(t, x_t^0)dt$ into (5.1.19), we recover (5.1.6) which proves our theorem since $\hat{\Psi}_t$ is by definition equal to Λ_t as given by (5.1.3). QED

Remark 5.1.3 It follows that if the detection problem (5.1.1), (5.1.2) is modified so that under hypothesis H_0 we only observe Brownian motion, (no signal present), then (5.1.1) remains the same but (5.1.2) is replaced by

$$(\Omega, \mathcal{F}_t, \mathcal{P}_0)\ H_0: \{dy_t = db_t.$$

The LR Λ_T of Theorem 5.1.1 then becomes

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$\Lambda_T = E_0\left[e^{\int_0^T h(t, x_t)dy_t - \frac{1}{2}\int_0^T |h(t, x_t)|^2 dt} \,\middle|\, \mathcal{F}_t^y\right]$$

where in this case $y_t \in M(\mathcal{F}_t, P_0)$ is a standard Brownian motion and Ψ_T is the likelihood-ratio used in nonlinear filtering.

5.2 EVALUATION OF LIKELIHOOD-RATIO

In this section we shall present the theory that allow us to evaluate the LR for the decision problem (5.1.1), (5.1.2). We shall show that the likelihood-ratio, as defined by (5.1.3) and related to Ψ_T through Theorem 5.1.1 can be written as a ratio of two densities integrated with respect to a new augmented state process $\tilde{x}_t \triangleq [x_t^1 \; x_t^0]^T$ over the space $R^n \otimes R^n$. These two densities are related to the unnormalized conditional density considered in the development of Chapter 2. Furthermore, we shall show that our method of solution will require no knowledge of the optimal filtering etimates for $h^i(t, x_t^i)$. Thus, in certain applications which result in infinite coupled filtering equations, computing Λ_T by (5.1.3) would not be possible but, as we shall now show, this approach will allow us to evaluate the two densities exactly so that Λ_T can be computed.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Theorem 5.2.1

In the likelihood-ratio test

$$\Lambda(y_t) = E_0(\Psi_t|\mathcal{F}_t^y) \underset{H_0}{\overset{H_1}{\gtrless}} \gamma, \tag{5.2.1}$$

the LR can be expressed as

$$\Lambda(y_t) = \frac{E_2[\exp\{\int_0^t h^1(s, x_s^1)dy_s - \frac{1}{2}\int_0^t |h^1(s, x_s^1)|^2 ds\} | \mathcal{F}_t^y]}{E_2[\exp\{\int_0^t h^0(s, x_s^0)\, dy_s - \frac{1}{2}\int_0^t |h^0(s, x_s^0)|^2 ds\} | \mathcal{F}_t^y]} \tag{5.2.2}$$

$$= \frac{\int_{R^n \otimes R^n} \rho^1(\tilde{x}, t)d\tilde{x}}{\int_{R^n \otimes R^n} \rho^0(\tilde{x}, t)d\tilde{x}} \tag{5.2.3}$$

where E_2 denotes expectation with respect to a measure P_2, the Wiener measure $\mu_w^d(dy)$

(see definition in Section 2.2.2) the process $\tilde{x}_t$ satisfies

$$(\Omega, \mathcal{F}_t, P_2): d\tilde{x}_t = \tilde{f}(t, \tilde{x}_t)dt + \tilde{\sigma}(t, \tilde{x}_t)d\tilde{w}_t, \ \ \tilde{x}_0 \tag{5.2.4}$$

with

$$\tilde{x}_t \triangleq \begin{bmatrix} x_t^1 \\ x_t^0 \end{bmatrix}, \quad \tilde{f}_t \triangleq \begin{bmatrix} f_t^1 \\ f_t^0 \end{bmatrix}, \quad \tilde{\sigma}_t \triangleq \begin{bmatrix} \sigma_t^1 & 0 \\ 0 & \sigma_t^0 \end{bmatrix}, \quad \tilde{w}_t \triangleq \begin{bmatrix} w^1 \\ w^0 \end{bmatrix},$$

and the density $\rho^i(\tilde{x}, t)$ satisfies the stochastic PDE

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$
\begin{aligned}
&d\rho^i(\tilde{x}, t) = L_{\tilde{x}}(t)^* \rho^i(\tilde{x}, t)dt + h^i(t, x^i)\, \rho^i(\tilde{x}, t)dy_t \\
&\lim_{t \downarrow t_0} \rho^i(\tilde{x}, t) = p_0(\tilde{x})
\end{aligned}
\tag{5.2.5}
$$

where

$$
L_{\tilde{x}}(t)(\cdot) \triangleq \frac{1}{2}\tilde{\sigma}\,\tilde{\sigma}^T \frac{\partial}{\partial \tilde{x}^2}(\cdot) + \tilde{f}^T \frac{\partial}{\partial \tilde{x}},\ p_0 = \text{the initial density.}
$$

Proof: Starting with the likelihood-ratio of Theorem 5.1.1, by (7.E.6) if $\mathcal{P}_2$ is a new probability measure we must have

$$
E_0[\Psi_t \mid \mathcal{F}_t^y] = \frac{E_2[\Psi_t \frac{d\mathcal{P}_0}{d\mathcal{P}_2} \mid \mathcal{F}_t^y]}{E_2[\frac{d\mathcal{P}_0}{d\mathcal{P}_2} \mid \mathcal{F}_t^y]}.
$$

where the function Ψ_t is given by (5.1.6). Let us now define the processes x_t^0, x_t^1, y_t on the probability space $(\Omega, \mathcal{F}_t, \mathcal{P}_2)$ by

$$
(\Omega, \mathcal{F}_t, \mathcal{P}_2):
\begin{cases}
dx_t^1 = f^1(t, x_t^1)dt + \sigma^1(t, x_t^1)dw_t^1 & , x_0^1 \\
dx_t^0 = f^0(t, x_t^0)dt + \sigma^0(t, x_t^0)dw_t^0 & , x_0^0 \\
dy_t = db_t^0\ . & , y_0 = 0
\end{cases}
\tag{5.2.6}
$$

Next we show that $\mathcal{P}_0 << \mathcal{P}_2$. Using the same reasoning that led to the definition of m_t in Theorem 5.1.1, we now define m'_t as

$$
m'_t \triangleq \int_0^t h^0(s, x_s^0)dy_s \in M_{loc}(\mathcal{F}_t, \mathcal{P}_2)
$$

and the exponential formula Ψ'_T as

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$\Psi'_T \triangleq E_2[\frac{d\mathcal{P}_0}{d\mathcal{P}_2} | \mathcal{F}_T] = e^{m'_T - \frac{1}{2}\langle m', m'\rangle_T} \tag{5.2.7}$$

where $E_2[\Psi'_T] = 1$. Similarly as in Theorem 5.1.1 we have $\mathcal{P}_0 << \mathcal{P}_2$ which follows from the weaker condition

$$\mathcal{P}_2\{\int_0^T \langle m', m'\rangle_t dt < \infty\} = \text{a.s.}.$$

given by Liptser and Shiryayev [103, Chp. 6, Example 4, pp. 221] which is satisfied since $h^0(t,x_t^0)$ satisfies the condition of (A6).

The exponential martingale Ψ'_T can be expressed as

$$\Psi'_T = e^{\int_0^T h^0(s,\, x_s^0) dy_s - \frac{1}{2}\int_0^T |h^0(s,\, x_s^0)|^2 ds} \quad . \tag{5.2.8}$$

Therefore, by (5.1.6) and the fact that y_t is a standard Brownian motion under $(\mathcal{F}_t, \mathcal{P}_2)$, it follows that

$$E_0[\Psi_t | \mathcal{F}_t^y] = \frac{E_2[\exp\{\int_0^t h^1(s,\, x_s^1) dy_s - \frac{1}{2}\int_0^t |h^1(x,\, x_s^1)|^2 ds\} | \mathcal{F}_t^y]}{E_2[\exp\{\int_0^t h^0(s,\, x_s^0) dy_s - \frac{1}{2}\int_0^t |h^0(s,\, x_s^0)|^2 ds\} | \mathcal{F}_t^y]}$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

which proves (5.2.2). Moreover, using the fact that y_t is a Wiener process under $(\mathcal{F}_t, P_2)$ and m'_t above is of class $M_{loc}(\mathcal{F}_t, P_2)$, the martingale translation theorem, (Theorem 7.E.3) shows that

$$\tilde{m}'_t = y_t - d\langle y., \int_0^{\cdot} h^0(s, x_s^0) dy_s \rangle \in M_{loc}(\mathcal{F}_t, P_0). \tag{5.2.9}$$

Again, as in Theorem 5.1.1, $\langle \tilde{m}', \tilde{m}' \rangle_t = It$, $\tilde{m}_t' \in M_{loc}(\mathcal{F}_t, P_0)$; thus $\tilde{m}'_t = b_t^0$ is a standard Brownian motion and (5.2.9) can be written as

$$dy_t = h^0(t, x_t^0)dt + db_t^0$$

with respect to $(\mathcal{F}_t, P_0)$. Therefore, as a result of the measure transformation, under measure P_0 the processes x_t^0, x_t^1, y_t satisfy (5.1.12). The equality between (5.2.2), (5.2.3) follows from the proof of the unnormalized conditional density equation presented in Chapter 2. QED

Remark 5.2.1 The likelihood-ratio presented in Theorem 5.2.1 and expressed as a ratio of two densities is a generalization of the likelihood-ratio found in Van Trees [128, Chp. 2] and examples therein, where it is expressed as a ratio of two Gaussian densities. This follows by defining the likelihood-ratio as $\Lambda_t \triangleq E_0[\frac{dP_1}{dP_0} \mid \mathcal{F}_T]$ and the fact that the P_i are Gaussian measures since the estimates $\hat{h}_t^i$ are treated as if they were nonrandom.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

5.3 PERFORMANCE BOUNDS

In this section we shall answer the second question associated with the detection problem of Section 5.1.1 using stochastic PDE's. As mentioned in Section 1.3 the question of performance bounds was addressed by Evans [52] and Hibey [74,75] by expressing the Chernoff bound in terms of an evolution PDE having as coefficients the state and signal estimates. This PDE can be solved exactly only for the case of linear systems. Thus, for nonlinear systems, they were led to a consideration of sub-optimal estimates. Our approach relates the Chernoff bound to stochastic PDE's similar to the ones given in Theorem 5.2.1 that can be solved exactly not only for linear systems but also for some cases of nonlinear systems. We now present the theorem that provides error bounds for P_F and P_M.

Theorem 5.3.1

Consider the detection problem described by (5.1.1), (5.1.2). The Chernoff bound on P_F given by (1.3.7) can be obtained from

$$P_F \leq e^{-s\ln\gamma} E_2[\sigma^s(h^1(t, x_t^1))\sigma^{1-s}(h^0(t, x_t^0))], \; s > 0 \tag{5.3.1}$$

and the Chernoff bound on P_M given by (1.3.10) can be obtained from

$$P_M \leq e^{-s\ln\gamma} E_2[\sigma^{s+1}(h^1(t, x_t^1))\sigma^{-s}(h^0(t, x_t^0))], \; s < 0 \tag{5.3.2}$$

where

$$\sigma(h^i(t, x_t^i)) \triangleq \int_{R^n \otimes R^n} \rho^i(\tilde{x}_t, t)d\tilde{x} \tag{5.3.3}$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

and $\rho^i(\tilde{x}, t)$ satisfies the stochastic PDE given by (5.2.5). The expectation E_2 is with respect to probability measure $\mathcal{P}_2$ which in this case is the Wiener measure $\mu_w^d(dy)$ on the space $C([0, T]; R^d)$.

Proof: From (1.3.7) and Theorem 5.2.1 we have

$$P_F \le e^{-s\ln\gamma} E_0[\Lambda_t^s] = e^{-s\ln\gamma} E_0\left\{\frac{\int_{R^n \otimes R^n} \rho^1(\tilde{x}, t)d\tilde{x}}{\int_{R^n \otimes R^n} \rho^0(\tilde{x}, t)d\tilde{x}}\right\}^s,$$

that is,

$$P_F \le e^{-s\ln\gamma} E_0\{\frac{\sigma^s(h^1(t,x_t^1))}{\sigma^s(h^0(t,x_t^0))}\}. \tag{5.3.4}$$

By the definition of the likelihood-ratio Ψ'_T given by (5.2.7) the unconditional expectation $E_0(\cdot)$ of (5.3.4) is related to the unconditional expectation $E_2(\cdot)$ through (7.E.5). Thus we can write (5.3.4) as

$$\begin{aligned} P_F &\le e^{-s\ln\gamma} E_2\{\frac{d\mathcal{P}_0}{d\mathcal{P}_2} \frac{\sigma^s(h^1(t, x_t^1))}{\sigma^s(h^0(t, x_t^0))}\} \\ &= e^{-s\ln\gamma} E_2\{E_2[\frac{d\mathcal{P}_0}{d\mathcal{P}_2} \frac{\sigma^s(h^1(t, x_t^1))}{\sigma^s(h^0(t, x_t^0))} | \mathcal{F}_t^y]\}. \end{aligned}$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

But $\sigma(h^i(t, x_t^i))$, i=0,1 is measurable with respect to $\mathscr{F}_t^y$, therefore

$$P_F \leq e^{-s\ln\gamma} E_2\{E_2(\frac{d\rho_0}{d\rho_2} | \mathscr{F}_t^y) \frac{\sigma^s(h^1(t, x_t^1))}{\sigma^s(h^0(t, x_t^0))}\}$$

$$= e^{-s\ln\gamma} E_2\{\sigma(h^0(t, x_t^0)) \frac{\sigma^s(h^1(t, x_t^1))}{\sigma^s(h^0(t, x_t^0))}\}.$$

The last equality follows from the result of Theorem 5.2.1 since $E_2 [\frac{d\rho_0}{d\rho_2} | \mathscr{F}_t^y]$ is the denominator of (5.2.3), which is by definition equal to $\sigma(h^0(t,x_t^0))$, thus (5.3.1) is shown.

The proof of (5.3.2) begins with

$$P_M \leq e^{-s\ln\gamma} E_0[\Lambda_t^{s+1}]$$

and follows the same procedure as above.

Finally, the fact that ρ_2 is a Wiener measure $\mu_w^d(dy)$ is due to $\sigma(h^i(t, x_t^i))$, i=0,1 being $\mathscr{F}_t^y$-adapted processes. QED

Remark 5.3.1 We notice that the performance bounds for P_F, P_M given by (5.3.1), (5.3.2), respectively, are a generalization of the performance bounds given by Van Trees [128, Chp. 2, equations 448, 449]. For cases described by Remark 5.2.1. Thus, if the h^i are known we can recover the performance bounds given by Van Trees by defining

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$p_{y|H_i}(Y|H_i) \triangleq \sigma(h^i(t, x_t^i))$, where $p_{y|H_i}(Y|H_i)$ is the probability density of the observations y under hypothesis H_i.

5.4 EXAMPLES

In this section we shall apply the results of Theorems 5.2.1, 5.3.1 to four detection problems; both linear and nonlinear examples will be presented to demonstrate the approach.

Example 5.4.1 Linear Decision Problem.

We consider the following linear detection problem:

$$(\Omega, \mathcal{F}_t, \mathcal{P}_1)\ H_1: \begin{cases} dx_t^1 = B^1 x_t^1 dt + dw_t^1 &, x_{t_0}^1 = x_0^1 \\ dy_t = C^1 x_t^1 dt + db_t^1 &, y_{t_0} = 0 \end{cases} \tag{5.4.1}$$

$$(\Omega, \mathcal{F}_t, \mathcal{P}_0)\ H_0: \begin{cases} dx_t^0 = B^0 x_t^0 dt + dw_t^0 &, x_{t_0}^0 = x_0^0 \\ dy_t = C^0 x_t^0 dt + db_t^0 &, y_{t_0} = 0 \end{cases} \tag{5.4.2}$$

where x^i, i=0,1 are of dimension one and, B^i, C^i, i=0,1 are constants. Using the notation of (5.1.1), (5.1.2) we define $f^i \triangleq B^i x_t^i$, $h^i \triangleq C^i x_t^i$ and $\sigma^i \triangleq 1$. Therefore, we have n=m=d=1.

From Theorem 5.2.1 under measure $\mathcal{P}_2$ we have the augmented system

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$
(\Omega, \mathscr{F}_t, P_2): \begin{cases} d\tilde{x}_t = A\tilde{x}_t dt + Bd\tilde{w}_t & , x^1_{t_0} = x^1_0,\ x^0_{t_0} = x^0_0 \\ A = \begin{bmatrix} B^1 & 0 \\ 0 & B^0 \end{bmatrix}, B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \tilde{w}_{t_0} = \begin{bmatrix} w^1_t \\ w^0_t \end{bmatrix}. \end{cases} \tag{5.4.3}
$$

The likelihood-ratio of Theorem 5.2.1 requires the solutions of $\rho^i(\tilde{x}, t), i = 0,1$

of the Fisk-Stratonovich stochastic PDE's

$$
\begin{aligned} \frac{d}{dt}\rho(\tilde{x}, t) &= \frac{Tr}{2} \frac{\partial^2}{\partial \tilde{x}^2}(BB^T \rho^i(\tilde{x}, t)) - \frac{\partial}{\partial \tilde{x}}(A\tilde{x}\, \rho^i(\tilde{x}, t)) \\ &- \frac{1}{2}(C^i x^i)^2 \rho^i(\tilde{x}, t) + C^i x^i \rho^i(\tilde{x}, t) \cdot \frac{dy_t}{dt} \quad , \rho(\tilde{x}, t_0) = p_0(\tilde{x}). \end{aligned} \tag{5.4.4}
$$

However, since we are required to solve (5.4.4) for i=0,1, we may avoid having to repeat the solution procedure twice by imbedding (5.4.4) into the more general stochastic PDE

$$
\begin{aligned} \frac{d}{dt}\rho^i(\tilde{x}, t) &= \frac{Tr}{2} \frac{\partial^2}{\partial \tilde{x}^2} (BB^T \rho^i(\tilde{x}, t)) - \frac{\partial}{\partial \tilde{x}} (A\tilde{x}\, \rho^i(\tilde{x}, t)) \\ &- \frac{1}{2}\, \tilde{x}^T \tilde{C}^T \tilde{C}\, \tilde{x} \rho^i(\tilde{x}, t) + \tilde{x}^T \tilde{C}^T \rho^i(\tilde{x}, t) \cdot \frac{dy_t}{dt} \end{aligned} \tag{5.4.5}
$$

where $\tilde{C} = [\alpha^1 C^1\ \alpha^0 C^0]$, $\alpha^0 = 1$, $\alpha^1 = 0$ if i=0, and $\alpha^0 = 0$, $\alpha^1 = 1$ if i=1.

We solve (5.4.5) by choosing a solution of the form

$$
\rho^i(\tilde{x}, t) = k^i_t\, e^{-\frac{1}{2} (\tilde{x} - \mu^i_t)^T \Sigma^{i^{-1}}_t (\tilde{x} - \mu^i_t)} \quad , \quad \rho^i(\tilde{x}, t_0) = \delta(\tilde{x} - \tilde{x}_0) \tag{5.4.6}
$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

where k_t^i is a scalar function, μ_t^i is a 2-dimensional vector, and Σ_t^i a 2x2 symmetric matrix having inverse $\Sigma_t^{i^{-1}}$. Substituting (5.4.6) into (5.4.5) we can identify the equations satisfied by μ_t^i, Σ_t^i, k_t^i. Thus, with dots denoting differentiation with respect to t, we find

$$\frac{d}{dt}\rho^i(\tilde{x},t) = \dot{k}_t^i\rho^i(\tilde{x},t) + [\frac{1}{2}(\tilde{x}-\mu_t^i)^T\Sigma_t^{i^{-1}}\dot{\Sigma}_t^i\Sigma_t^{i^{-1}}(\tilde{x}-\mu_t^i) + (\tilde{x}-\mu_t^i)^T\Sigma_t^{i^{-1}}\dot{\mu}_t^i]\rho^i(\tilde{x},t) \tag{5.4.7}$$

$$(\frac{\partial}{\partial\tilde{x}}\rho(\tilde{x},t))^T = -(\tilde{x}-\mu_t^i)^T\Sigma_t^{i^{-1}}\rho^i(\tilde{x},t) \tag{5.4.8}$$

$$\frac{\partial^2}{\partial\tilde{x}^2}\rho^i(\tilde{x},t) = -\Sigma_t^{i^{-1}}\rho^i(\tilde{x},t) + \Sigma_t^{i^{-1}}(\tilde{x}-\mu_t^i)(\tilde{x}-\mu_t^i)^T\Sigma_t^{i^{-1}}\rho^i(\tilde{x},t). \tag{5.4.9}$$

Substituting (5.4.7) , (5.4.8), (5.4.9) into (5.4.5) and equating coefficients of the state vector $\tilde{x}$, we find expressions for $\dot{\Sigma}_t^i$, $\dot{\mu}_t^i$, $\dot{k}_t^i$ given by

$$\dot{\Sigma}_t^i = A\,\Sigma_t^i + \Sigma_t^i\,A^T - \Sigma_t^i\,\tilde{C}^T\tilde{C}\,\Sigma_t^i + BB^T \quad , \Sigma_{t_0}^i = 0 \tag{5.4.10}$$

$$\dot{\mu}_t^i = A\mu_t^i - \Sigma_t\,\tilde{C}^T\tilde{C}\,\mu_t + \Sigma_t^i\,\tilde{C}^T\cdot\frac{dy_t}{dt} \quad , \mu_{t_0}^i = [x_0^1\ x_0^0]^T \tag{5.4.11}$$

$$\dot{k}_t^i = \mu_t^{i^T}\Sigma_t^{i^{-1}}\dot{\mu}_t^i + \frac{Tr}{2}(B^T\Sigma_t^{i^{-1}}\mu_t^i\mu_t^{i^T}\Sigma_t^{i^{-1}}B) - \frac{1}{2}\mu_t^{i^T}\Sigma_t^{i^{-1}}\dot{\Sigma}_t^i\Sigma_t^{i^{-1}}\mu_t^i$$
$$- \frac{1}{2}Tr(B^T\Sigma_t^{i^{-1}}B) - Tr(A). \qquad , k_{t_0} = k_0 = 1.$$

Writing (5.4.6) as

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$\rho^i(\tilde{x}, t) = (2\pi)\ (\det \Sigma_t^i)^{1/2}\ k_t^i\ \frac{e^{-\frac{1}{2}(\tilde{x}-\mu_t^i)^T \Sigma_t^{i^{-1}} (\tilde{x}-\mu_t^i)}}{(2\pi)\ (\det \Sigma_t^i)^{1/2}} \tag{5.4.12}$$

and using normalization, we deduce

$$\int_{R^n \otimes R} \rho^i(\tilde{x}, t) d\tilde{x} = (2\pi)(\det \Sigma_t^i)^{1/2}\ k_t^i. \tag{5.4.13}$$

Thus, the likelihood-ratio test as given by Theorem 5.2.1 is evaluated by

$$\Lambda(y_t) = \left(\frac{\det \Sigma_t^1}{\det \Sigma_t^0}\right)^{1/2} \left(\frac{k_t^1}{k_t^0}\right) \underset{H_0}{\overset{H_1}{\gtrless}} \gamma$$

where k_t^i are driven by the observation process $\{y_s,\ t_0 \le s \le t\}$ which is a Brownian motion. It is important to note that $\Sigma_t^i, i=0,1$ satisfies a Riccati equation with zero initial condition. On the other hand $k_t^i, i=0,1$ satisfy a stochastic differential equation driven by standard Brownian motion y_t having a known initial condition $k_{t_0}^i = k_0^i = 1, i=0,1$.

By Theorem 5.3.1, the Chernoff bounds for P_F, P_M are given by

$$P_F \le e^{-s \ln \gamma}\ E_2\{[2\pi(\det \Sigma_t^1)^{1/2}\ k_t^1]^s [2\pi(\det \Sigma_t^0)^{1/2}\ k_t^0]^{1-s}\}\ ,\ s > 0$$

$$P_M \le e^{-s \ln \gamma}\ E_2\{[2\pi(\det \Sigma_t^1)^{1/2}\ k_t^1]^{s+1} [2\pi(\det \Sigma_t^0)^{1/2}\ k_t^0]^{-s}\}\ ,\ s < 0.$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Remark 5.4.1 The general case $x_t^i \in R^n, h^i(t, x_t^i) \in R^d$ is handled similarly, however it requires much more notation. For this reason we shall treat the multidimensional case last by considering a nonlinear detection problem. Next we treat the version of the above example when there is no signal present under hypothesis H_0.

Example 5.4.2 Linear Decision Problem.

Consider the detection problem as stated in Remark 5.4.1.

$$(\Omega, \mathscr{F}_t, \rho_1)\ H_1: \begin{cases} dx_t = B\,x_t dt + dw_t & , x_{t_0} = x_0 \\ dy_t = x_t dt + db_t & , y_{t_0} = 0 \end{cases} \tag{5.4.14}$$

$$(\Omega, \mathscr{F}_t, \rho_0)\ H_0: \{dy_t = db_t \qquad , y_{t_0} = 0. \tag{5.4.15}$$

Since there is no signal under hypothesis H_0,using the notation established in Example 5.4.1, $C^0 = 0$. The likelihood-ratio test of Theorem 5.2.1 is determined once the densities $\rho^i(x, t)$, i=0,1 are determined, where

$$\begin{aligned} \frac{d}{dt}\rho^1(x,t) &= \frac{1}{2}\frac{\partial^2}{\partial x^2}\rho^1(x,t) - \frac{\partial}{\partial x}(Bx\,\rho^1(x,t)) \\ &\quad - \frac{1}{2}x^2\rho^1(x,t) + x\,\rho^1(x,t)\cdot\frac{dy_t}{dt} \quad , \rho^1(x,t_0) = p_0(x) \end{aligned} \tag{5.4.16}$$

$$\frac{\partial}{\partial t}\rho^0(x,t) = \frac{1}{2}\frac{\partial^2}{\partial x^2}\rho^0(x,t) - \frac{\partial}{\partial x}(B\,\rho^0(x,t)) \ , \ \rho^1(x,t_0) = p_0(x).$$

The equation satisfied by ρ^0 (x, t) is a Fokker-Planck equation which integrates to one, so, the likelihood-ratio of Theorem 5.2.1 is consistent with the one presented in Remark 5.1.3, as expected. Therefore, needing only to solve for $\rho^1(t, x)$, we assume a solution of the form

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$\rho(x,t) = (2\pi)^{-1/2}\,(\det \Sigma_t)^{-\frac{1}{2}}\,k_t^{-1}\,e^{-\frac{1}{2}(x-\mu_t)^T\Sigma_t^{-1}(x-\mu_t)} \tag{5.4.17}$$

$$\rho(x,t_0) = \delta(x - x_0)$$

where

$$\frac{\partial}{\partial t}\rho(x,t) = -\frac{1}{2}\frac{\dot{\det}\,\Sigma_t}{\det \Sigma_t}\rho(x,t)$$
$$- k_t^{-1}\,\dot{k}_t\,\rho(x,t) + [\frac{1}{2}(x-\mu_t)^T\Sigma_t^{-1}\dot{\Sigma}_t\Sigma_t^{-1}(x-\mu_t) + (x-\mu_t)^T\Sigma_t^{-1}\dot{\mu}_t]\,\rho(x,t)$$

$(\frac{\partial}{\partial x}\rho(x,t))^T$ is given by (5.4.8), and

$\frac{\partial^2}{\partial x^2}\rho(x,t)$ is given by (5.4.9).

Substituting the above derivatives into (5.4.16) we obtain

$$-\frac{1}{2}\frac{\dot{\det}\,\Sigma_t}{\det \Sigma_t} - k_t^{-1}\,\dot{k}_t + \frac{1}{2}\,(x-\mu_t)^T\Sigma_t^{-1}\,\dot{\Sigma}_t\,\Sigma_t^{-1}(x-\mu_t) + (x-\mu_t)\Sigma_t^{-1}\dot{\mu}_t$$
$$= -\frac{1}{2}\Sigma_t^{-1} + \frac{1}{2}\,\Sigma_t^{-1}(x-\mu_t)\,(x-\mu_t)^T\,\Sigma_t^{-1} - B + (x-\mu_t)^T\,\Sigma_t^{-1}\,Bx - \frac{1}{2}\,x^2 + x\cdot\frac{dy}{dt}.$$

$$\dot{\Sigma}_t = 1 + 2\,\Sigma_t B - \Sigma_t^2 \qquad , \quad \Sigma_{t_0} = 0$$

$$\dot{\mu}_t = (B - \Sigma_t)\mu_t + \Sigma_t\cdot\frac{dy_t}{dt}\,, \quad \mu_{t_0} = x_0$$

$$\dot{k}_t = k_t(-\frac{1}{2}\frac{\dot{\det}\,\Sigma_t}{\det \Sigma_t} + \frac{1}{2}\,\mu_t^2 + \frac{1}{2}\,\Sigma_t^{-1} + B - \mu_t\cdot\frac{dy_t}{dt})\,, \quad k_{t_0} = 1\,.$$

The likelihood-ratio test is computed by

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$\Lambda(y_t) = (k_t)^{-1} \underset{H_0}{\overset{H_1}{\gtrless}} \gamma$$

where as before k_t satisfies a stochastic differential equation driven by the observation process y_t.

By Theorem 5.3.1 and (5.4.17) the performance bounds P_F, P_M are then given by

$$P_F \le e^{-s\ln\gamma} E_0[(k_t)^{-s}] , \; s > 0$$

$$P_M \le e^{-s\ln\gamma} E_0[(k_t)^{-s-1}] , \; s < 0$$

where $\mathcal{P}_0$ is a Wiener measure on the space $C([0, T]; R^1)$.

Remark 5.4.2 We note that the likelihood-ratio test of Example 5.4.1 as well as that of Example 5.4.2 depend on the observed process $\{y_s, t_0 \le s \le t\}$, in terms of differntial equations of Riccati type.

Example 5.4.3 Nonlinear Decision Problem.

This example is a special case of the nonlinear filtering problem treated by Beneš [5] . We model the detection problem as follows:

$$(\Omega, \mathcal{F}_t, \mathcal{P}_1)\; H_1 : \begin{cases} dx_t^1 = f^1(x_t^1)dt + dw_t^1 & , x_{t_0}^1 = x_0^1 \\ dy_t = c^1 x_t^1 dt + db_t^1 & , y_{t_0} = 0 \end{cases} \tag{5.4.18}$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$(\Omega, \mathcal{F}_t, P_0)\ H_0: \begin{cases} dx_t^0 = f^0(x_t^0)dt & , x_{t_0}^0 = x_0^0 \\ dy_t = c^0 x_t^0 dt + db_t^0 & , y_{t_0} = 0 \end{cases} \tag{5.4.19}$$

where x^i, $i = 0, 1$ is of dimension one, c^i, $i = 0, 1$ is a constant number. We shall assume that $f^1(x^1)$ is the gradient of a potential function $F(x^1)$ given by

$$F(x^1) = \int_0^{x^1} f^1(u)du \tag{5.4.20}$$

satisfying

$$\left|\frac{\partial F(x^1)}{\partial x^1}\right|^2 + \frac{\partial^2 F(x^1)}{\partial (x^1)^2} = \gamma[x^1]^2 + \beta x^1 + \delta. \tag{5.4.21}$$

The function $f^0(x^0)$ is assumed to be linear in x^0, thus we set $f^0(x^0) \triangleq B^0 x^0$. Notice that the function $\tanh(x^1)$ satisfies (5.4.21).

As shown earlier, we need to solve the stochastic PDE

$$\frac{\partial}{\partial t}\rho^i(\tilde{x}, t) = L_{\tilde{x}}(t)^* \rho^i(\tilde{x}, t) + c^i x^i \rho^i(\tilde{x}, t) \cdot \frac{dy_t}{dt} \tag{5.4.22}$$

$$\rho^i(\tilde{x}, t_0) = \delta(\tilde{x} - \tilde{x}_0).$$

where

$$L_{\tilde{x}}(t)^* \rho^i(\tilde{x}, t) = \frac{\mathrm{Tr}}{2} \frac{\partial^2}{\partial \tilde{x}} \{\tilde{\sigma}\, \tilde{\sigma}^T \rho^i(\tilde{x}, t)\} - \frac{\partial}{\partial \tilde{x}} (\tilde{f} \rho^i(\tilde{x}, t)) - \frac{1}{2} \tilde{x}^T \tilde{c}^T \tilde{c} \tilde{x} \rho(t, \tilde{x}).$$

The drift and diffusion coefficients $\tilde{f}$, $\tilde{\sigma}$, respectively, are a consequence of the augmented system

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$(\Omega, \mathscr{F}_t, \mathscr{P}_2): \begin{cases} d\tilde{x}_t = \tilde{f}\,(\tilde{x}_t)\,dt + \tilde{\sigma}\,d\tilde{w}_t \quad , \tilde{x}_{t_0} = \tilde{x}_0 \\ \tilde{f} \triangleq \begin{bmatrix} f^1(x_1) \\ f^0(x_0) \end{bmatrix}, \tilde{\sigma} \triangleq \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \end{cases} \qquad (5.4.23)$$

Here, we redefine $x_i \triangleq x^i$, $i=0,1$, $\tilde{c} \triangleq [\alpha_1 c^1 \quad \alpha_0 c^0]$ to avoid complicated notation. The objective for using α_i, $i = 0,1$ and its importance was established in Example 5.4.1. We show in Appendix 7.H that the solution to (5.4.22) is given by

$$e^{-F(x_1)}\rho^i(\tilde{x}, t) = k_t^i\, e^{-\frac{1}{2}(\tilde{x} - \mu_t^i)^T \Sigma_t^{i^{-1}} (\tilde{x} - \mu_t^i)} \qquad (5.4.24)$$

where

$$\dot{\Sigma}_t^i = \tilde{F}\Sigma_t^i + \Sigma_t^i\tilde{F}^T - \Sigma_t^i(\tilde{c}^T\tilde{c} + \tilde{A})\,\Sigma_t^i + \tilde{E}\tilde{E}^T \qquad , \Sigma_{t_0} = 0 \qquad (5.4.25)$$

$$\dot{\mu}_t^i = \tilde{F}\mu_t^i - \Sigma_t^i(\tilde{c}^T\tilde{c} + \tilde{A})\mu_t^i + \Sigma_t\tilde{B} + \Sigma_t\,\tilde{c}^T \cdot \frac{dy_t}{dt} \quad , \quad \mu_{t_0} = \tilde{x}_0 \qquad (5.4.26)$$

$$\begin{aligned} \dot{k}_t^i &= -\frac{1}{2}\mu_t^{i\,T}\Sigma_t^{i^{-1}}\dot{\Sigma}_t^i\Sigma_t^{i^{-1}}\mu_t^i + \mu_t^{i\,T}\Sigma_t^{i^{-1}}\dot{\mu}_t^i - \frac{Tr}{2}(\tilde{E}\Sigma_t^{i^{-1}}\tilde{E}) + \tilde{D} \\ &\quad + \frac{1}{2}\mu_t^{i\,T}\Sigma_t^{i-1}\tilde{E}\tilde{E}^T\Sigma_t^{i-1}\mu_t^i \qquad , \quad k_{t_0}{}^i = 1 \end{aligned} \qquad (5.4.27)$$

where,

$$\tilde{A} = \begin{bmatrix} \gamma & 0 \\ 0 & 0 \end{bmatrix}, \tilde{E} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}^T, \tilde{D} = -\frac{\delta}{2} - B^0$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$\tilde{F} = \begin{bmatrix} 0 & 0 \\ 0 & B^0 \end{bmatrix}, \tilde{B} = \begin{bmatrix} -\beta^0/2 \\ 0 \end{bmatrix}.$$

A simpler expression for k_t^i is given by

$$\dot{k}_t^i = -\frac{1}{2} \mu_t^{i^T} (\tilde{c}^T \tilde{c} + \tilde{A}) \mu_t^i \mu_t^{i^T} \tilde{B} - \frac{\mathrm{Tr}}{2} (\tilde{E}^T \Sigma_t^{-1} \tilde{E}) + \tilde{D} + \mu_t^{i^T} \tilde{c}^T \cdot \frac{dy}{dt}. \tag{5.4.28}$$

Next, we evaluate the integral

$$\int_{R^2} \rho^i(\tilde{x}, t) d\tilde{x} = \int_{R^2} e^{F(x_1)} k_t^i e^{-\frac{1}{2}(\tilde{x} - \mu_t^i)^T \Sigma_t^{i^{-1}} (\tilde{x} - \mu_t^i)} d\tilde{x} \tag{5.4.29}$$

by setting $f^1(x^1) = \tanh(x_1)$. It is shown in Appendix 7.H that (5.4.29) becomes

$$\int_{R^2} \rho^i(\tilde{x}, t) d\tilde{x} = (2\pi)(\det \Sigma_t^i)^{1/2} k_t^i \{ e^{\frac{1}{2} G^T \Sigma_t^i G} \cosh(\mu_t^{i^T} G + G^T \mu_t^i) \} \tag{5.4.30}$$

where $G^T = [1 \ \ 0]$. Hence, the likelihood-ratio test of theorem 5.2.1 can be evaluated by

$$\Lambda(y_t) = \left(\frac{\det \Sigma_t^1}{\det \Sigma_t^0}\right)^{1/2} \frac{k_t^1}{k_t^0} e^{\frac{1}{2} G^T (\Sigma_t^1 - \Sigma_t^0) G} \frac{\cosh \frac{1}{2}(\mu_t^{1^T} G + G^T \mu_t^1)}{\cosh \frac{1}{2}(\mu_t^{0^T} G + G^T \mu_t^0)} \tag{5.4.31}$$

where Σ_t^i, k_t^i, μ_t^i, $i = 0, 1$ satisfy (5.4.25), (5.4.26), (5.4.28).

The performance bounds can also be evaluated by referring to Theorem 5.3.1, expressions (5.3.1), (5.3.2) which in this case are given by

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$P_F \le e^{-s\ln\gamma} E_2\Big[(2\pi)\frac{(\det\Sigma_t^1)^{s/2}}{(\det\Sigma_t^0)^{(s-1)/2}}\frac{(k_t^1)^s}{(k_t^0)^{s-1}}\frac{\{e^{(1/2)G^T\Sigma_t^1 G}\cosh(\mu_t^{1^T}G+G^T\mu_t^1)\}^s}{\{e^{(1/2)G^T\Sigma_t^0 G}\cosh(\mu_t^{0^T}G+G^T\mu_t^0)\}^{s-1}}\Big]$$

$$P_M \le e^{-s\ln\gamma} E_2\Big[(2\pi)\frac{(\det\Sigma_t^1)^{(s+1)/2}}{(\det\Sigma_t^0)^{s/2}}\frac{(k_t^1)^{s+1}}{(k_t^0)^{s}}\frac{\{e^{(1/2)G^T\Sigma_t^1 G}\cosh(\mu_t^{1^T}G+G^T\mu_t^1)\}^{s+1}}{\{e^{(1/2)G^T\Sigma_t^0 G}\cosh(\mu_t^{0^T}G+G\mu_t^0)\}^{s}}\Big].$$

Here, as before, the probability measure $\mathcal{P}_2$ is the Wiener measure $\mu_w^d(dy)$.

Remark 5.4.3 The dimension of the state process x_t^0 of Example 5.4.3 can also be extended to the n-dimensional case as long as $f^0(x_t^0)$ is a linear combination of the components of x_t^0. Notice that the LR (5.4.31) can be evaluated exactly, a situation that does not arise if traditional methods were used.

Example 5.4.4 Nonlinear Decision Problem.

This example is a generalization of the multidimensional filtering problem treated by Beneš [4]. The two-dimensional case was presented by Zeitouni and Bobrovsky [135] whereas the n-dimensional case can be shown to have a finite-dimensional statistics using the estimation algebra. Here, we recognize that the n-dimensional filtering problem is a special case of the quasipotential of a dynamical system used to obtain certain approximations such as exit probabilities. We model the detection problem as follows.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$(\Omega, \mathcal{F}_t, P)\ H_1: \begin{cases} dx_t^1 = \nabla F(x_t)dt + g(x_t) + Bdw_t, & x_{t_0} = x_0 \\ dy_t = Cx_tdt + db_t & , y_{t_0} = 0 \end{cases} \quad (5.4.32)$$

$$(\Omega, \mathcal{F}_t, P_0)\ H_0: \{dy_t = db_t \quad , y_{t_0} = 0 \quad (5.4.32)$$

where, $x_t \in R^n$, $w_t \in R^n$, $y_t \in R^d$. We shall assume

(i) $BB^T = I_n$,

(ii) $F(x_t)$ is a potential function,

(iii) $< \nabla F(x_t), g(x_t) >_1 = 0$,

(iv) $g(x_t)$ an n-dimensional linear function of x,

(v) $\Delta F(x) + ||\nabla F(x)||^2 = \frac{1}{2} x^T \tilde{A}' x + \tilde{B}'^T + \tilde{C}'$.

First, we shall present an expression for evaluating the likelihood-ratio presented by Theorem 5.2.1 which requires the solution of the stochastic PDE

$$\begin{aligned} &\frac{\partial \rho}{\partial t}(x, t) = (L(t)^* - \frac{1}{2} x^T C^T C x)\rho(x, t) + x^T C^T \rho(x, t) \cdot \frac{dy_t}{dt} \\ &\lim_{t \downarrow t_0} \rho(x, t) = \delta(x - x_0) \end{aligned} \quad (5.4.33)$$

where L(t) is the backward Kolmogorov's operator associated with the diffusion process (5.4.32). Proceeding in a similar fashion as in Appendix H, we define

$$\hat{\rho}(x, t) \triangleq e^{-F(x)} \rho(x, t) \quad (5.4.34)$$

where $\hat{\rho}(x,t)$ can be shown to satisfy the stochastic PDE

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$\frac{\partial \hat{p}}{\partial t}(x,t) = \frac{Tr}{2}(B^T \frac{\partial^2 \hat{p}}{\partial x^2}(x,t)B) - \frac{\partial \hat{p}}{\partial x}(x,t)g(x) - (\frac{1}{2}\nabla^2 F(x) + \frac{1}{2}||\nabla F(x)||^2 + \frac{1}{2}x^T C^T Cx + Tr(\frac{\partial g(x)}{\partial x}))\hat{p}(x,t) + x^T C^T \hat{p}(x,t) \cdot \frac{dy_t}{dy}. \tag{5.4.35}$$

By assumptions (iv), (v) we have

$$\nabla^2 F(x) + ||\nabla F(x)||^2 + 2Tr(\frac{\partial g(x)}{\partial x}) + x^T C^T Cx = x^T \tilde{A}x + \tilde{B}x + \tilde{C} \tag{5.4.36}$$

where we define the linear function $g(x) \triangleq Ax$. The solution $\hat{p}(x,t)$ of (5.4.35) is given by

$$\hat{p}(x,t) = (\frac{1}{2\pi})^{n/2} |\Sigma_t|^{-1/2} k_t^{-1} e^{-\frac{1}{2}(x-\mu_t)^T \Sigma_t^{-1}(x-\mu_t)} . \tag{5.4.37}$$

The n-dimensional vector μ_t, nxn matrix Σ_t and scalar function k_t satisfy

$$\dot{\Sigma}_t = A\Sigma_t + \Sigma_t A^T - \Sigma_t \tilde{A}\Sigma_t + BB^T \quad , \Sigma_{t_0} = 0 \tag{5.4.38}$$

$$\dot{\mu}_t = (A - \Sigma_t \tilde{A})\mu_t - \frac{1}{2}\Sigma_t \tilde{B} + \Sigma_t C^T \cdot \frac{dy_t}{dt} \ , \ \mu_{t_0} = x_0 \tag{5.4.39}$$

$$\frac{\dot{k}_t}{k_t} = -\frac{1}{2}\frac{\dot{\det}\Sigma_t}{\det\Sigma_t} + \frac{1}{2}\mu_t^T \Sigma_t^{-1}\dot{\Sigma}_t \Sigma_t^{-1}\mu_t - \mu_t^T \Sigma_t^{-1}\dot{\mu}_t - \frac{1}{2}B^T \Sigma_t^{-1} B - \frac{1}{2}B^T \Sigma_t^{-1}\mu_t \mu_t^T \Sigma_t^{-1} B - \frac{1}{2}\tilde{C}$$
$$k_{t_0} = 1. \tag{5.4.40}$$

Let us now assume that the potential function F(x) is given by

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$F(x) = \int_0^{\frac{1}{2}x^TQx} \tanh(u)du$, and Q is a positive definite nxn matrix. Then

$$F(x)=\ln\cosh(\frac{1}{2}x^TQx)$$

and by (5.4.34) we deduce

$$\rho(x,t)=\cosh(\frac{1}{2}x^TQx)k_t^{-1}(\frac{1}{2\pi})^{n/2}(\det\Sigma_t)^{-1/2}e^{-\frac{1}{2}(x-\mu_t)^T\Sigma_t^{-1}(x-\mu_t)}.$$

Again, using the methodology presented in Appendix 7.H,

$$\int_{R^n}\rho(x,t)dx=\frac{(\det\Sigma_t)^{-1/2}k_t^{-1}}{2(2\pi)^{n/2}}[\int_{R^n}(e^{-\frac{1}{2}(x-B_1\mu_t)^T\tilde{Q}_1^{-1}(x-B_1\mu_t)}+e^{-\frac{1}{2}(x-B_2\mu_t)\tilde{Q}_2^{-1}(x-B_2\mu_t)})dx]$$

where,

$$\tilde{Q}_1^{-1}=Q+\Sigma_t^{-1},\ \tilde{Q}_2^{-1}=-Q+\Sigma_t^{-1},\ B_1=\tilde{Q}_1\Sigma_t^{-1},\ B_2=\tilde{Q}_2\Sigma_t^{-1}$$

and

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$\int_{R^n}\rho(x,t)dx=\frac{(\det\Sigma_t)^{-1/2}k_t^{-1}}{2(2\pi)^{n/2}}\Big[\int_{R^n}\Big(e^{-\frac{1}{2}\mu_t^T(\Sigma_t^{-1}+B_1^T\tilde{Q}_1^{-1}B_1)\mu_t}\times e^{-\frac{1}{2}(x-B_1\mu_t)^T\tilde{Q}_1^{-1}(x-B_1\mu_t)}\Big)dx$$

$$+\int_{R^n}\Big(e^{-\frac{1}{2}\mu_t^T(\Sigma_t^{-1}+B_2^T\tilde{Q}_2^{-1}B_2)\mu_t}\times e^{-\frac{1}{2}(x-B_2\mu_t)^T\tilde{Q}_2^{-1}(x-B_2\mu_t)}\Big)dx\Big]$$

$$=\frac{1}{2}(\det\Sigma_t)^{-1/2}\Big[\det\tilde{Q}_1\, e^{-\frac{1}{2}\mu_t^T(\Sigma_t^{-1}+B_1^T\tilde{Q}_1^{-1}B_1)\mu_t}+\det\tilde{Q}_2\, e^{-\frac{1}{2}\mu_t^T(\Sigma_t^{-1}+B_2^T\tilde{Q}_2^{-1}B_2)\mu_t}\Big]. \tag{5.4.41}$$

Therefore, by Theorem 5.2.1, the likelihood-ratio test is evaluated by

$$\Lambda(y_t)=\frac{1}{2}(\det\Sigma_t)^{-1/2}k_t^{-1}\Big(\det\tilde{Q}_1\, e^{-\frac{1}{2}\mu_t^T(\Sigma_t^{-1}+B_1^T\tilde{Q}_1^{-1}B_1)\mu_t}+\det\tilde{Q}_2\, e^{-\frac{1}{2}\mu_t^T(\Sigma_t^{-1}+B_2^T\tilde{Q}_2^{-1}B_2)\mu_t}\Big)$$

and error probabilities P_F, P_M satisfy the inequalities

$$P_F\le e^{-s\ln\gamma}E_2\Big[\Big(\int_{R^n}\rho(x,t)dx\Big)^s\Big],\ s>0$$

$$P_M\le e^{-s\ln\gamma}E_2\Big[\Big(\int_{R^n}\rho(x,t)dx\Big)^{s+1}\Big],\ s<0.$$

Remark 5.4.4 As stated at the beginning of Section 5.2, the bounds on the error probabilities derived using our approach preclude having to generate signal estimates $\hat{h}^i$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

that may come from an infinite-dimensional filter because of the moment-closure problem. However, the need to perform an expectation with respect to measure ρ_2, even though ρ_2 is a Wiener measure, still remains and may require numerical approximation in some cases. Whenever the stochastic PDE does not have a closed form solution, the need for approximations is required which is propagated to the last stage.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

CHAPTER 6

CONCLUSION

6.1 SUMMARY AND MAIN CONTRIBUTIONS

In this thesis we presented results in nonlinear estimation, nonlinear stochastic partially observed control, and nonlinear decision using measure transformations.

6.1.1 The Nonlinear Estimation Problem

Several authors have previously addressed this problem, using either Lie algebraic methods or probabilistic methods. Our approach is basically probabilistic and relies heavily on gauge transformations and on the exact linearization of stochastic systems using diffeomorphisms and measure transformations. As pointed out by Brockett [27] diffeomorphisms and gauge transformations result in equivalent filtering problems. Therefore, by first linearizing the diffusion equation and measurement equation and then applying a gauge transformation we derive sufficient conditions for solving a Feynman-Kac formula, and this solution is related to the unnormalized conditional density. The above linearization techniques were also used by Cohen and Levine [34] who independently derived a set of sufficient statistics for obtaining the solution to the unnormalized conditional density for the global case. Their approach is completely geometric and does include provide our derivation of the initial-boundary value problem presented in Sections 2.5, 3.5 for the case when the linearization is only valid locally. Moreover, we provide lower and upper bounds on functions of state estimates which

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

appear here for the first time. The extension of the Bobrovsky and Zakai [23, 24] lower bound to degenerate diffusions also appears here for the first time and is due to the linearization of nonlinear stochastic systems.

6.1.2 The Partially Observed Stochastic Control Problem

The stochastic control problem considered in Chapter 4 is treated here for the first time. The derivation of the stochastic PDE satisfied by the unnormalized conditional density is based on the derivations presented by Pardoux [114] and Kunita [96]. As pointed out in Section 4.1 this stochastic PDE has as a special case the stochastic PDE derived by Pardoux [114] who considered the case when the diffusion process is of nondegenerate form with an additional condition satisfied. The equations satisfied by the decomposed measure-valued processes in both forward and backward variables, and the equations satisfied by the inverse measure-valued process are extensions of the decomposition of Kunita [94, 96], who first provided the decomposition of the unnormalized conditional density which is our original motivation. The similarity between adjoint-processes in Euclidean space and in L^2 space are exploited here for the first time. Both approaches in obtaining the minimum principle and equations satisfied by the adjoint-processes are new. Also the case when correlation is allowed between the state process and the observation process is addressed here for the first time. Furthermore, the case when no correlation is present but the diffusion process is allowed to depend on the control variable is also new, in that we present an explicit equation satisfied by the adjoint-process, whereas Elliott and Yang [50], who treated the same problem using different methods, do not provide an explicit equation for the adjoint-

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

process. Finally, our derivations are rigorous compared to the ones presented by Bensoussan [15] in that no approximations are considered, whereas the above author is forced to approximate the Sobolev space by using a finite-dimensional basis to approximate and then derive the equation satisfied by the adjoint-process.

6.1.3 The Nonlinear Decision Problem

The derivations of the results for the decision problem in Chapter 5 appear to be new. Although similar results for error probabilities using Chernoff bounds have previously been established by Evans [52] and Hibey [74, 75], who express these bounds in terms of the solution of a Feynman-Kac formula whose coefficients are the optimal estimates, expressing these bounds in terms of stochastic PDE's are first presented here. Expressing the generalized likelihood-ratio in terms of a ratio of two stochastic PDE's integrated over the whole space also appear here for the first time. Finally, to the best of our knowledge, the exact calculation of the generalized likelihood-ratio and error bounds (assuming the Wiener expectation can be performed exactly) for binary nonlinear decision problems appears here for the first time. Notice also that no approximations were considered, whereas if one uses traditional methods certain approximations involving suboptimal filtering estimates have to be introduced for the nonlinear decision problem only.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

6.2 TOPICS FOR FURTHER RESEARCH

1. A method for solving the initial-boundary value problem encountered in Chapter 3 would be very desirable. Furthermore, solving filtering problems that are defined locally instead of globally is a very interesting problem.

2. In Chapter 4 we have been exclusively concerned with control problems having bounded drift and diffusion coefficients. The extension to the linear case (i.e., when the drift and diffusion coefficients as well as the signal satisfy a linear growth condition) would be very interesting. Indeed we believe that the methodology used to derive the minimum principle and adjoint process can be extended to the case when assumption (A′1) through (A′10) are weakened to cover the case of stochastic equations with linear growth, at least for the case when $h(t, x_t)$ is a bounded function.

3. Recalling Approach 1 presented in Chapter 4, where we derived the necessary conditions of optimality using deterministic control ideas, it would be desirable repeat the same approach when constraints are incorporated (i.e., terminal constraints).

4. A method of solving the control problem of Chapter 4 when the admissible control set is not convex would be very desirable.

5. Treating the control problem of Chapter 4 when the admissible controls are of wide sense would be a challenging problem.

6. Hibey [74, 75] derives Chernoff bounds for the case of discontinuous observations. It would be interesting to pursue the decision problem of Chapter 5 in this context when the observation process is a Poisson point process whose intensity is a given function of the unobserved state process. Filtering results for Poisson point processes

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

using the unnormalized conditional density are found in Pardoux [114, 115] and Gertner [64].

7. Asymptotic estimates to the solution of the DMZ equation when the noise terms are multiplied by a small nonrandom parameter and the time is a fixed interval were given by Hijab [76] in terms of the solution to a deterministic control problem. These bounds could be used to find asymptotic estimates for the error probabilities. It would be interesting to derive similar asymptotic estimates when the diffusion process is defined up to a stopping time τ and use then to obtain asymptotic estimates for the sequential decision problem.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

7. APPENDICES

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

APPENDIX 7.A
DIFFERENTIAL GEOMETRY AND RELATED TOPICS

In this section, we present only those concepts of differential geometry that will be useful to us in Chapter 3, where linearization of stochastic differential equations is considered. For more detailed information on state-feedback linearization we refer to Isidori [79], Nijmeijer and Shaft [108] and Brockett, Millman and Sussmann [30], Su [124], Hunt, Su and Meyer [77], Cheng and Isidori [32] and Dayawansa, Boothby and Elliott [41]. For detailed information on differential geometry on manifolds, we suggest Boothby [25].

Throughout this section we define certain concepts of differential geometry and state the notation used for understanding the result of state-feedback linearization of control systems. We begin with the definition of a manifold, define the Lie bracket of vector fields and define the Lie bracket of a scalar field with respect to a vector field. Next, we give the characterization of a distribution, and definitions of involutive and integrable distribution. Finally, we present the famous Frobenius Theorem and its application to state-feedback linearization.

Definition: Manifold.

A manifold M of dimension n is a topological space having the following properties:

(i) M is Hausdorff;

(ii) M is Locally Euclidean;

(iii) M has a countable basis of open sets.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Every open set containing an open neighborhood of the point x is referred to as a neighborhood of x.

Definition: Vector Field.

Let M be an n-dimensional manifold. A vector field f on M is a mapping assigning to each point $x \in M$ a tangent vector $f(x)$ in the tangent space of M at x.

Definition: Tangent Space to M.

The tangent space to M at x, denoted T_xM, is the set of all tangent vectors at x.

Definition: Lie Bracket of Vectors Fields.

Suppose we are given two vector fields on R^n (or any manifold M). The Lie bracket [f, g] is also a vector field on $R^n(M)$ and defined as

$$[f, g] \triangleq \frac{\partial g}{\partial x} f - \frac{\partial f}{\partial x} g$$

where $\frac{\partial g}{\partial x}, \frac{\partial f}{\partial x}$ denote n x n Jacobian matrices. It represents the Lie derivative of one vector field with respect to another. It is also denoted by $ad_f^1(g)$ and by induction we can define

$$ad_f^k(g) = [f, ad_f^{k-1}(g)], \; ad_f^o(g) = g.$$

For more information on the properties of Lie bracket, see Isidori [79, pp. 9].

Definition: Lie Derivative of a Function Along a Vector Field.

Let $h: R^n \rightarrow R$ be a scalar field, with gradient denoted by dh which is a row vector field

$$(\frac{\partial h}{\partial x_1}, \dots, \frac{\partial h}{\partial x_n})$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

on R^n. The Lie derivative of h along a vector field $f = (f_1, ..., f_n)^T$, denoted by $\langle dh, f\rangle_1$, is a scalar field defined by

$$\langle dh, f\rangle_1 = \frac{\partial h}{\partial x_1} f_1 + ... + \frac{\partial h}{\partial x_n} f_n.$$

The above definition is often written as $L_f h$. Also, it follows easily that

$$L_g L_f h = \frac{\partial}{\partial x} (L_f h) g,$$

and, by induction,

$$L_f^k h = \frac{\partial}{\partial x}(L_f^{k-1} h) f.$$

<u>Remark 7.A.1</u> If M is a smooth manifold of dimension n and x any point of M, the tangent space T_xM to M at x is an n-dimensional vector space over the field R. If (U,ϕ) is a local system of coordinates around x, then the tangent vectors $\{\frac{\partial}{\partial \phi_1}, ..., \frac{\partial}{\partial \phi_n}\}_x$ form a basis for T_xM.

<u>Definition</u>: Distribution.

Suppose we are given d vector fields $f_1, ..., f_d$, all defined on R^n. Then at any point $x \in R^n$, the vectors $f_1, ..., f_d$ span a vector space (a subspace of R^n). Δ called a distribution and denoted by

$$\Delta \triangleq \text{span}\ \{f_1, ..., f_d\}.$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Remark 7.A.2 A distribution Δ defined on a manifold M is nonsingular if there exists an integer d such that $\dim(\Delta) = d$ for all $x \in M$, that is, the vector fields $f_1, ..., f_d$ are linearly independent $\forall x \in M$.

Definition: Involutive Distribution.

A distribution $\Delta = \text{span}\{f_1, ..., f_d\}$ defined on R^n is called involutive if there exist scalar fields γ_{ijk} such that

$$[f_i, f_j] = \sum_{k=1}^{d} \gamma_{ijk} f_k, \; 1 \le i, j \le d, \; i \ne j.$$

It can be shown that the distribution Δ is completely integrable if and only if there are n-d linearly independent scalar fields $h_1, ..., h_{n-d}$ such that

$$\langle dh_i, f_j \rangle_1 = 0, \quad 1 \le i \le n-d, \quad 1 \le j \le d.$$

The concepts of involutiveness and complete integrability are connected by the Frobenius theorem.

Theorem 7.A.1 Frobenius Theorem.

A nonsingular distribution is completely integrable if and only if it is involutive.

Proof: See Boothby [ ].

We shall be interested in the application of the above concepts and definitions in a neighborhood of a given point. Concepts and definitions of this type will be called local.

Deterministic State-Feedback Linearization.

Consider a nonlinear control system of affine form

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$\Sigma_1: \dot{x}_t = f(x_t) + g(x_t)u \tag{7.A.1}$$

where $x \in R^n$ and f(x), g(x) are C^∞ vector fields on R^n (R^n can also be replaced by a C^∞ n-dimensional manifold M). State-feedback linearization characterizes the Σ_i, i = 1, 2 equivalent class of systems which contains controllable linear systems. That is, the control system (7.A.1) can be transformed into the linear controllable system

$$\Sigma_2: \dot{z}_t = Az_t + Bv \tag{7.A.2}$$

where A and B are, respectively, n x n and n x 1. The two operations leading to system (7.A.2) are state-feedback $u = \alpha(x) + \beta(x)v$ and coordinate transformation $z = \Phi(x)$ with $\Phi(x)$ being a diffeomorphism.

<u>Definition</u>: Locally Linearizable Systems.

The nonlinear control system (7.A.1) is said to locally linearizable at a given point x^0 if there exist a neighborhood U of x^0, a feedback $u = \alpha(x) + \beta(x)v$ defined on U, and a coordinate transformation $z = \Phi(x)$ also defined on U, such that the corresponding closed loop equation

$$\dot{x} = f(x) + g(x)\,\alpha(x) + g(x)\,\beta(x)v \tag{7.A.3}$$

in the coordinate $z = \Phi(x)$, is linear and controllable, that is, such that

$$\frac{\partial \Phi}{\partial x}\,(f(x) + g(x)\,\alpha(x))\,\Big|_{x=\Phi^{-1}(z)} = Az$$

$$\frac{\partial \Phi}{\partial x}\,(g(x))\;\beta(x)\,\Big|_{x=\Phi^{-1}(z)} = B$$

for some matrix $A \in R^{nxn}$ and vector $B \in R^n$ satisfying the condition

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$\text{rank}\,(B, AB, \ldots, A^{n-1}B) = n.$$

The previous definition implies that Φ is a diffeomorphism onto an open set of R^n when defined on U. The equivalent definition of a globally linearizable system on M requires that there exists a C^∞ diffeomorphism Φ: $M \rightarrow R^n$ such that Φ transforms system (7.A.1) to (7.A.2) and Φ is onto R^n; global results are presented by Dayawansa, Boothby and Elliott [41].

We are now in a position to state theorems regarding necessary and sufficient conditions for local and global state-feedback linearization of nonlinear control systems. Results on local linearization are found in Isidori [79] for systems with multiple inputs, and Su [124] for the single input case. Hunt, Su and Meyer [77] give a sufficient conditions for global linearizability, again for the single input case. For extending locally linearizable systems to globally linearizable system, see Boothby [25]. Results on global linearization for multiple inputs are found in Dayawansa, Boothby and Elliott [41].

Theorem 7.A.2

Suppose we are given the single input system Σ_1 defined on a C^∞-manifold M of dimension n, where f, g are C^∞ vector fields. Then (a) Σ_1 is locally linearizable in a neighborhood U of x^0 if and only if

(i) $\{g, ad_f(g), \ldots, ad_f^{n-1}(g)\}x^0 = T_{x^0}M$,

(ii) $\Delta = \text{span}\{g, ad_f(g), \ldots, ad_f^{n-2}(g)\}_{x^0}$ is involutive near x^0,

Proof: See Isidori [79, Theorem 2.6, p. 165];

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

and (b) Σ_1 is globally linearizable if and only if

(i) (a)(i) is satisfied $\forall x \in M$,

(ii) $\eta(ad^{i-1}(f)g) = (-1)^n \delta_{in}$, $i \leq n$, is closed

(iii) $g, ad_{\hat{f}} g, \ldots, ad_{\hat{f}}^{n-1}(g)$ are complete where $\hat{f} \triangleq f - L_f^{n-1}(\eta(f))g$.

Proof: See Dayawansa, Boothby and Elliott [41, Thm. 4].

The procedure that leads to the local construction of the coordinate and feedback transformations $z = \Phi(x)$ and $u = \alpha(x) + \beta(x)v$, respectively, is as follows:

- Construct vector fields g, $ad_f(g)$, ..., $ad_f^{n-1}(g)$ and check conditions (a)(i), (a)(ii);

- If both are satisfied then, using Frobenius Theorem 7.A.1 there exists a scalar field $h(x)$ such that

$$\langle dh(x), ad_f^j(g)\rangle_1 = 0, \quad 0 \leq j \leq n-2 \quad \text{around } x^0, \text{ and}$$

- set $\alpha(x) = -\dfrac{L_f^n h(x)}{L_g L_f^{n-1} h(x)}$, $\beta(x) = \dfrac{1}{L_g L_f^{n-1} h(x)}$

$$\Phi(x) = (h(x), L_f h(x), \ldots, L_f^{n-1} h(x))^T;$$

then the new coordinate system Σ_2 of (7.A.1) has the form

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$A = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & 0 \\ \vdots & & \ddots & & \vdots \\ & & & & 1 \\ 0 & 0 & \dots & \dots & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}. \tag{7.A.4}$$

The multi-input multi-output state-feedback linearization is found in Isidori [79, Thm. 2.4, pp. 250].

<u>Remark 7.A.3</u> If condition (a)(i) of Theorem 7.A.2 is satisfied, then the nonlinear system under consideration is said to be controllable. Nonlinear controllability and observability is investigated by Hermann and Krener [73].

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

APPENDIX 7.B
MARTINGALES AND RELATED TOPICS

In this section we borrow some concepts from the theory of martingales that are of extreme importance in analyzing stochastic processes. Since some of these concepts are quite technical, a complete introduction is not possible. Therefore, we present here only those concepts that are suitable to our needs with references cited as needed. However, for a detailed introduction to the theory, we refer to Liptser and Shiryayev [103], Kallianpur [84], Wong and Hajek [132], Karatzas and Shreve [86], Elliott [47] and van Schuppen [127]. For a first time exposure to this subject we suggest Oksendal [111].

<u>Definition</u>: Filtration.

Suppose $(\Omega, \mathcal{F}, \mathcal{P})$ is a complete probability space. A filtration $\{\mathcal{F}_t, t \geq 0\}$ of $(\Omega, \mathcal{F})$ is a nondecreasing family of sub-σ-fields $\mathcal{F}_t$, $t \in [0, T]$, of $\mathcal{F}$ such that $\mathcal{F}_s \subseteq \mathcal{F}_t$ for $0 \leq s \leq t < \infty$.

<u>Remark 7.B.1</u> The family of σ-fields $\{\mathcal{F}_t, t \geq 0\}$ can be considered as describing the history of some phenomenon and $\mathcal{F}_t$ is sometimes called the σ-field of events prior to time t. Let us define $\mathcal{F}_{t^-}$ to be the minimum σ-field of events strictly prior to $t > 0$ and $\mathcal{F}_{t^+}$ to be the σ-field of events immediately after $t \geq 0$. We say that the filtration $\{\mathcal{F}_t\}$ (i) is right-(left-)continuous if $\mathcal{F}_t = \mathcal{F}_{t^+}(\mathcal{F}_{t^-})$ holds for every $t \geq 0$, and (ii) complete if $\mathcal{F}_0$ contains all the $\mathcal{P}$-null sets in $\mathcal{F}$. Throughout this thesis, we will assume that conditions (i) and (ii) are satisfied by saying that $\mathcal{F}_t$ satisfies the usual conditions.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Definition: Adapted Process.

A stochastic process x_t, $t \in [0, T]$, is said to adapted to the family $\{\mathcal{F}_t, t \in [0, T]\}$ if, for all $t \in [0, T]$, x_t is $\mathcal{F}_t$ measurable. We shall denote this by writing $x_t \sim \mathcal{F}_t$. Note that, $x_t \sim \mathcal{F}_t^x \triangleq \sigma(x_s, s \leq t)$.

Stopping Times. A nonnegative random variable $\tau: \Omega \rightarrow [0, \infty]$, is called a stopping time of a given family $\mathcal{F}_t$, $t \in [0, \infty)$, if $\{\omega \in \Omega; \tau(\omega) \leq t\} \in \mathcal{F}_t$ for every $t \in [0, \infty)$. A similar notion is defined for a family $\mathcal{F}_t$, $t \in [0, T]$, by assuming $\tau \in [0, T]$.

The importance of stopping times is significant when we investigate the properties of a process defined only locally. For more discussion on stopping times we refer to Wong and Hajek [132], Kallianpur [84], and van Schuppen [127, Chp. 1].

Definition: Martingale.

A continuous stochastic process m_t, $t \in [0, T]$ is called a martingale with respect to a given family of σ-fields $\mathcal{F}_t$, $t \in [0, T]$, satisfying the usual conditions if:

(i) m_t is adapted to $\mathcal{F}_t$ for all $t \in [0, T]$;

(ii) $E|m_t| < \infty$ for all $t \in [0, T]$;

(iii) $E[m_t|\mathcal{F}_s] = m_s$ a.s for all $s, t \in [0, T]$, $s < t$.

We note that the above definition requires the specification of some family of σ-fields $\mathcal{F}_t$ and some measure say $\wp$. We shall denote the class of all martingales by M, and we emphasize the dependence on $\wp$ by writing $m_t \in M(\mathcal{F}_t, \wp)$.

Next, we introduce the following classes of martingales with continuous sample paths.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Definition: Local Martingale.

A stochastic process m_t, $t\in[0, \infty)$, which is adapted to some increasing family of σ-fields $\mathcal{F}_t$ $t\in[0, \infty)$, satisfying the usual conditions, is a local martingale with respect to $\mathcal{F}_t$, $t\in[0, \infty)$ if there is a sequence of stopping times $\{\tau_n\}_{n\geq 1}$ such that τ_n converges to ∞ a.s (or, in the case of the interval $[0, T]$, converging to T a.s) and for each n, $m_{t\wedge\tau_n}$ is a martingale with respect to $\mathcal{F}_t$.

The class of martingales defined above is denoted by M_{loc} and we shall use the notation $m_t\in M_{loc}(\mathcal{F}_t, P)$.

Definition: Square and Locally-Square Integrable Martingales.

The class of square-integrable martingales, denoted by M_2, is defined as

$$M_2 \triangleq \{m_t \in M(\mathcal{F}_t);\ E|m_t|^2<\infty\}.$$

If the time interval $[0, T] \rightarrow [0, \infty)$ then $E|m_t|^2 < \infty \rightarrow \sup_t E|m_t|^2 < \infty$.

The class of locally square-integrable martingales, denoted by M_{2loc}, is defined as

$$M_{2loc} \triangleq \{m_t \in M_{loc}(\mathcal{F}_t);\ \text{there exist a sequence}$$

of stopping times $\{\tau_n\}_{n\geq 1}$ such that τ_n converges to ∞ a.s and for each n, $m_{t\wedge\tau_n} \in M_2(\mathcal{F}_t)\}$.

Moreover, for continuous martingales we have $M_{loc} = M_{2loc}$ which can be shown by a stopping time argument.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Definition: Predictable Processes.

The quadratic variation process and the stochastic integration theory that we shall define later require the notion of a predictable σ-field and predictable process. Suppose that we are given a family of σ-fields $\mathscr{F}_t$, $t \in [0, \infty)$ satisfying the usual conditions. The smallest σ-fields $\bar{\mathscr{F}}_t$ of subsets of $[0, \infty) \times \Omega$ with respect to which all processes $x_t(\omega)$ are (i) adapted to $\mathscr{F}_t$ and (ii) for each ω, is left-continuous a.s., is called a predictable σ-field. A stochastic process x_t, $t \in [0, \infty)$ is called predictable if it is measurable with respect to the predictable σ-field $\bar{\mathscr{F}}_t$, that is, any process which is measurable with respect to $([0, \infty) \times \Omega, \bar{\mathscr{F}}_t)$.

It follows that if $x_t \sim \mathscr{F}_t$ and its sample paths are continuous, a.s., then x_t is a predictable process with respect to $\mathscr{F}_t$. For more discussion on predictable processes, we refer to Kallianpur [84, pp. 48-49]. For discussion on processes that are not left-continuous having left-hand limits we refer to van Schuppen [127, pp. 7-9].

Theorem 7.B.1 The Quadratic Variation Process.

To every martingale $m_t \in M_{loc}(\mathscr{F}_t)$ there corresponds a process $\langle m, m\rangle_t - \langle m, m\rangle_0 = a_t$, called the quadratic variation process of m, with the properties:

(i) $a_0 = 0$ almost surely;

(ii) $a_t \geq a_s$ almost surely for $t \geq s$;

(iii) a_t is continuous and $\mathscr{F}_t$-adapted;

(iv) $m^2_t - a_t$ is a local martingale.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Remark 7.B.2 Let $m_t \in M_{loc}(\mathscr{F}_t)$ with variation process $<m, m>_t$. If E $<m, m>_t < \infty$, for any $t \in [0, T]$, then $m_t \in M_2(\mathscr{F}_t)$ for $t \in [0, T]$. Moreover, if $m_t, n_t \in M_2, (\mathscr{F}_t)$ then we define the quadratic covariation process

$$\langle m,n\rangle > \triangleq \frac{1}{2}(\langle m+n, m+n\rangle_t - \langle m,m\rangle - \langle n,n\rangle_t).$$

Remark 7.B.3 If the two processes m_t, n_t take values in R^k and R^r, respectively, and $m_t, \in M_2(\mathscr{F}_t)$, $n_t \in M_2(\mathscr{F}_t)$ then $<m, n>_t$ with $<m, n>_0 = 0$ is a matrix with each element defined as $<m^i, n^j>_t$, $i = 1, \ldots, k$ and $j = 1, \ldots, r$.

Remark 7.B.4 The Wiener process $x_t \in R^n$, $t \in [0, T]$ defined above has $E(x_t) = 0$ for all t and covariance function given by $E(x_t\, x_s) = \sigma \min(t, u)$ where σ is an n x n diagonal matrix. If $\sigma = I$, the identity matrix, then $x_t \in R^n$ is called a standard Brownian motion. The following conditions given by Kunita and Watanabe [97, Them. 2.3] are sufficient conditions for a process to be a Wiener process: (i) $x_t \in M_{loc}(\mathscr{F}_t)$ and (ii) $<x, x>_t = It$.

All of the processes encountered in this thesis will be modeled as semimartingales (SM) or local semimartingales (SM_{loc}), which are defined as follows.

Definition: SM and SM_{loc}.

An n-dimensional continuous stochastic process x_t, $t \in [0, T]$ of the form $x_t = x_0 + a_t + m_t$, where $a_0 = m_0 = 0$ a.s, is called a continuous semimartingale if for each i (i) E $[x_0^i] < \infty$, (ii) $a_t^i \sim \mathscr{F}_t$, the function $a_t^i(\omega_t)$ is, for almost all ω, continuous, (iii) a_t^i is of bounded variation, (iv) $E\,|\,a^i\,|_t < \infty$, (v) the function $t \to m_t^i(\omega)$ is, for

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

almost all ω, continuous, and (v) $m_t \in M(\mathscr{F}_t)$. It is called a local semimartingale if $m_t \in M_{loc}(\mathscr{F}_t)$ and (iv) is removed.

Stochastic Integration.

We shall be concerned with representing integrals when the integration is with respect to certain classes of martingales m, in either Ito or Fisk-Stratonovich (F-S) forms, in both forward and backward variables. Background materials on the subject can be found in Wong and Hajek [132], Liptser and Shiryayev [103], Karatzas and Shreve [86], Kallianpur [84], and Kunita [96]. A wonderful introduction on this subject for the case of continuous stochastic processes is found in Kunita [96]. (The martingales m that we shall consider will be Brownian motion process, adapted to increasing as well as decreasing filtrations).

In all that follows, we assume that we are given either an increasing family of σ-fields $\mathscr{F}_s^t$, $t \in [0, T]$ with s a fixed time $s \in [0, T]$ or a family of decreasing σ-fields $\mathscr{F}_s^{t_1}$, $s \in [0, t_1]$ with t_1 a fixed time $t_1 \in [0, T]$, such that $m_t \sim \mathscr{F}_s^t$ or $m_t \sim \mathscr{F}_s^{t_1}$ respectively. We shall use the notation $(h.m)_t \triangleq h_0\, m_0 + \int_0^t h_u\, dm_u$ to denote forward integrals and the notation $(h.m)_s \triangleq h_{t_1}\, m_{t_1} + \int_s^{t_1} h_u\, \hat{d}m_u$ to denote backward direction integrals. We define the following two classes of integrand processes:

If $m \in M_2(\mathscr{F}_s^t)$, then $\mathscr{L}_2(\langle m, m\rangle_t) \triangleq \{h_t;\, h_t$ is predictable and $E[\int_0^T |h_u|^2 d\langle m, m\rangle_u] < \infty\}$.

If $m \in M_{loc}(\mathscr{F}_s^t)$ then $\mathscr{L}_{2loc}(\langle m, m\rangle_t) \triangleq \Big\{h_t;\ h_t$ is predictable and there exist an increasing

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

sequence of stopping times $\{\tau_n\}$ converging to ∞ a.s. such for all n, $E[\int_0^{\tau_n} |h_u|^2 d\langle m, m\rangle_u] < \infty\}$.

Moreover, if $h_t \in M_{2loc}(\mathscr{F}_s^t)$ we shall say that h_t is locally in $\mathscr{L}_2$ ($<m, m>_t$).

The definition of the two classes of integrand processes above are given in terms of forward integrals. The definition of integrand processes with respect to a backward variable can be established similarly.

With the above definitions and notations in mind, we state the following properties of stochastic integrals.

Theorem 7.B.2 Stochastic Ito Integrals $\mathscr{L}_2$.

If $m_t \in M_2(\mathscr{F}_s^t)$ and $h_t \in \mathscr{L}_2(\langle m, m\rangle_t)$, then there exists a unique stochastic integral $(h.m)_t \in M_2(\mathscr{F}_s^t)$ such that for all $n_t \in M_2(\mathscr{F}_s^t)$, $\langle (h.m), n\rangle_t = (h.\langle m,n\rangle)_t = h_s\langle m,n\rangle_s + \int_s^t h_u d\langle m, n\rangle_u$.

Remark 7.B.5 If the martingales m_t and n_t are in the class $M_{2loc}(\mathscr{F}_s^t)$ and h_t is locally in $\mathscr{L}_2(\langle m, m\rangle_t)$ then there exists a unique stochastic integral $(h.m)_t \in M_{2loc}(\mathscr{F}_s^t)$ and the properties of Theorem 7.B.2 remain valid.

Remark 7.B.6 If $x_t \in SM_{loc}$ with the decomposition $x_t = x_s + \alpha_t + m_t$ where $m_t \in M_{loc}(\mathscr{F}_s^t)$ and h_t is predictable and locally bounded, then we define

$(h.x)_t = \int_s^t h_u \, dx_u$ as

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$(h.x)_t \triangleq h_0 x_0 + (h.\alpha)_t + (h.m)_t.$$

The integral $(h\,.\,\alpha)_t \triangleq \int_s^t h_u \, d\alpha_u$ is a stochastic Stieltjes integral that is well-defined because α_t is of bounded variation. Furthermore, we have $(h.x)_t$ is adapted to $\mathcal{F}_s^t$. The proof analogous results for the backward direction integrals are found in Kunita [96].

Everything that we have said until now about Ito integrals remains valid for the case of Fisk-Stratonovich integrals. There is however an important difference: the latest integral is not always well-defined in that the limit does not always exist. However even if it does exist, it does not coincide with the value of the Ito integral. The Ito and F-S integrals are related by

$$\int_s^t f(x_r) \circ dw_r \triangleq \int_s^t f(x_r)_r dw_r + \frac{1}{2}\left(\langle f(x), w\rangle_t - \langle f(x), w\rangle_s\right). \qquad (7.B.1)$$

Proof: See Kunita [96].

We are now in the position to introduce the stochastic differential rules of Ito and F-S in the forward and backward direction.

Theorem 7.B.3 The Forward Ito Differential Rule.

Let $X_t(x)$, $t \in [s, T]$ be a real-valued continuous local semimartingale adapted to $\mathcal{F}_s^t$, $t \in [0, T]$, s fixed, $x_s = x$ having solution $X_{s,t}(x)$, and let $f: R^n \rightarrow R^1$ be a twice

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

continuously differentiable function with $f_{x^j} = \dfrac{\partial f}{\partial x^j}$, $1 \le j \le n$ and $f_{x^i x^j} = \dfrac{\partial^2 f}{\partial x^i \partial x^j}$,

$1 \le i, j \le n$. Then $f(X_{s,t}(x))$, $t \in [s, T]$ is again a local $\mathcal{F}_s^t$-semimartingale and satisfies:

$$f(X_{s,t}(x)) - f(x) = \sum_{j=1}^{n} \int_s^t f_{x^j} (X_{s,r}(x))\, dX_{s,r}^j(x) + \frac{1}{2} \sum_{i,j}^{n} \int_s^t f_{x^i x^j} (X_s(x))\, d\langle X^i, X^j\rangle_r \quad (7.B.2)$$

where $d\langle X^i, X^j\rangle_r = \gamma_r^{ij}$.

Proof: See for example, Kallianpur [84].

Remark 7.B.7 By Remark 7.B.6 the differential form (7.B.2) for the case when f also depends on t, is defined as,

$$df(X_{s,t}(x)) = L(t)\, f(X_{s,t}(x))\, dt + \sum_{j=1}^{m} f_{x^j} (X_{s,t}(x)) dm_t^j \quad (7.B.3)$$

where

$$L(t) \triangleq \sum_{j=1}^{n} a^i \frac{\partial}{\partial x^j} + \frac{1}{2} \sum_{i,j}^{n} \gamma_t^{ij} \frac{\partial^2}{\partial x^i \partial x^j} + \frac{\partial}{\partial t}.$$

Theorem 7.B.4 The Backward Ito Differential Rule.

Let $\hat{X}_{s,t_1}(x)$, $s \in [0, t_1]$ be a real-value continuous backward local semimartingale adapted to $\mathcal{F}_s^{t_1}$, $s \in [0, t_1]$, t_1 fixed and $X_{t_1} = x$, having solution $\hat{X}_{s,t_1}(x)$. Let $f: R^n \to R^1$ be a twice continuously differentiable function, with f_{x^j}, $f_{x^i x^j}$ as defined in Theorem

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

7.B.3. Then $f(\hat{X}_{s,t_1}(x))$, $s \in [0, t_1]$ is again a local $\mathcal{F}_s^{t_1}$ - backward semimartingale for $s \in [0, t_1]$ and satisfies:

$$\hat{X}_{s,t_1}(x) - f(x) = \sum_{j=1}^{n} \int_s^{t_1} f_{x^j} (\hat{X}_{r,t_1}(x))\, d\hat{X}^j_{r,t_1}(x) + \frac{1}{2} \sum_{i,j}^{n} \int_s^{t_1} f_{x^i x^j}(\hat{X}_{r,t_1}(x)) d\langle \hat{X}^i, \hat{X}^j \rangle_r \qquad (7.B.4)$$

where $d\langle X^i, X^j \rangle_r = \gamma_r^{ij}$.

Remark 7.B.8 Again, using Remark 7.B.6 the differential form of (7.B.4) for the case when f depends on s also, is defined as

$$df(\hat{X}_{s,t_1}(x)) = - L(s)\; f(\hat{X}_{s,t_1}(x))ds - \sum_{j=1}^{m} f_{x^j} (\hat{X}_{s,t_1}(x))\, \hat{d}m_s^j \qquad (7.B.5)$$

where L(s) is the same as defined in Remark 7.B.7 with the time derivative having opposite sign.

Theorem 7.B.5 The Fisk-Stratonovich Forward Differential Equation.

Let $X_t(x)$, $t \in [s, T]$ be a real-valued continuous local semimartingale adapted to $\mathcal{F}_s^t$, $t \in [s, T]$, s fixed, $x_s = x$, having solution $X_{s,t}(x)$. Let $f: R^n \to R^1$ be a three times continuous differentiable function, with f_{x^j}, as defined before, then $f(x_{s,t}(x))$, $t \in [s, T]$ is again a local $\mathcal{F}_s^t$-semimartingale for $t \in [s, T]$ and satisfies

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$f(X_{s,t}(x)) - f(x) = \sum_{j=1}^{n} \int_{s}^{t} f_{x^j}(X_{s,r}(x)) \circ dX^{j}_{s,r}(x) \tag{7.B.6}$$

Remark 7.B.9 The advantage of writing differential rules with respect to F-S integrals rather than Ito integrals is obvious from (7.B.6). That is F-S differential rules obey the classical rules of calculus and sometimes provides a more physical interpretation. However, the expectation of a F-S integral is not zero in general, and the moment or variance is not easily computed, unless it is transformed to an Ito integral. The above equation (7.B.6) can be transformed to an Ito equation via (7.B.1). The differential form of (7.B.6) is defined as

$$df(X_{s,t}(x)) = \sum_{j=1}^{\eta} f_{x^j}(X_{s,t}(x)) \circ dX^{j}_{s,t} \tag{7.B.7}$$

Theorem 7.B.6 The Fisk-Stratonovich Backward Differential Equation.

If $\hat{X}_{s,t_1}(x)$ be defined similarly as in Theorem 7.B.4, then the corresponding function $f(\hat{X}_{s,t}(x))$ satisfies

$$f(\hat{X}_{s,t_1}(x)) - x = \sum_{j=1}^{m} \int_{s}^{t_1} f_{x^j}(\hat{X}_{r,t_1}(x)) \cdot d\hat{X}^{j}_{r,t_1}. \tag{7.B.8}$$

Remark 7.B.10 The differential form of (7.B.8) is defined as in Remark 7.B.8 by omitting the term resulting from the double summation. It is also related to Ito backward Integral form in a similar manner as (7.B.1) see Kunita [96].

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

APPENDIX 7.C
STRONG AND WEAK SOLUTIONS

We are now in the position to introduce the concepts of weak and strong solutions of stochastic differential equations. Throughout, we shall assume that stochastic integrals are well defined. Background information on this subject can be found in Kallianpur [84, Chps. 5, 7], Liptser and Shiryayev [103, Chp. 4, pp. 126-151], Elliott [47] and Karatzas and Shreve [86].

Definition: Strong Solution.

Let $(\Omega, \mathcal{F}, P)$ be a complete probability space with $\mathcal{F}_t$, $t \in [0, T]$ a family of sub-σ-fields of $\mathcal{F}$ satisfying the usual conditions. Further, assume w_t is d-dimensional Wiener process, $w_t \in M(\mathcal{F}_t)$, $\mathcal{F}_t^{\eta,w}$ and η an n-dimensional random variable, $\eta \sim \mathcal{F}_0$. We shall say that an n-dimensional process ξ_t, $t \in [0, T]$ defined on $(\Omega, \mathcal{F}, P)$ is a strong solution of the stochastic differential equation

$$d\xi_t = a(t, \xi_t)\, dt + b(t, \xi)_t\, dw_t \tag{7.C.1}$$

with initial condition $\xi_0 = \eta$ if for each $t \in [0, T]$ the following assertions are satisfied:

(i) ξ_t is continuous and $\mathcal{F}_t$-adapted for each $t \in [0, T]$

(ii) $P\left(\int_0^T |a(t,\xi)|dt < \infty\right) = 1$ a.s.,

(iii) $P\left(\int_0^T \| b(t, \xi) \|^2 < \infty\right) = 1$ a.s. and

(iv) $\xi_t = \eta + \int_0^t a(s, \xi)\, ds + \int_0^t b(s, \xi) dw_s$ w.p.1, $0 \le t \le T$.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Moreover, the stochastic differential equation (7.C.1) has a unique strong solution, if for any two strong solutions ξ_t^1, $t \in [0, T]$, ξ_t^2, $t \in [0, T]$, we have

$$P\{\sup_{0 \leq t \leq T} |\xi_t^1 - \xi_t^2| > 0\} = 0$$

Remark 7.C.1 The definition of the strong solution assumes that a set up $\{(\Omega, \mathcal{F}, P), \mathcal{F}_t, w, \omega, a, b\}$ is given in advance. If in this case $\mathcal{F}_t = \mathcal{F}_t^{w,\eta} \triangleq \sigma\{\eta, w_s;\ 0 \leq s \leq t\}$,

then the process ξ_t, $t \in [0, T]$ is such that with each t the ξ_t is $\mathcal{F}_t^{w,\eta}$-measurable (i.e ξ_t is determined by the past trajectory of the Wiener process).

The simplest conditions guaranteeing the existence and uniqueness of strong solutions of (7.C.1) are given in the following theorem.

Theorem 7.C.1

Let the coefficient of (7.C.1) satisfy the following Lipschitz and growth conditions; there exists a positive constant K such that for all $t \in [0, T]$, ξ, $\hat{\xi} \in R^n$,

$$|a(t,\xi) - a(t,\hat{\xi})|^2 + ||b(t,\xi) - b(t - \hat{\xi})||^2 \leq K|\xi - \hat{\xi}|^2$$
$$|a(t,\xi)|^2 + ||b(t,\xi)||^2 \leq K\ (1 + |\xi|^2).$$

Then (7.C.1) has a unique, strong solution $\xi_t \sim \mathcal{F}_t$ for all $t \in [0, T]$.

Proof: See Liptser and Shiryayev [103, Chp. 4, pp. 129-132].

The Lipschitz condition guarantees uniqueness whereas the linear growth condition guarantees existence (no finite explosion times).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The other type of solution which has proven useful, especially in applications to stochastic control theory, is the so-called weak solution (or solution in a weak sense), that is, any two solutions which have the same finite dimensional distribution. For example, if Liptschitz conditions are not satisfied, then we consider weak solutions; see Karatzas and Shreve [86] for specific examples.

Definition: Weak Solution.

Suppose on some complete probability space $(\Omega, \mathcal{F}, P)$ we can define an increasing family of σ-fields $\mathcal{F}_t \subseteq \mathcal{F}$, an n-dimensional random variable η with prescribed distribution function $P(\eta)$, continuous processes ξ_t, $t \in [0, T]$, w_t, $t \in [0, T]$, such that condition (ii), (iii) (iv) of the previous definition are satified, and

(i) $\quad w_t \in M(\mathcal{F}_t, P)\,, \quad \xi_t \sim \mathcal{F}_t,$

Then, ξ_t, $t \in [0, T]$ is called a weak solution associated with the model $(\Omega, \mathcal{F}, \mathcal{F}_t, P, w_t, \xi_t)$.

Moreover, the stochastic differential equation (7.C.1) is said to have a unique weak solution (solution in a weak sense), if for any two weak solutions ξ^i_t, $t \in [0, T]$ associated with $(\Omega^i, \mathcal{F}^i, \mathcal{F}^i_t, P^i, w^i_t, \xi^i_t)$, $(i=1, 2)$, ξ^1_t, $t \in [0, T]$ and ξ^1_t, $t \in [0, T]$ have the same distributions (i.e coincide), that is, if

$$P^1\{\omega^1;\ \xi^1(\omega^1) \in A\} = P^2\{\omega^2;\ \xi^2(\omega^2) \in A\}$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Remark 7.C.2 The definition of weak solution requires that the functions $a(t, \xi_t)$, $b(t, \xi_t)$ be prescribed, and it is assumed that we can construct a probability space $(\omega, \mathcal{F}, \mathcal{P})$.

Next, we present a theorem that provides existence and uniqueness of weak solutions to stochastic differential equations via the Radon-Nikodym theorem.

Theorem 7.C.2 Let $(C([0, T]; R^n), B_T^n)$ be a measurable space of the continuous functions ξ_t, $t \in [0,T]$, $\xi_0 = 0$, $B_T^n \triangleq \sigma\{\xi_s, 0 \leq s \leq T\}$ and let by $\mathcal{P}_w$ be a Wiener measure on $(C([0, T]; R^n), B_T^n)$. Suppose that $a(t, \xi)$ of (7.C.1) is such that

$$\mathcal{P}_w\{\xi; \int_0^T a^2(t,\xi_t)\, dt < \infty\} = 1 \text{ a.s.}$$

$$E_w\{e^{\int_0^t a(t, \xi_t) d\tilde{w}_t - \frac{1}{2}\int_0^T |a(t, \xi_t)|^2 dt}\} = 1 \text{ a.s.}$$

where E_w denotes expectation with respect to measure $\mathcal{P}_w$. Then (7.C.1) has a unique weak solution, provided $b(t, \xi_t)$ is an n x n constant matrix.

Proof: See Liptser and Shiryayev [103, Thm. 4.11 pp. 147-148] or Kallianpur [84, Chp. 7, Thm. 7.4.1, pp. 180].

Remark 7.C.3 Theorem 7.C.2 has received significant attention in the theory of stochastic optimal control. In particular, suppose the dynamic programming equation is degenerate. Then no conclusion can be made as to whether there exists a twice

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

continuously differential solution so, no existence of optimal control can be concluded using the verification theorem presented by Fleming and Rishel [56]. However, one can show existence of weak (generalized) solutions. That is, if the uncontrolled stochastic differential equation possesses a density in an L^p space, (i.e., unique in probability law) and if the partial derivatives of the solution to the dynamic programming equation are functions in L^p space, then there exists a control function minimizing the Hamiltonian function. Results of these nature were first derived by Beneš [3] and later by Rishel [118], Davis [36], Fleming and Rishel [56], Fleming and Pardoux [53].

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

APPENDIX 7.D
STOCHASTIC DIFFERENTIAL MAPS

In this section we shall introduce certain concepts of stochastic differential geometry and their application to stochastic differential equations which are of primary importance in pursuing the partially observed stochastic control problem of Chapter 4. As a starting point we shall give the representation of stochastic differential equations using the definition of vector fields defined on a manifold M. Although the concepts and notation will appear to be complex, their importance in analyzing SDE's is very rewarding. For more detailed discussion as well as background material on this subject we refer to Kunita [91, 92, 95, 96], Elworthy [51], Blagoveščenskii and Friedlin [22], Ikeda and Watanabe [78], Malliavin [104], Bismut [17, 18, 19].

Suppose we are given a continuous stochastic process x_t, $t\in[0,T]$, $x_t \in SM$ having a differential form

$$d\xi_t = f(t,\xi_t)dt + \sigma(t,\xi_t)\cdot dw_t \tag{7.D.1}$$

where $w_t\in R^m$ is a vector Brownian motion, and $\xi_t\in R^n$ is the solution of (7.D.1) with given initial condition ξ_s=x. Furthermore, assume f, σ_j, j = 1, ..., m are smooth vector fields on a manifold M and the solution ξ_t sometimes denoted by $\xi_{s,t}$ (x, ω), is in M. Since the coefficients of (7.D.1) are vector fields on M assigning to each point $x\in M$ the tangent vectors f (x), σ_j (x)$\in T_xM$, each vector field in T_xM can be represented in terms of the local coordinates $(x^1, ..., x^n)$ in a neighborhood of x. Thus, any vector field X_j can be represented by

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$X_j = \sum_{t=1}^{n} X_j^i \frac{\partial}{\partial x^i}.$$

If we denote the representation of f(x) by X_0, and that of σ_j by X_j, we can write (7.D.1) as

$$d\xi_t = X_0(t,\xi_t)dt + \sum_{j=1}^{m} X_j(t,\xi_t)\cdot dw_t^j. \tag{7.D.2}$$

where $f^i \triangleq X_0^i$, $i = 1, ..., n$ and $\sigma_j^i \triangleq X_j^i$, $i = 1, ..., n$, $j = 1, ..., m$.

Remark 7.D.1 Let $M = R^n$, and let $X_0(t, x), ..., X_m(t, x)$ be continuous in (t, x), continuously differentiable in t, and twice continuously differentiable in x with first and second derivatives in x bounded. Then the existence and uniqueness of (7.D.2) follows from the existence and uniqueness of the Ito differential equation

$$d\xi_t = \tilde{X}_0(t,\xi_t)dt + \sum_{j=1}^{m} X_j(t,\xi_t)dw_t^j, \tag{7.D.3}$$

where

$$\tilde{X}_0(t,x) = X_0(t,x) + \frac{1}{2}\sum_{j=1}^{m}\sum_{k=1}^{d} X_j^k(t,x)\frac{\partial X_j}{\partial x^k}(t,x)$$

because the coefficient of (7.D.3) are Lipschitz continuous. Indeed, the F-S integral and Ito integral are related by

$$\int_s^t X_j^i(r,\xi_r)\cdot dw_r^j = \int_s^t X_j^i(r,\xi_r)dw_r^j + \frac{1}{2}\langle X_j^i(t,\xi_t) - X_j^i(s,\xi_s),\ w_t^j - w_s^j\rangle. \tag{7.D.4}$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Theorem 7.D.1

Let $M = R^n$. Suppose $\xi_{s,t}(x) \sim \mathcal{F}_s^t$, $t \in [s,T]$ is the solution of (7.D.2) satisfying the conditions of Remark 7.D.1 and let f: $R^n \to R^1$ is a function of C^3-class. Then

$$\text{(i)} \quad f(\xi_{s,t}(x)) = f(x) + \int_s^t X_0(r) f(\xi_{s,r}(x)) dr + \sum_{j=1}^m \int_s^t X_j(r) f(\xi_{s,r}(x)) \circ dw_r^j. \quad (7.D.5)$$

Assume further that the coefficients $X_0(t, x), \ldots, X_m(t, x)$ are of C^4-class in x. Then the solution $\xi_{s,t}(x)$ is a backward semimartingale adapted to $\mathcal{F}_s^t$, $s \in [0,t]$, for t fixed and

$$\begin{aligned}\text{(ii)} \quad f(\xi_{s,t}(x)) &= f(x) + \int_s^t X_0(r)(f \cdot \xi_{r,t})(r,x) dr \\ &+ \sum_{j=1}^m \int_s^t X_j(r)(f \cdot \xi_{r,t})(r,x) \circ \hat{d}w_r^j.\end{aligned} \quad (7.D.6)$$

where "$\cdot$" denotes composition.

Proof: See Kunita [95, 96].

Remark 7.D.2 A similar theorem holds for the Ito stochastic differential equation (7.D.3) which is also found in Kunita [96]. The backward SDE (7.D.6) is used by the above author to derive the backward Kolmogorov equation. Results of this nature are also given by Kunita [95, 96].

Theorem 7.D.2

Suppose the solution map $\xi_{s,t}(x)$ of (7.D.2) is defined on an M-paracompact connected C^∞-manifold and the vector fields $X_0, \ldots, X_m$ are of class C^k, $k \geq 5$. Further,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

let $\xi_{s,t}(x)$ be an $\mathscr{F}_s^t$-adapted process with lifetime τ (i.e., $s \leq t < \tau$) and $\xi_{s,t}(\cdot,\omega): D_s^t(\omega) \to M$, $D_s^t(\omega) \triangleq \{x;\ \tau(s,x,\omega) > t\}$. Then the map $\xi_{s,t}(\cdot,\omega)$ is a local C^{k-2} diffeomorphism map from $D_s^t(\omega)$ into M for any t a.s.. If M is a compact manifold, then $\xi_{s,t}(\cdot,\omega)$ is a C^{k-2}-diffeomorphism of M for all t a.s. (i.e., $D_s^t(\omega) = M$ and $\xi_t(\cdot,\omega)$ becomes an onto map or $s < t$). Furthermore, if $\xi_{s,t*}$ denotes the differential map of $\xi_{s,t}$ for any $f \in C^\infty(M)$, then $f(\xi_{s,t}(x))$ is a backward semimartingale with respect to s and satisfies the backward SDE

$$\text{(i)}\quad f(\xi_{s,t}(x)) = f(x) + \int_s^t \xi_{u,t*}(X_0)f(\xi_{u,t}(x))du + \sum_{j=1}^m \int_s^t \xi_{u,t*}(X_j)f(\xi_{u,t}(x)) \cdot \hat{d}w_u^j, \tag{7.D.7}$$

the solution map $\xi_{s,t}(x,\omega)$ of (7.D.2) satisfies the backward SDE

$$\text{(ii)}\quad \hat{d}\xi_{s,t} = -\xi_{s,t*}(X_0)(\xi_{s,t})ds - \sum_{j=1}^m \xi_{s,t*}(X_j)(\xi_{s,t}) \cdot \hat{d}w_s^j. \tag{7.D.8}$$

Moreover, the inverse map $\xi_{s,t}^{-1}(x,\omega)$ defined by $\eta_{s,t} \triangleq \xi_{s,t}^{-1}$ satisfies the backward SDE

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$\hat{d}\eta_{s,t}(x) = X_0(\eta_{s,t}(x))ds + \sum_{j=1}^{m} X_j(\eta_{s,t}(x)) \cdot \hat{d}w_s^j \tag{7.D.9}$$

with $\eta_{t,t}(x) = x$ which is again an onto map if the set $D_s^t(w)$ of (7.D.9) is equal to M a.s for any $s < t$.

Proof: The proof of this theorem requires results from Kunita [92, 95], Malliavin [104], Ikeda and Watanabe [78] and Bismut [17].

The following definition is from Kunita [95].

Definition: Differential Map - Stochastic Field.

Given a point x of $D_s^t(\omega)$, the differential map $(\xi_{s,t*})_x$ of the map $\xi_{s,t}$ is defined as a linear map from T_xM into $T_{\xi_{s,t}(x)}M$ such that

$$(\xi_{s,t*})_x X_x f \triangleq X_x(f \cdot \xi_{s,t}) \text{ , for all } X_x \in T_xM. \tag{7.D.10}$$

Given a vector field X on M, if we denote by $R_s^t(\omega)$ the range of the map $\xi_t(\cdot, \omega)$ and by X_x the restriction of X at $x \in M$, then the new stochastic vector field $(\xi_{s,t*})(x)$ with domain $R_s^t(\omega)$ is defined for $x \in R_s^t$ by

$$\xi_{s,t*}(X)_x = (\xi_{s,t*})_{\xi_{s,t}^{-1}(x)} X_{\xi_{s,t}^{-1}(x)}) \tag{7.D.11}$$

Finally, it can be shown using (7.D.10), (7.D.11) that

$$\xi_{s,t*}(X)f(x) = X(f \cdot \xi_{s,t}) (\xi_{s,t}^{-1}(x)). \tag{7.D.12}$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Remark 7.D.3 If we apply identity (7.D.12) to the representation (7.D.7) we deduce

$$\xi_{u,t*}(X_0)f(\xi_{u,t}(x)) = X_0(f\cdot\xi_{u,t})(\xi_{u,t}\xi_{u,t}^{-1}(x))$$

and thus (7.D.7) is equivalent to (7.D.6) as expected.

Remark 7.D.3 If the SDE (7.D.2) is defined on R^n with coefficient $X_0, \ldots, X_m$ of class C^k having bounded first and second derivatives, then the solution map $\xi_t(\cdot, \omega)$ is a diffeomorphism of C^{k-1}-class for any t a.s. The proof was first given by Blagoveščenskii and Friedlin [22].

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

APPENDIX 7.E
ABSOLUTE CONTINUITY OF MEASURES AND RELATED TOPICS

In this section we discuss some results on absolute continuity of measures and its widely used applications. The Radon-Nikodym derivative in the detection set-up is referred to as the likelihood-ratio. This transformation of probability measure involves martingales in a manner first introduced by Cameron and Martin [31] in the context of Wiener integrals and later for Wiener processes by Girsanov [67]. Extensive work on this topic is found in Liptser and Shiryayev [103, Chps. 4-7] and Kallianpur [84, Chp. 7]. The importance of this concept to detection problems is established by Duncan [45] and Kailath [80, 81, 82]. The significance of measure transformations to the field of stochastic control was first noted by Beneš [3] and then by Rishel [118] and Davis [36].

Theorem 7.E.1 The Exponential Formula.

Assume $\{x_t, 0 \le t \le T\}$ is a real-valued local martingale, with having $x_0 = 0$ a.s. that is adapted to the family of σ-fields $\mathcal{F}_t$, $t \in [0, T]$ satisfying the usual conditions. Then there exists a unique local martingale $L_t \sim \mathcal{F}_t$ with values in R^1 satisfying the stochastic differential equation

$$dL_s = L_s dx_s, \tag{7.E.1}$$

Moreover, $L_t = L_0 \exp^{x_t - \frac{1}{2}\langle x, x\rangle_t}$, denoted by $\mathcal{E}(x_t)$ and $\mathcal{P}(L_t > 0, t \in [0,T]) = 1$ a.s..

Proof: See van Schuppen [127].

We now present the result on absolute continuity of measures and its relation to martingales.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Definition Absolute Continuity of Measures

Suppose we are given a measurable space $(\Omega, \mathscr{F})$ and two probability measures $\tilde{P}$, P defined on it. We say that measure $\tilde{P}$ is absolutely continuous with respect to P (i.e., $\tilde{P} << P$), if for all $A \in \mathscr{F}$ such that $P(A) = 0$ we have $\tilde{P}(A) = 0$.

A consequence of absolute continuity of measures is the existence of an integrable function $L(\omega)$ such that $\tilde{P}(A) = \int_A L(\omega)\ P(d\omega)$, $\omega \in \Omega$ where $L(\omega)$ is the Radon-Nikodym derivative sometimes denoted as $L = \frac{d\tilde{P}}{dP}$.

Theorem 7.E.2 Measure Transformation (Girsanov's Theorem).

1. Suppose we are given a probability space $(\Omega, \mathscr{F}, P)$, a family of sub-σ-fields $\mathscr{F}_t \subset \mathscr{F}$, $t \in [0, T]$ and a local martingale $m_t \in (\mathscr{F}_t, P)$ such that $m_0 = 0$, $\langle m, m\rangle_T < \infty$ P-a.s, $t \in [0,T]$ satisfying $E[\mathscr{E}(x_T)] = 1$. Then, the formula $\mathscr{E}(x_T) = E\,[\frac{d\tilde{P}}{dP} \mid \mathscr{F}_T]$ introduces a new probability measure $\tilde{P}$ on $(\Omega, \mathscr{F}_t)$ such that $\tilde{P} << P$. If we also have $\tilde{P}\{\langle m,m\rangle_t < \infty\} = 1$ then $P << \tilde{P}$, thus, $P \sim \tilde{P}$.

2. Suppose we are given a measurable space $(\Omega, \mathscr{F})$ and two probability measures $\tilde{P}$, P defined on it. If we assume $\tilde{P}$ and P equivalent measures, then $\frac{d\tilde{P}}{dP} > 0$ P-a.s.

Let $\mathscr{F}_t \subset \mathscr{F}$, $t \in [0,T]$ be a family of sub-σ-fields satisfying the usual conditions and define

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$L_T \triangleq E\,[\frac{d\tilde{P}}{dP} \mid \mathscr{F}_T]$. Then, there exists a process $\tilde{x}_t \in M_{loc}(\mathscr{F}_t, \tilde{P})$, $\tilde{x}_0 = 0$, $t \in [0, T]$

such that $L_T = \mathscr{E}(\tilde{x}_T)\, P$- a.s, $t \in [0,T]$.

Proof: See van Schuppen [127, pp. 38].

We now present the theorem which is a consequence of transformation of measures. As we shall see, each time Theorem 7.E.2 is applied, certain martingales need to be defined with respect to the new induced measure.

Theorem 7.E.3 Martingale Translation.

Suppose

1. $(\Omega, \mathscr{F}, P)$ is a probability space, $\mathscr{F}_t \subset \mathscr{F}$, $t \in [0, T]$;
2. $y_t \in M_{loc}(\mathscr{F}_t, P)$, $y_t \in R^n$, $y_0 = 0$ a.s;
3. $x_t \in M_{loc}(\mathscr{F}_t, P)$, $x_t \in R$ such that $x_0 = 0$, P- a.s, $E(L_T) = 1$, where

$$L_T \triangleq E\,[\frac{d\tilde{P}}{dP} \mid \mathscr{F}_T] = e^{\,x_T - \frac{1}{2}\langle x, x\rangle_T} \tag{7.E.2}$$

Then, $\tilde{y}_t \in M_{loc}(\mathscr{F}_t, \tilde{P})$, $\tilde{y}_t \in R^n$ is such that

$$\tilde{y}_t = y_t - v_t,\; v_t^i = \langle y^i, x\rangle_t\,,\; i = 1, \ldots, n \tag{7.E.3}$$

and

$$\langle \tilde{y}, \tilde{y}\rangle_t = \langle y, y\rangle_t.$$

Proof: See Kallianpur [84, Theorem 7.1.2, pp. 168].

Remark 7.E.1 Suppose $(\Omega, \mathscr{F}_t)$ is a measurable space with $\tilde{P}$, P defined on it such that $\tilde{P} \sim P$. If the likelihood-ratio L_T is defined by

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$L_T \triangleq E\,[\frac{d\tilde{P}}{dP} \mid \mathscr{F}_T]$$

then

$$L_T^{-1} = \tilde{E}\,[\frac{dP}{d\tilde{P}} \mid \mathscr{F}_T] \tag{7.E.4}$$

where $\tilde{E}$ denotes expectation with respect to measure $\tilde{P}$. Suppose $f(\omega)$, $\omega \in \Omega$ is an integrable function adapted to $\mathscr{F}_t$ defined under measure P. Then we have

$$E[f(\omega)] = \tilde{E}[L_T f(\omega)] \tag{7.E.5}$$

and

$$E\,[f(\omega) \mid \mathscr{F}_s] = \frac{\tilde{E}[f(\omega) L_T \mid \mathscr{F}_s]}{\tilde{E}[L_T \mid \mathscr{F}_s]}\text{, for any } s \le t \in [0,T]. \tag{7.E.6}$$

<u>Remark 7.E.2</u> Suppose the Radon-Nikodyn derivative L_T defined by (7.E.2) is $\int f(t,\omega)dw$ such that $x_T = \int_0^T f(t,\omega)dw_t$. If for some $\delta > 0$ the condition

$$\sup_{t \in [0,T]} E \exp(\delta |f(t,\omega)|^2) < \infty \quad P\text{-a.s.} \tag{7.E.7}$$

is satisfied then $E[\,L_t\,]=1$, $t \in [0,T]$. Condition (7.E.7) is very important for our later work. It can be shown that if $|a(t, \xi)| \le k(1 + |\xi|)$, $|b(t, \xi)|\; k \le \infty$ of equation (7.C.1) are satisfied then (7.E.7) is satisfied. This result is found in Liptser and Shiryayev [103, Chp. 6, pp. 220-221, Chp. 4, Thm. 4.7, pp. 137-141], and in Gihman and Skorohod [65].

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Remark 7.E.3 Assume that the stochastic differential equation (7.C.1) is a diffusion process of dimension one and the solution $\xi_{0,t}$ is an $\mathcal{F}_t^{\xi}$-adapted process for all $t \in [0,T]$, with $\xi_0 = 0$. Then

$$1. \quad P_{\xi}\left(\int_0^T a^2(t, \xi_t)dt < \infty\right) = 1 \text{ a.s.} \tag{7.E.8}$$

$$2. \quad P_w\left(\int_0^T a^2(t, w_t)dt < \infty\right) = 1 \text{ a.s.} \tag{7.E.9}$$

imply that $P_{\xi} \sim P_w$. The likelihood-ratios are given by

$$\frac{dP_w}{dP_{\xi}} = \exp\left\{-\int_0^t a(s,\xi_s)\, d\xi_s + \frac{1}{2}\int_0^t a^2(s,\xi_s)\, ds\right\} \tag{7.E.10}$$

$$\frac{dP_{\xi}}{dP_w} = \exp\left\{\int_0^t a(s,w_s)dw_s - \frac{1}{2}\int_0^t a^2(s,w_s)ds\right\}. \tag{7.E.11}$$

If we remove condition (7.E.8) we have $P_w << P_{\xi}$, similarly, if we remove condition (7.E.9) we only have $P_{\xi} << P_w$. Here, P_w denotes the Wiener measure $P_w(B) = P_w(\omega;\ w \in B)$, P_{ξ} denotes the measure $P_{\xi}(B) = P_{\xi}(\omega; \xi \in B)$, and by the definition of diffusion process both are defined on $\mathcal{F}_t^{\xi} = \sigma\{\xi_s;\ 0 \le s \le t\}$. The results above are found in Liptser and Shiryayev [103, Chp. 7, Thm. 7.7, pp. 248]. The solution of (7.C.1) is described by its probability law P_{ξ} (i.e., weak solution) which is a consequence of Theorem 7.C.2.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Furthermore, suppose we assume condition (7.E.9) with ρ_ξ the weak solution to (7.C.1). It then follows that $E_w\,[\frac{d\rho}{d\rho_w}\,|\,\mathscr{F}_t^\xi] = 1$.

The following theorem describes independecne of σ-fields.

Theorem 7.E.4

We say that two σ-fileds $\mathscr{F}_t^1, \mathscr{F}_t^2$ are conditionally independent given $\mathscr{F}_t$ if, and only if, for any $B^2 \in \mathscr{F}_t^2$

$$\rho(B^2\,|\,\mathscr{F}_t, \mathscr{F}_t^1) = \rho(B^2\,|\,\mathscr{F}_t) \text{ a.s.}$$

with the superscripts 1 and 2 interchanged.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

APPENDIX 7.F
MARKOV PROCESSES, BACKWARD AND FORWARD EQUATIONS

Markov processes play an essential role in the theory of random processes. The most important feature of Markov process is the evolutionary character of its behavior: the state of the process at present completely determines its probabilistic behavior in the future. This feature allows us to derive evolutionary equations for the determination of the probabilistic characteristics of the process. Our main references for the material presented in this section are found in Gihman and Skorohod [66, Chp. 1, pp. 8 -16, Vol. 2], Dynkin [46] and Fleming and Rishel [56].

We begin with an introduction on transition functions of markov processes. We then present two families of operators which can be associated with transition probabilities and finally we give the backward and forward equations which are related to the above transition probabilities. This result are essential in understanding the proves given in subsequent development.

Transition Functions of Markov Processes

Suppose we are given a stochastic process $\{x_t, t \in [0,T]\}$, with state space some complete separable metric space Σ (the state space of the process x_t). Denote by $P(\beta,t; x,s)$ the probability of the event $x_t \in \beta$ given that $x_s = x$, $s < t$ where $\beta \subset B(\Sigma)$, ($B(\Sigma)$ denotes the Borel set of Σ). The function

$$P(\beta,t;x,s) \triangleq \mathrm{Prob}(x_t \in \beta \,|\, x_s = x) \tag{7.F.1}$$

is called the transition probability of x_t. If x_t is a Markov process (without an after effect), from the properties of conditional expectations we have

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$P(\beta,t;x,s) = \int_{\Sigma} P(\beta,t;y,u)\ P(dy,u;x,s),\ s \le u \le t \qquad (7.F.2)$$

which is called the Chapman-Kolmogorov equation.

Definition: Markov Process

A stochastic process $\{x_t, t \in [0,T]\}$, with state space Σ, is a Markov process if:

1. $P(x_t \in \beta \mid x_{t_1}, \ldots, x_{t_m}) = P(x_t \in \beta \mid x_{t_m})$ where $t_1 < t_2 < \ldots < t_m < t$ is in $[0,T]$;

2. $P(\beta,t; .,s)$ is $B(\Sigma)$ measurable for fixed s, t, β and $P(., t; x,s)$ is a probability measure on $B(\Sigma)$ for fixed s, x, t;

3. The Chapman-Kolmogorov equation holds for $s < u < t$ in $[0, T]$.

The Backward Equation

Suppose $\bar{B}(\Sigma)$ denotes the space of all bounded, real valued, Borel measurable functions f on Σ, with the norm $||f|| = \sup_{x \in \Sigma} |f(x)|$. For every $f \in \bar{B}(\Sigma)$ and $s,t \in [0,T]$ with $s < t$, let

$$T_{s,t}f(x) \triangleq \int_{\Sigma} f(y)P(dy,t;x,s) = E[f(x_t) \mid x_s = x]. \qquad (7.F.3)$$

The family of operators $T_{s,t}$ determined by the above relation is called the semigroup of operators associated with the transition probability $P(\beta,t;x,s)$. By the Chapman-Kolmogorov equation it follows easily that $T_{s,t}$ is a semigroup, i.e.,

$$T_{s,t} = T_{s,u}T_{u,t} \qquad s < u < t. \qquad (7.F.4)$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Let us consider some fixed time $t_1 \in [0,t^*]$ and define the infinitesimal operator L(s) on some subspace D of $\bar{B}(\Sigma)$ such that for each $s \in (0,t_1)$ the limit

$$L(s)f(x) = \lim_{h \downarrow 0} \frac{T_{s-h,s}f(x) - f(x)}{h}, \quad s \in (0,t_1) \tag{7.F.5}$$

exists and

$$\lim_{h \downarrow 0} T_{t_1-h,t_1}f(x) = f(x).$$

The limit is understood in a weak or strong sense. For more discussion on the limit see Dynkin [46, p. 19 - 22, Vol. 1]. The operator L(s) is well-defined only under suitable restrictions on the function f for which the above limit exists. These functions are said to be in the domain D.

The semigroup $T_{s,t}$ is uniquely determined by the infinitesimal operator L(s) above. Moreover, for a diffusion process, L(s) takes the form of a second order partial differential operator in the sense that the domain D of L(s) contains all f of class C^2 and

$$L(s)f \triangleq \frac{1}{2} \sum_{i,j=1}^{n} a^{ij}(s,x) \frac{\partial^2 f}{\partial x^i \partial x^j} + \sum_{i=1}^{n} b^i(s,x) \frac{\partial f}{\partial x^i} \tag{7.F.6}$$

where b, σ are drift and diffusion coefficients of the diffusion equation satisfied by the process x_t with $a = \sigma\, \sigma^T$. Finally, by combining (7.F.3), (7.F.5) it can be shown that $T_{s,t_1}f(x)$ satisfies

$$\begin{aligned} &\frac{\partial}{\partial s} T_{s,t_1}f(x) = -L(s)T_{s,t_1}f(x), \quad s \in [0,t_1] \\ &\lim_{s \uparrow t_1} T_{s,t_1}f(x) = f(x). \end{aligned} \tag{7.F.7}$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The Kolmogorov's backward equation is a special case of (7.F.7) and is obtained by letting $f(y) \rightarrow I_{\{\omega;\ y \in \beta\}}$ in (7.F.6) so that $T_{s,t_1} f(x) \rightarrow P(\beta,t;x,s)$ in (7.F.7).

The Forward Equation

Suppose we now define on $B(\Sigma)$ a probability measure $\rho_s(\beta) = \text{Prob}\ \{x_s \in \beta\}$. Then it follows from the general formulas of probability theory that the probability $\rho_s(\beta)$ of the event $\{x_t \in \beta\}$ should be defined as

$$\rho_t(\beta) \triangleq \text{Prob}\{x_t \in \beta\} \triangleq \int P(\beta,t;x,s)\rho_s(dx)\,. \tag{7.F.8}$$

If we denote by $\tilde{M}$ the set of all finite measures on $B(\Sigma)$ and define

$$m_{t,s} \triangleq T^*_{t,s}\, m \tag{7.F.10}$$

where

$$m_{t,s}(\beta) = \int P(\beta,t;y,s)m(dy) \quad , s \leq t, \quad \beta \in B(\Sigma) \tag{7.F.10}$$

then $T^*_{t,s}$ is an operator which maps $\tilde{M}$ into $\tilde{M}$. Again, using the Chapman-Kolmogorov formula we can write

$$T^*_{t,s} = T^*_{t,u} T^*_{u,s} \quad , \quad s < u < t. \tag{7.F.11}$$

The operator $T^*_{t,s}$ is now directed differently than $T_{s,t}$. Using a similar approach as the one presented for the backward equation one can show for s a fixed time $s,t \in [0,T]$ that $T^*_{t,s} m(B)$ satisfies

$$\begin{aligned} &\frac{\partial}{\partial t} T_{t,s} m(\beta) = L(t)^*\, T_{t,s} m(\beta) \\ &\lim_{t \downarrow s} T_{s,t} m(\beta) = m(\beta)\,. \end{aligned} \tag{7.F.12}$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The Kolmogorov's forward equation is a special case of (7.F.12) by letting $m(\beta) \rightarrow I_{\{\omega;\, x\in\beta\}}$ in (7.F.10) so that $T^*_{t,s}m(\beta) \rightarrow P(\beta,t;x,s)$ in (7.F.12).

Remark 7.F.1 The backward and forward equations are adjoint to each other. The time derivative of (7.F.7), (7.F.12) defined in an L^2 norm is zero.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

APPENDIX 7.G
DERIVATION OF SUFFICIENT STATISTICS OF THEOREM 3.4.1

Reffering to (3.4.15) and Theorem 3.4.1 we seek for the solution of

$$\frac{\partial}{\partial t}q(z,t) = L(t)^{*}q(z,t) - Q(t,z)q(z,t) \tag{7.G.1}$$

with initial condition $q(z,0) = p_0(z)$. To solve, set

$$q(z,t) = k_t e^{-\frac{1}{2}(z-\mu_t)^T\Sigma_t^{-1}(z-\mu_t)} \tag{7.G.2}$$

where z_t, μ_t are n-dimensional vectors, and Σ_t is an nxn symmetric matrix.

Differentiating (7.G.2) with respect to time:

$$\frac{\partial}{\partial t}q(z,t) = \dot{k}_t q(z,t) + [(z-\mu_t)^T\Sigma_t^{-1}\dot{\mu}_t + \frac{1}{2}(z-\mu_t)^T\Sigma_t^{-1}\dot{\Sigma}_t\Sigma_t^{-1}(z-\mu_t)]q(z,t) \tag{7.G.3}$$

Differentiating (7.G.2) with respect to z:

$$\frac{\partial}{\partial z}q(z,t) = -\Sigma_t^{-1}(z-\mu_t)q(z,t). \tag{7.G.4}$$

Differentiating (7.G.4) with respect to z:

$$\frac{\partial^2}{\partial z^2}q(z,t) = -\Sigma_t^{-1}q(z,t) + \Sigma_t^{-1}(z-\mu_t)(z-\mu_t)^T\Sigma_t^{-1}q(z,t) \tag{7.G.5}$$

Since, L(t) is the forward operator associated with the diffusion process x_t of (3.4.9), and Q is given by (3.4.16) then

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$\frac{\partial}{\partial t}q(z,t) = \frac{1}{2}B^T\frac{\partial^2}{\partial z^2}q(z,t)B - q(z,t)Tr(A) - Az\frac{\partial}{\partial z}q(z,t)$$
$$- \frac{1}{2}(z^T\tilde{A}z + \tilde{B}^Tz + \tilde{C}) + \frac{1}{2}c_n + z^Tc^Tcz - y_tz_2 - \frac{1}{2}z_1^2 \tag{7.G.6}$$

where, A,B, are the drift and diffusion coefficients of (3.4.9) respectively, and c the last row of drift coefficient A. Therefore, by substituting (7.G.3) - (7.G.5) into (7.G.1),

$$\dot{k}_t + (z-\mu_t)^T\Sigma_t^{-1}\dot{\mu}_t + \frac{1}{2}(z-\mu_t)^T\Sigma_t^{-1}\dot{\Sigma}_t\Sigma_t^{-1}(z-\mu_t) =$$
$$\frac{1}{2}\{B^T[-\Sigma_t^{-1} + \Sigma_t^{-1}(z-\mu_t)(z-\mu_t)^T\Sigma_t^{-1}]B\} - Tr(A) + (z-\mu_t)^T\Sigma_t^{-1}Az \tag{7.G.7}$$
$$- \frac{1}{2}(z^T\tilde{A}z + \tilde{B}^Tz + \tilde{C}) + \frac{1}{2}c_n + \frac{1}{2}z^Tc^Tcz - y_tz_2 - \frac{1}{2}z_1^2.$$

and equating coefficients of $z^T(\cdot)z$, , $z^T(\cdot)$, and $z^0(\cdot)$ we can obtain an equation for $\dot{\Sigma}_t$, $\dot{\mu}_t$, $\dot{k}_t$, respectively.

$$\frac{1}{2}\Sigma_t^{-1}\dot{\Sigma}_t\Sigma_t^{-1} = \frac{1}{2}\Sigma_t^{-1}BB^T\Sigma_t^{-1} + \Sigma_t^{-1}A - \frac{1}{2}\tilde{A}'$$

where

$$\tilde{A}' = \tilde{A} + \begin{array}{cc} & \scriptstyle n-1 \\ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} & \scriptstyle n-1 \end{array} - \frac{1}{2}c^Tc.$$

Thus,

$$\dot{\Sigma}_t = A\Sigma_t + \Sigma_tA^T - \Sigma_t\tilde{A}'\Sigma_t + BB^T$$
$$\Sigma_0 = 0. \tag{7.G.8}$$

Next,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$\Sigma_t^{-1}\dot{\mu}_t - \frac{1}{2}\Sigma_t^{-1}\dot{\Sigma}_t\Sigma_t^{-1}\mu_t - \frac{1}{2}\Sigma_t^{-1}\dot{\Sigma}_t\Sigma_t^{-1}\mu_t =$$
$$-\frac{1}{2}\Sigma_t^{-1}BB^T\Sigma_t^{-1}\mu_t - \frac{1}{2}\Sigma_t^{-1}BB^T\Sigma_t^{-1}\mu_t - A^T\Sigma_t^{-1}\mu_t - \frac{1}{2}\tilde{B}'$$

where

$$\tilde{B}' = \tilde{B} + H^T, \quad \tilde{H}^T = [0\ 2\ 0\ \ldots\ 0]^T y_t.$$

Thus,

$$\dot{\mu}_t = \dot{\Sigma}_t\Sigma_t^{-1}\mu_t - BB^T\Sigma_t^{-1}\mu_t - \Sigma_t A^T\Sigma_t^{-1}\mu_t - \frac{\Sigma_t}{2}\tilde{B}'.$$

Substituting (7.G.8) into the last equation, we have,

$$\begin{aligned} \dot{\mu}_t &= (A - \Sigma_t\tilde{A}')\mu_t - \frac{1}{2}\Sigma_t\tilde{B} - \frac{1}{2}\Sigma_t H^T y_t \\ \mu_0 &= z(t=0). \end{aligned} \qquad (7.G.9)$$

where $H^T \triangleq \dfrac{\tilde{H}^T}{y_t}$

Finally, an equation for k_t can be obtained as above; however, since k_t is cancelled during the normalization of density $\rho(z,t)$, there is no need to simplify (7.G.7) further.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

APPENDIX 7.H
DERIVATION OF SUFFICIENT STATISTICS OF EXAMPLE 5.4.3

Reffering to Example 5.4.3, we shall show that the solution to stochastic PDE (5.4.22) is given by a set of sufficient statistics Σ_t^i, μ_t^i, k_t^i satisfying

$$\dot{\Sigma}_t^i = \tilde{F}\Sigma_t^i + \Sigma_t^i\tilde{F}^T - \Sigma_t^i(\tilde{c}^T\tilde{c} + \tilde{A})\Sigma_t^i + \tilde{E}\tilde{E}^T \quad ,\Sigma_t^0 = 0 \tag{7.H.1}$$

$$\dot{\mu}_t^i = \tilde{F}\mu_t^i - \Sigma_t^i(\tilde{c}^T\tilde{c} + \tilde{A})\mu_t^i + \Sigma_t\tilde{B} + \Sigma_t\tilde{c}^T \bullet \frac{dy}{dt} \quad ,\mu_{t_0} = \tilde{x}_0 \tag{7.H.2}$$

$$\dot{k}_t^i = -\frac{1}{2}\mu_t^{i\,T}\Sigma_t^{i\,-1}\dot{\Sigma}_t^i\Sigma_t^{i\,-1}\mu_t^i + \mu_t^{i\,T}\Sigma_t^{i\,-1}\dot{\mu}_t^i + \frac{1}{2}\mu_t^{i\,T}\Sigma_t^{i\,-1}\tilde{E}\tilde{E}^T\Sigma_t^{i\,-1}\mu_t - \frac{Tr}{2}(\tilde{E}\Sigma_t^{i\,-1}\tilde{E}) + \tilde{D} \;,k_{t_0} = 1. \tag{7.H.3}$$

Define a new density $\hat{\rho}$ related to ρ though a gauge transformation

$$\hat{\rho}^i(\tilde{x},t) = e^{-F(x_1)}\rho^i(\tilde{x},t).$$

From (5.4.22) we have

$$\frac{\partial}{\partial t}\hat{\rho}^i(\tilde{x},t) = \frac{\partial}{\partial t}(e^{-F(x_1)}\rho^i(\tilde{x},t)) = e^{-F(x_1)}\frac{\partial}{\partial t}\rho^i(\tilde{x},t)$$

$$= e^{-F(x_1)}\{(L_{\tilde{x}}(t)^* e^{F(x_1)}\hat{\rho}^i(\tilde{x},t)) + \tilde{x}^T\tilde{c}^T e^{F(x_1)}\hat{\rho}^i(\tilde{x},t)\bullet\frac{dy_t}{dt}\}$$

$$= e^{-F(x_1)}\{\frac{Tr}{2}\frac{\partial}{\partial\tilde{x}}(\tilde{\sigma}\tilde{\sigma}^T\frac{\partial}{\partial\tilde{x}}F(x_1)e^{F(x_1)}\hat{\rho}^i(\tilde{x},t) + \tilde{\sigma}\tilde{\sigma}^T e^{F(x_1)}\frac{\partial}{\partial\tilde{x}}\hat{\rho}^i(\tilde{x},t))$$

$$- Tr(\frac{\partial}{\partial\tilde{x}}\tilde{f})e^{F(x_1)}\hat{\rho}^i(\tilde{x},t) - \tilde{f}^T\frac{\partial}{\partial\tilde{x}}F(x_1)e^{F(x_1)}\hat{\rho}^i(\tilde{x},t)$$

$$- e^{F(x_1)}\tilde{f}^T(\tilde{x})\frac{\partial}{\partial\tilde{x}}\hat{\rho}(\tilde{x},t) - \frac{1}{2}\tilde{x}^T\tilde{c}^T\tilde{c}\tilde{x}e^{F(x_1)}\hat{\rho}^i(\tilde{x},t) + \tilde{x}^T\tilde{c}^T e^{F(x_1)}\hat{\rho}^i(\tilde{x},t)\bullet\frac{dy_t}{dt}\}$$

After some algebra we obtain

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$\frac{\partial}{\partial t}\rho(\tilde{x}, t) = \frac{1}{2}\frac{\partial^2}{\partial x_1^2}\rho^i(\tilde{x}, t) + (-\frac{1}{2}\frac{\partial}{\partial x_1}f^1(x_1) - \frac{1}{2}(f^1(x_1))^2)\rho^i(\tilde{x}, t)$$

$$-f^0(x_0)\frac{\partial}{\partial x_0}\rho^i(\tilde{x}, t) - \frac{\partial}{\partial x_0}f^0(x_0)\rho^i(\tilde{x}, t) - \frac{1}{2}\tilde{x}^T\tilde{c}^T\tilde{c}\tilde{x}\rho^i(\tilde{x}, t) + \tilde{x}^T\tilde{c}^T\rho^i(\tilde{x}, t)\cdot\frac{dy_t}{dt}.$$

From (5.4.21) and the definition of f^0, $\tilde{h}$ we deduce

$$\begin{aligned}\frac{\partial}{\partial t}\rho(\tilde{x},t) &= \frac{1}{2}\frac{\partial^2}{\partial x_1^2}\rho^i(\tilde{x}, t) - \frac{1}{2}(\gamma x_1^2 + \beta x_1 + \delta)\rho^i(\tilde{x}, t)\\ &- B^0x_0\frac{\partial}{\partial x_0}\rho^i(\tilde{x}, t) - B^0\rho^i(\tilde{x}, t) - \frac{1}{2}\tilde{x}^T\tilde{c}^T\tilde{c}\tilde{x}\rho^i(\tilde{x}, t)\\ &+ \tilde{x}^T\tilde{c}^T\rho^i(\tilde{x}, t)\cdot\frac{dy_t}{dt}.\end{aligned} \tag{7.H.4}$$

We now define

$$\tilde{A} \triangleq \begin{bmatrix}\gamma & 0\\ 0 & 0\end{bmatrix}, \quad \tilde{B} = \begin{bmatrix}-\frac{\beta}{2}\\ 0\end{bmatrix}, \quad \tilde{D} = -\frac{\delta}{2} - B^0$$

$$\tilde{E} = \begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}, \quad \tilde{F} = \begin{bmatrix}0 & 0\\ 0 & B^0\end{bmatrix}$$

and rewrite (7.H.4) as

$$\begin{aligned}\frac{\partial}{\partial t}\rho(\tilde{x}, t) &= \frac{Tr}{2}(\tilde{E}\frac{\partial^2}{\partial \tilde{x}}\rho^i(\tilde{x}, t)\tilde{E}) - \frac{\partial}{\partial \tilde{x}}\rho^{i\,T}(\tilde{x}, t)\tilde{F}\tilde{x} + \tilde{x}^T(-\frac{\tilde{c}^T\tilde{c}}{2} - \frac{\tilde{A}}{2})\tilde{x}\rho^i(\tilde{x}, t)\\ &+ \tilde{x}^T\tilde{B}\rho^i(\tilde{x}, t) + \tilde{x}^T\tilde{c}^T\rho^i(\tilde{x}, t)\cdot\frac{dy}{dt}.\end{aligned} \tag{7.H.5}$$

To solve, we set

$$\rho^i(\tilde{x}, t) = k_t^i e^{-\frac{1}{2}(\tilde{x} - \mu_t^i)^T\sigma_t^{i\,-1}(\tilde{x} - \mu_t^i)}, \quad \rho^i(\tilde{x}, t) = \delta(\tilde{x} - x_0)$$

and assume Σ_t is a symmetric matrix of dimension 2x2.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The partial derivatives satisfy

$$\frac{d}{dt}\hat{\rho}(\tilde{x},t) = \dot{k}_t^i \hat{\rho}^i(\tilde{x},t) + [\frac{1}{2}(\tilde{x}-\mu_t^i)^T \Sigma_t^{i^{-1}} \dot{\Sigma}_t^i \Sigma_t^{i^{-1}} (\tilde{x}-\mu_t^i) + (\tilde{x}-\mu_t^i)^T \Sigma_t^{i^{-1}} \dot{\mu}_t^i] \hat{\rho}^i(\tilde{x},t)$$

$$\frac{\partial}{\partial \tilde{x}}\hat{\rho}^i(\tilde{x},t) = -(\tilde{x}-\mu_t^i)^T \Sigma_t^{i^{-1}} \hat{\rho}^i(\tilde{x},t)$$

$$\frac{\partial^2}{\partial \tilde{x}^2}\hat{\rho}^i(\tilde{x},t) = -\Sigma_t^{i^{-1}} \hat{\rho}^i(\tilde{x},t) + \Sigma_t^{i^{-1}} (\tilde{x}-\mu_t^i)(\tilde{x}-\mu_t^i)^T \Sigma_t^{i^{-1}} \rho^i(\tilde{x},t)$$

Substituting into (7.H.5) and equating coefficients we derive the expressions (7.H.1)-(7.H.3).

Next, we shall evaluate the integration

$$\int_{R^2} \rho^i(\tilde{x},t)d\tilde{x} = \int_{R^2} e^{F(x_1)} \hat{\rho}^i(\tilde{x},t)d\tilde{x}, \quad F(x_1) = \int_0^{x_1} \tanh(u)du$$

where x_{1_0} denotes the value of state variable x_1 at time $t = t_0$. Substituting $F(x_1) = \ln\cosh(x_1)$

$$\int_{R\otimes R} \rho^i(t,\tilde{x})d\tilde{x} = \int_{R\otimes R} \cosh(x_1)\hat{\rho}(\tilde{x},t)d\tilde{x} = k_t^i \int_{R\otimes R} \cosh(x_1)\hat{\rho}(\tilde{x},t)d\tilde{x}. \qquad (7.H.6)$$

At this point we concentrate in evaluating the integral

$$\int_{R^2} \cosh(x_1)\hat{\rho}(\tilde{x},t)d\tilde{x} = \int_{R^2} \frac{e^{x_1}+e^{-x_1}}{2} e^{-\frac{1}{2}(x-\mu_t^i)^T \Sigma_t^{i^{-1}} (\tilde{x}-\mu_t^i)} d\tilde{x} \qquad (7.H.7)$$

by using the property that any Gaussian density integrates to one.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

First, we define $x_1 \triangleq \frac{1}{2}G^T\tilde{x} + \frac{1}{2}\tilde{x}^TG$, $G^T = [1\ 0]$. If we express the exponent of the integrand of (7.H.7) as a perfect square we have

$$\int_{R^2} \cosh(x_1)\hat{\rho}(\tilde{x}, t)d\tilde{x} = \frac{1}{2}e^{\frac{1}{2}(\mu_t^{i\,T}G + G^T\mu_t^i + G^T\Sigma_t^iG)}\int_{R^2} e^{-\frac{1}{2}(\tilde{x}-(\mu_t^i+\Sigma_t^iG))^T\Sigma_t^{i\,-1}(\tilde{x}-(\mu_t^i+\Sigma_T^iG))}d\tilde{x}$$

$$+\frac{1}{2}e^{-\frac{1}{2}(\mu_t^{i\,T}G + G^T\mu_t^i - G^T\Sigma_t^iG)}\int_{R^2} e^{-\frac{1}{2}(\tilde{x}-(\mu_t^i-\Sigma_t^iG))^T\Sigma_t^{i\,-1}(\tilde{x}-(\mu_t^i-\Sigma_t^iG))}d\tilde{x}$$

$$= \frac{1}{2}(2\pi)(\det\Sigma_t^i)^{1/2}\left[e^{\frac{1}{2}(\mu_t^{i\,T}G + G^T\mu_t^i + G^T\Sigma_t^iG)} + e^{-\frac{1}{2}(\mu_t^{i\,T}G + G^T\mu_t^i - G^T\Sigma_t^iG)}\right]. \qquad (7.H.8)$$

Combining (7.H.6) and (7.H.8) we finally deduce

$$\int_{R^2} \rho^i(\tilde{x}, t)d\tilde{x} = (2\pi)(\det\Sigma_t^i)^{1/2}k_t^i e^{\frac{1}{2}G^T\Sigma_t^iG}\cosh\left(\frac{1}{2}(\mu_t^{i\,T}G + G^T\mu_t^i)\right). \qquad (7.H.9)$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

PLEASE NOTE

Copyrighted materials in this document have not been filmed at the request of the author. They are available for consultation, however, in the author's university library.

288-290

University Microfilms International

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

8. BIBLIOGRAPHY

[1] Baras, J. S., Blankenship, G. L. and Hopkins, W. E., "Existence, Uniqueness, and Asymptotic Behavior of Solutions to a Class of Zakai Equations with Unbounded Coefficients," IEEE Trans. Aut. Control, AC-28,2, (1983), pp. 203-214.

[2] Beneš, V. E. and Shepp, L. A., "Wiener Integrals Associated with Diffusion Processes," Teoriya Veroiatnostei i Prim., 3, (1968), pp. 498-501.

[3] Beneš, V. E., "Existence of Optimal Stochastic Control Laws," SIAM J. Control, 9, (1971), pp. 446-475.

[4] Beneš, V. E., "Exact Finite-Dimensional Filters for Certain Diffusions with Nonlinear Drift," Stochastic, 5, (1981), pp. 65-92.

[5] Beneš, V. E., "New Exact Nonlinear Filters with Large Lie Algebras," Systems and Control Letters, 5, (1985), pp. 217-221.

[6] Beneš, V. E. and Karatzas, I., "Estimation and Control for Linear Partially Observed Systems with Non-Gaussian Initial Distributions," Stochastic Processes and Their Applications, 14, (1983), pp. 223-248.

[7] Beneš, V. E., "Nonlinear Filtering in Sonar Detection and Estimation I: Probabilities of Detection Errors," Applied Stochastic Analysis, Stochastic Monographs, Vol. 5, Davis, M. H. A. and Elliott, R. J., eds., Gordon and Breach Science Publishers, New York, (1990), pp. 447-510.

[8] Bensoussan, A., "Optimal Control of Partially Observed Diffusions," Advances in Filtering and Optimal Stochastic Control, Fleming, W.H. and Gorostizar, L, eds., Springer-Verlag, Berlin, (1982), pp. 38-53.

[9] Bensoussan, A., Stochastic Control by Functional Analysis Methods, North-Holland, 1982.

[10] Bensoussan, A., "Maximum Principle and Dynamic Programming Approaches of the Optimal Control of Partially Observed Diffusion," Stochastics, 9, (1983), pp. 169-222.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

[11] Bensoussan, A., "Variational Methods in Stochastic Control," Nonlinear Filtering and Stochastic Control, Mitter, S.K. and Moro, A., eds., Lecture Notes in Mathematics, 972, Springer-Verlag, Berlin, (1983), pp. 1-39.

[12] Bensoussan, A., "Non Linear Filtering Theory," Recent Advances in Stochastic Calculus, Baras, J.S. and Mirelli, V., eds., Springer-Verlag, (1990), pp. 27-64.

[13] Bensoussan, A., Glowinski, R. and Rascanu, A., "Approximation of the Zakai Equation by the Splitting up Method," SIAM J. Optimization, 28, 6 (1990), pp. 1420-1431.

[14] Bensoussan, A., "Splitting up Method in the Context of Stochastic PDE," Stochastic Partial Differential Equations and their Applications, Rozovskii, B. L. and Sowers, R. B., eds., Lecture Notes in Control and Information Sciences, 176, Springer-Verlag, Berlin, (1991), pp. 23-31.

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

[16] Bismut, J. M., "An Introductory Approach to Duality in Optimal Stochastic Control, "SIAM Review, 20, 1, (1978), pp. 62-78.

[17] Bismut, J. M., "A Generalized Formula of Ito and Some Other Properties of Stochastic Flows," Z. Wahrscheinlichkeitstheorie Ver. Geb., 55, (1981) pp. 331-350.

[18] Bismut, J. M., "Martingales, the Malliavin Calculus and Hypoellipticity Under General Hormander's Conditions," Z. Wahrscheinlichkeitstheorie, Ver. Geb., 56, (1981), pp. 469-505.

[19] Bismut, J. M., "An Introduction to the Stochastic Calculus of Variations," Conference on Stochastic Systems, (1982), pp. 33.72.

[20] Bismut, J. M., "Partially Observed Diffusions and Their Control," SIAM J. Control and Optimization, 20, 2, (1982), pp. 302-309.

[21] Blankenship, G. L., Chang-Huan Liu and Marcus, S.I., "Asymptotic Expansions and Lie Algebras for Some Nonlinear Filtering Problems," IEEE Trans. on Automatic Control, AC-28, 7, (1983), pp. 787-797.

[22] BlagoveščČenskii, Y. N. and Freidlin, "Certain Properties of Diffusion Processes Depending on a Parameter," Soviet Mathematics, 2, (1961), pp. 633-636.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

[23] Bobrovsky, B. Z. and Zakai, M., "A Lower-Bound on the Estimation Error for Markov Processes," IEEE Trans. Automatic Control, AC-20, (1975), pp. 785-788.

[24] Bobrovsky, B. Z. and Zakai, M., "A Lower-Bound on the Estimation Error for Certain Diffusion Processes," IEEE Trans. Information Theory, IT-22,1, (1976), pp. 44-53.

[25] Boothby, W. M., An Introduction to Differentiable Manifolds and Riemannian Geometry, Academic Press, New York, (1975).

[26] Brockett, R. W., "Remarks on Finite Dimensional Nonlinear Estimation, "Analyse des System, Lobry, C., ed., Bordeaux, (1978), pp. 47-57.

[27] Brockett, R. W., "Classification and Equivalence in Estimation Theory," Proc. 18th IEEE Conference on Decision and Control, Ft. Lauderdale, FL., (1979), pp. 172-175.

[28] Brockett, R. W. and Clark, J. M. C., "The Geometry of the Conditional Density Equation," Analysis and Optimization of Stochastic Systems, Jacobs, O.L.R. et al., eds., New York, (1980), pp. 299-309.

[29] Brockett, R. W., "Nonlinear Systems and Nonlinear Estimation Theory," Stochastic Systems: The Mathematics of Filtering and Identification and Applications, Hazewinkel, M. and Williams, J.C., eds., Reidel Publishing Company, Holland, (1981), pp. 441-478.

[30] Brockett, R. W., Millman, R. S. and Sussmann, H. J., Differential Geometric Control Theory, Birkhäuser, Boston, (1982).

[31] Cameron, R. H. and Martin, W. T., "Transformations of Wiener Integrals under Translations, "Ann. Math., 45, 2, (1944), pp. 386-396.

[32] Cheng, D., Tarn, T. J. and Isidori, A., "Global Feedback Linearization of Nonlinear Systems," Proc. 23rd Conference on Decision and Control, Las Vegas, NV, (1984), pp. 74-83.

[33] Clark, J. M. C., "The Design of Robust Approximations to the Stochastic Differential Equations of Nonlinear Filtering," Communication Systems and Random Process Theory, Skwirzynski, J.K., ed., Nato Advanced Study Institute Series, Sijthoff and Noorshoff, Alphen aan des Rijn, (1978).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

[34] Cohen, M. and Levine J., "Deterministic Feedback Linearization, Girsanov Transformations and Finite-Dimensional Filters," Systems and Control Letters, 13, (1989), pp. 81-92.

[35] Curtain, R., "Stochastic Evolution Equations with General White Noise Disturbance," J. of Mathematic Anal. and Appl., 60, (1977), 570-595.

[36] Davis, M. H. A., "On the Existence of Optimal Control Policies in Stochastic Control," SIAM J. Control, 11, 4, (1973), pp. 587-594.

[37] Davis, M. H. A., "Martingale Methods in Stochastic Control," Stochastic Control and Stochastic Differential Systems, Lecture Notes in Control and Information Sciences, 16, Springer-Verlag, Berlin, (1979), pp. 85-117.

[38] Davis, M. H. A., "Pathwise Solutions and Multiplicative Functionals in Nonlinear Filtering," Proc. 18th Conference on Decision and Control, Ft. Lauderdale, FL, (1979), pp. 176-181.

[39] Davis, M. H. A., "On a Multiplicative Functional Transformation Arising in Nonlinear Filtering Theory," Z. Wahrscheinlichkeitstheorie Ver. Geb., 54, (1980), pp. 125-139.

[40] Davis, M. H. A., "Pathwise Non-Linear Filtering," Stochastic Systems: The Mathematics of Filtering and Identification and Applications, Hazewinkel, M. and Williams, J. C., eds., Reidel Publishing Company, Holland, (1981), pp. 505-528.

[41] Dayawansa, W., Boothby, W. M. and Elliott, D. L., "Global State and Feedback Equivalence of Nonlinear Systems," Systems and Control Letters, 6, (1985), 229-234.

[42] Di Masi, G. B., Pratelli, M. and Runggaldier, W. J., "An Approximation to the Nonlinear Filtering Problem with Error Bound," Stochastics, 14, 4, (1985), pp. 247-271.

[43] Di Masi, G. B. and Runggaldier, W. J., "Continuous-Time Approximations for the Nonlinear Filtering Problem, "Applied Mathematics and Optimization, 7, 3, (1981), pp. 233-245.

[44] Doss, H., "Liens Entre Equations Differentielles Stochastiques at Ordinairies," Ann. Inst. H. Poincaré, B, 13, (1977), pp. 99-125.

[45] Duncan, T. E., "Evaluation of Likelihood Functions," Information and Control, 13, (1968), pp. 62-74.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

[46] Dynkin, E. B., Markov Processes, Vol. 1-2, Springer-Verlag, Berlin, (1965).

[47] Elliott, R. J., Stochastic Calculus and Applications, Applications of Mathematics, Springer-Verlag, Berlin, 18, (1982).

[48] Elliott, R. J. and Glowinski, R., "Approximations to Solutions of the Zakai Filtering Equation," Stochastic Analysis and Applications, 7, 2, (1988), pp. 145-168.

[49] Elliott, R. J. and Kohlmann, "On the Existence of Optimal Partially Observed Controls," Applied Mathematics and Optimization, 9, (1982), pp. 41-63.

[50] Elliott, R. J. and Yang, H., "Control of Partially Observed Diffusions, "Journal of Optim. Th. Appl., 71, 3, (1991), pp. 485-501.

[51] Elworthy, K. D., "Stochastic Dynamical Systems and Their Flows," Stochastic Analysis, Friedman, A. and Pinsky, M., eds., Academic Press, New York, (1978), pp. 79-95.

[52] Evans, J. E., "Chernoff Bounds on the Error Probability for the Detection of Non-Gaussian Signals," IEEE Tran. Information Theory, IT-20, 5, (1974), pp. 569-577.

[53] Fleming, W. H. and Pardoux, E., "Optimal Control for Partially Observed Diffusions," SIAM J. Control and Optimization, 20, 2, (1982), pp. 261-285.

[54] Fleming, W. H. and Pardoux, E., "Stochastic Control with Partial Observations," Nonlinear Filtering and Stochastic Control, Mitter, S.K. and Moro, A., eds., Lecture Notes in Mathematics, 972, Springer-Verlag, Berlin, (1982), pp. 240-245.

[55] Fleming, W. H. and Nisio, M., "On Stochastic Relaxed Control for Partially Observed Diffusions," Nagoya Mathematical Journal, 93, (1984), pp. 71-108.

[56] Fleming, W. H. and Rishel, R. W., Deterministic and Stochastic Optimal Control, Springer-Verlag, Berlin, (1975).

[57] Fleming, W.H., "Exit Probabilities and Optimal Stochastic Control," Applied Mathematics and Optimization, 4, (1978), pp. 329-346.

[58] Fleming, W. H. and Miter, S. K., "Optimal Control and Nonlinear Filtering for Nondegenerate Diffusion Processes," Stochastics, 8, (1982), pp. 63-77.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

[59] Flochinger, P. and Le Gland, F., "Time-Discretization of the Zakai Equation for Diffusion Processes Observed in Correlated Noise," Stochastics, 35, (1991), pp. 233-256.

[60] Friedman, A., Partial Differential Equations of Parabolic Type, Prentice-Hall, N.J., (1964).

[61] Friedman, A., Stochastic Differential Equations and Applications, Vol. 1, Academic Press, New York, (1976).

[62] Friedman, A., Stochastic Differential Equations and Applications, Vol. 2, Academic Press, New York, (1976).

[63] Fujisaki, M., Kallianpur, G. and Kunita, H., "Stochastic Differential Equations for the Nonlinear Filtering Problems," Osaka J. Math., 9, (1972), pp. 19-40.

[64] Gertner, I., "An Alternative Approach to Nonlinear Filtering," Stochastic Processes and Their Applications, 7, (1978), pp. 231-246.

[65] Gihman, I. I., and Skorohod, A. V., Stochastic Differential Equations, Springer-Verlag, New York, (1972).

[66] Gihman, I. I. and Skorohod, A. V., The Theory of Stochastic Processes, Vol. 2, Springer-Verlag, Berlin, (1975).

[67] Girsanov, I. V., "On Transforming a Certain Class of Stochastic Processes by Absolutely Continuous Substitution of Measures," Theory of Probability and its Applications, 5, (1960), pp. 285-301.

[68] Haussmann, V. G., "The Maximum Principle for Optimal Control of Diffusions with Partial Information," SIAM J. Control and Optim., 25, 1, (1987), pp. 341-361.

[69] Hazewinkel, M., "On Deformations, Approximations and Nonlinear Filtering," Systems and Control Letters, 1, 1, (1981), pp. 32-36.

[70] Hazewinkel, M. and Marcus, S.I., "On the Relationship Between Lie Algebras and Nonlinear Estimation," Proc. IEEE Conf. on Decision and Control, Albuquerque, NM, (1980), pp. 66-71.

[71] Hazewinkel, M. and Marcus, S. I., "On Lie Algebras and Finite Dimensional Filtering," Stochastics, 7, (1982), pp. 29-62.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

[72] Hazewinkel, M., Marcus, S. I. and Sussmann, H. J., "Nonexistence of Finite-Dimensional Filters for Conditional Statistics of the Cubic Sensor Problem," Systems and Control Letters, 3, (1983), pp. 331-340.

[73] Hermann, R. and Krener, A. J., "Nonlinear Controllability and Observability," IEEE Trans. Automatic Control, AC-22, (1977), pp. 728-740.

[74] Hibey, J. L., "Error-Probability Bounds for Continuous-Time Decision Problems," Ph. D. Dissertation, Washington University, Saint Louis, Missouri, August, 1976.

[75] Hibey, J. L., Snyder, D. L. and van Schuppen, J. H., "Error-Probability Bounds for Continuous-Time Problems," IEEE Trans. Information Theory, IT-24, (1978), pp. 608-622.

[76] Hijab, O., "Asymptotic Nonlinear Filtering and Large Deviations," Advances in Filtering and Optimal Stochastic Control, Fleming, W. H. and Gorostiza, L. G., eds., Lecture Notes in Control and Information Sciences, 42, Springer-Verlag, New York, (1982), pp. 170-177.

[77] Hunt, L. R., Su, R. and Meyer, G., "Global Transformations of Nonlinear Systems," IEEE Trans. Automatic Control, AC-28, 1, (1983), pp. 24-30.

[78] Ikeda, N. and Watanabe, S., Stochastic Differential Equations and Diffusion Processes, North-Holland, (1981).

[79] Isidori, A., Nonlinear Control Systems, Springer-Verlag, Second Edition, (1989).

[80] Kailath, T., "A Note on Least-Squares Estimates For Likelihood Ratios," Information and Control, 13, (1968), pp. 534-540.

[81] Kailath, T., "A General Likelihood-Ratio Formula for Random Signals in Gaussian Noise," IEEE Trans. on Information Theory, IT-15, 3, (1969), pp. 350-361.

[82] Kailath, T., "A Further Note on a General Likelihood Formula for Random Signals in Gaussian Noise," IEEE Trans. on Information Theory, IT-16, 4, (1970), pp. 393-396.

[83] Kalman, R. E. and Bucy, R. S., "New Results in Linear Filtering and Prediction Theory," Transactions ASME J. Basic Engrg. 83D, (1961), pp. 95-108.

[84] Kallianpur, G., Stochastic Filtering Theory, Springer-Verlag, New York, (1980).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

[85] Kallianpur, G. and Striebel, C., "Estimation of Stochastic Systems: Arbitrary System Process with White Noise Observation Errors, "The Annals of Mathematical Statistics, 39, 3, (1968), pp. 785-801.

[86] Karatzas, I. and Shreve, Brownian Motion and Stochastic Calculus, Springer-Verlag, Berlin, (1988).

[87] Krishnaprasad, P. S., Marcus, S. J. and Hazewinkel, M., "System Identification and Nonlinear Filtering: Lie Algebras," Proc. IEEE Conference on Decision and Control, IEEE, New York, (1981), pp. 330-334.

[88] Krishnaprasad, P. S., Marcus, S. I. and Hazewinkel, M., "Current Algebras and the Identification Problem," Stochastics, 11, (1983), pp. 65-101.

[89] Krylov, N. V., Controlled Diffusion Processes, Springer-Verlag, (1980).

[90] Krylov, N. V. and Rosovskii, B. L., "On the Cauchy Problem for Linear Stochastic Partial Differential Equations," Math. USSR Izvestija, 11, 6, (1977), pp. 1267-1284.

[91] Kunita, H., "On the Representation of Solutions of Stochastic Differential Equations," Seminaire des Probabilities XIV, Lecture Notes in Mathematics, 784, (1980), pp. 282-303.

[92] Kunita, H., "On the Decomposition of Solutions of Stochastic Differential Equations," Stochastic Integrals, Proceedings of the Durham Symposium, (1980), Lecture Notes in Mathematics, Williams, D., ed., Springer-Verlag, Berlin, pp. 211-255.

[93] Kunita, H., "Cauchy Problem for Stochastic Partial Differential Equations Arising in Nonlinear Filtering Theory," Systems and Control Letters, 1, 1, (1981), pp. 37-41.

[94] Kunita, H., "Densities of a Measure-Valued Process Governed by a Stochastic Partial Differential Equation," Systems and Control Letters, 1, 2, (1981), pp. 100-104.

[95] Kunita, H., "On Backward Stochastic Differential Equations," Stochastics, 6, (1982), pp. 293-313.

[96] Kunita, H., "Stochastic Differential Equations, Stochastic Partial Differential Equations," Nonlinear Filtering and Stochastic Control, Mitter, S.K. and Moro, A., eds., Lecture Notes in Mathematics, 972, Springer-Verlag, (1981), pp. 100-169.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

[97] Kunita, H. and Watanabe, S., "On Square Integrable Martingales," Nagoya Math. Journal, 30, (1967), pp. 209-245.

[98] Kushner, H. J., "Dynamical Equations for Optimal Non-Linear Filtering," Journal on Differential Equations, 3, (1967), pp. 179-190.

[99] Kushner, H. J., Probability Method for Approximations in Stochastic Control and Elliptic Equations, Academic Press, New York, (1977).

[100] Kushner, H. J., "A Robust Discrete State Approximation to the Optimal Nonlinear Filter for Diffusions," Stochastics, 3, (1979), pp. 75-83.

[101] Kwakernaak, H., "A Minimum Principle for Stochastic Control Problems with Output Feedback," 1, 1, (1981), pp. 74-77.

[102] Le Gland, F., "Splitting-up Approximation for SPDE's and SDE's with Application to Nonlinear Filtering," Stochastic Partial Differential Equations and Their Applications, Rozovskii, B.L. and Sowers, R.B., eds., Lecture Notes in Control and Information Sciences, 176, Springer-Verlag, Berlin, (1992), pp. 177-187.

[103] Liptser, R. S. and Shiryayev, A. N., Statistics of Random Processes I., Springer-Verlag, New York, (1977).

[104] Malliavin, P., "Stochastic Calculus of Variations and Hypoelliptic Operators," Stochastic Differential Equations, Proc. Inter. Symp. SDE, Kyoto, John Wiley, (1976), pp. 195-264.

[105] Marcus, S. I., "Low Dimensional Filters for a Class of Finite State Estimation Problems with Poisson Observations," Systems and Control Letters, 1, 4, (1982), pp. 237-241.

[106] Mitter, S. K., "On the Analogy Between Mathematical Problems of Non-Linear Filtering and Quantum Physics," Richerche di Automatica, 10, 2, (1979), pp. 163-216.

[107] Naidu, D. S., Hibey, J. L., and Charalambous, C., "Neighboring Optimal Guidance for an Aeroassisted Optimal Transfer Vehicle in the Presence of Modelling Uncertainties," To Appear in IEEE Trans. Aerospace and Electronic Systems, (1993).

[108] Nijmeijer, H. and van der Schaft, A.J., Nonlinear Dynamical Control Systems, Springer-Verlag, New York, (1990).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

[109] Ocone, D. L., "Topics in Nonlinear Filtering Theory," Ph.D. Thesis, Massachusetts Inst. of Technology, Cambridge, (1980).

[110] Ocone, D. L., Baras, J. S. and Marcus, S. I., "Explicit Filters for Diffusions with Certain Nonlinear Drifts," Stochastics, 8, (1982), pp. 1-16.

[111] Oksendal, B., Stochastic Differential Equations, Springer-Verlag, Third Edition, Berlin, (1992).

[112] Pardoux, E., "Stochastic Partial Differential Equations and Filtering of Diffusion Processes," Stochastic, 3, (1979), pp. 127-167.

[113] Pardoux, E., "Backward and Forward Stochastic Partial Differential Equations Associated with a Non-Linear Filtering Problem," Proc. 18th Conf. on Decision and Control, Ft. Lauderdale, FL (1979), pp. 166-171.

[114] Pardoux, E., "Equations of Nonlinear Filtering and Applications to Stochastic Control with Partial Observations," Nonlinear Filtering and Stochastic Control, Mitter, S. K. and Moro, A., eds., Lecture Notes in Mathematics, 972, Springer-Verlag (1981), pp. 208-248.

[115] Pardoux, E., "Non-Linear Filtering, Prediction and Smoothing," Stochastic Systems: The Mathematics of Filtering and Identification and Applications, Hazewinkel, M. and Williams, J. C., eds., Reidel Publishing Company, Holland, 1981, pp. 529-559.

[116] Prohorov, Y. V. and Rozanov, Y. A., Probability Theory, Springer-Verlag, New York, (1969).

[117] Roth, Z. S. and Loparo, K. A., "Optimal Filter Realization for a Class of Nonlinear Systems with Finite-Dimensional Estimation Algebras," Systems and Control Letters, 4, (1984), pp. 23-26.

[118] Rishel, R. W., "Weak Solutions of a Partial Differential Equation of Dynamic Programming," SIAM J. Control, 9, 4, (1971), pp. 519-528.

[119] Runggaldier, W. H. and Di Masi, G. B., "An Approximation for the Nonlinear Filtering Problem with Error Bound," Stochastics, 14, 4, (1985), pp. 247 - 271.

[120] Sheu, S. J., "Solution of Certain Parabolic Equations with Unbounded Coefficients and Its Applications to Nonlinear Filtering," Stochastics, 10, (1983), pp. 31-46.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

[121] Snyder, D. L. and Rhodes, I. B., "Filtering and Control Performance Bounds with Implications on Asymptotic Separation, "Automatica, 8, (1972), pp. 747-753.

[122] Stephen, S. T., "Recent Resutls on Nonlinear Filteirng: New Classes of Finite-Dimensional Filters," Proc. 29th Conf. on Decision and Control, Honolulu, Hawaii, (1990), pp. 231-233.

[123] Striebel, C., "Optimal Control of Discrete Time Stochastic Systems," Lecture Notes in Economics and Mathematical Systems, 110, Springer-Verlag, Berlin, (1975).]

[124] Su, R., "On the Linear Equivalents of Nonlinear Systems," Systems and Control Letters, 2, 1, (1982), pp. 48-52.

[125] Sussmann, H. J., "Existence and Uniqueness of Minimal Realization of Nonlinear Systems," Math. Systems Theory, 10, (1977), pp. 263-284.

[126] Sussmann, H. J., "On the Gap Between Deterministic and Stochastic Ordinary Differential Equations," Ann. Probability, 6, (1978), pp. 19-41.

[127] van Schuppen, J. H., "Estimation Theory for Continuous Time Processes, a Martingale Approach," Ph.D. Dissertation, University of California, Berkley, (1971).

[128] Van Trees, H. L., Detection, Estimation, and Modulation Theory, John Wiley and Sons, New York, (1958).

[129] Wei, J. and Norman E., "On the Global Representation of the Solutions of Linear Differential Equations as a Product of Exponentials," Proc. Am. Math. Society, (1964).

[130] Wonham, W. M., "On the Separation Theorem of Stochastic Control," SIAM J. Control, 6,2, (1968), pp. 312 - pp. 326.

[131] Wong, E., Stochastic Processes in Information and Dynamical Systems, McGraw-Hill, New York, (1971).

[132] Wong, E. and Hajek, B., Stochastic Processes in Engineering Systems, Springer-Verlag, New York, (1985).

[133] Wong, W. S. "On a New Class of Finite Dimensional Estimation Algebras," System and Control Letters, 9, (1987), pp. 79-83.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

[134] Zakai, M., "On the Optimal Filtering of Diffusion Processes," Z. Wahrscheinlichkeitstheorie, Verg. Geb., 11, (1969), pp. 230 - 243.

[135] Zeitouni, B. Z. and Bobrovsky, B. Z., "On the Reference Probability Approach to the Equations of Nonlinear Filteing, "Stochastics, 19, (1986), pp. 133 - 149.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

9. AUTOBIOGRAPHICAL STATEMENT

Charalambos D. Charalambous was born in Cyprus on July 4, 1962. He recieved the B.S. and M.S. degrees in electrical engineering from Old Dominion University, Norfolk, VA, in May 1987 and December 1988, respectively.

Mr Charalambous has been suported by teaching assistanhip from 1988 through 1992, and by research assistanhip from 1988 through 1992.

He is the author of the following articles.

"Fuel Optimal Trajectories for Aeroassisted Coplanar Orbital Transfer problem," 27th IEEE Conference on Decision and Control, Austin, Texas, December 1988, and IEEE Transactions on Aerospace and Electronics Systems, Vol. 26, No.2, March 1990, pp.347-381.

"Neighboring Optimal Guidance for an Aeroassisted Orbital Transfer Vehicle in the Presence of Modelling Uncertainties," AIAA Guidance, Navigation, and Control Conference, Portland, Oregon, August 1990, and to Appear in IEEE Transactions on Aerospace and Electronics Systems.

"Neighboring Optimal Guidance for an Aeroassisted Orbital Transfer Vehicle in the Presence of Modelling and Measurement Uncertainties," AIAA 30th Aerospace Science Meeting, Reno, Nevada, January 1992.

"Differential Geometry Theory and Meassure Transformations Applied to Nonlinear Estimation and Control," SIAM Conference on Control and its Applications," Minneapolis, Minnesota, September 1992.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.