1973

Decentralized Filtering and Control in Interconnected Stochastic-Systems.

Charles Woodford Sanders Jr
Louisiana State University and Agricultural & Mechanical College

74-18,364

SANDERS, Charles Woodford, Jr., 1943-
DECENTRALIZED FILTERING AND CONTROL IN INTERCONNECTED STOCHASTIC SYSTEMS.

The Louisiana State University and Agricultural and Mechanical College, Ph.D., 1973
Engineering, electrical

University Microfilms, A XEROX Company, Ann Arbor, Michigan

Decentralized Filtering and Control

in

Interconnected Stochastic Systems

A Dissertation

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

in

The Department of Electrical Engineering

by
Charles Woodford Sanders, Jr.
B.S., Louisiana State University, 1965
M.S., Louisiana State University, 1968
December, 1973

TABLE OF CONTENTS

ACKNOWLEDGEMENT

During the course of this work the author has been a teaching assistant, a National Science Foundation Trainee, and a Dissertation Year Fellow. He is very grateful for this financial support, and for the freedom for independent research which it has provided.

It has been the author's pleasure to study under the direction of Professor Edgar C. Tacker of the Departments of Engineering Science, Electrical Engineering, and Chemical Engineering. In addition to his unswerving quest for knowledge and understanding, the energy, dedication to purpose, and friendship of Professor Tacker have been a constant source of encouragement throughout this period.

A number of the ideas presented in this work grew out of many enjoyable and helpful conversations with Professor Tacker and Thomas D. Linton. It is with pleasure that the author acknowledges this assistance.

Finally, the author would like to express his deep appreciation to his wife, Keri, for her constant support and encouragement during a sometimes difficult and trying period.

LIST OF FIGURES

LIST OF NOMENCLATURE

The following list of symbols and the page on which their initial definition appears should prove helpful to the reader. Symbols having only a temporary meaning are not listed.

ABSTRACT

Filtering and control techniques are developed for use in large-scale stochastic systems with particular attention focused on the information flow constraints which are often present in such a system. A decentralized filter having complexity constraints is derived using a variational approach and the relationships of this filter to the centralized Kalman filter are investigated. In particular it is shown that the decentralized filter can be designed by employing a computationally simple modification to the standard Kalman filtering Riccati equations. Qualitative aspects of the filter's performance are given in terms of an observability condition, and quantitative performance data are generated for the case of state estimation in a two-area electric power system.

In the decentralized control of large-scale stochastic systems, complexity constraints are imposed on the structure and a variational approach is used to develop necessary conditions for optimality. An apparently new derivation of the separation principle for centralized information patterns is employed to investigate the applicability of the separation principle for decentralized information patterns.

A design methodology is developed for an interaction modeling approach to decentralized controller design and the structural aspects of the local controllers are considered in detail.

CHAPTER I

INTRODUCTION AND PROBLEM FORMULATION

Systems engineering encompasses a wide range of general problem areas which can be both theoretically interesting and of practical importance. From one viewpoint, the central problem of interest to the systems engineer is that of controlling a given system. That is, a plant or collection of interconnected components which under "ideal operating conditions" is capable of performing some desired task, has been designed and is to be put into operation in the real world. The control design problem is then to design an auxiliary system which enables the plant to operate in the real world in a manner which in some prespecified sense is optimal or at least satisfactory. The design of such a controller must, of course, be carried out subject to certain constraints. Figure 1.1 gives a schematic representation of this very general problem and serves to introduce some of the notation used throughout this paper.

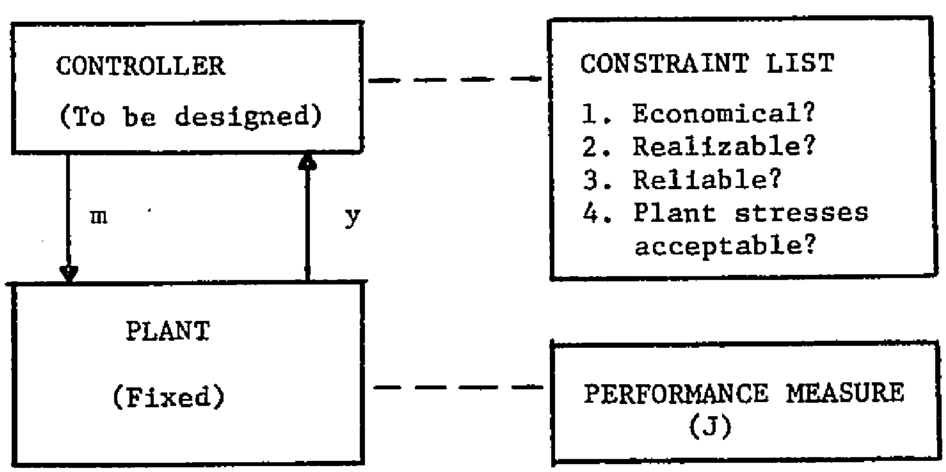

Fig. 1.1 General control problem.

As indicated in the above figure, y represents the observations which are available to the controller for use in generating the control signal, m. Loosely speaking, the control problem is then to determine a functional relationship between the observations and the control which satisfies the constraints and gives a satisfactory value of the performance measure. Solution of this design problem may proceed either analytically or experimentally with experiments being carried out on: (1) the actual system, (2) a physical prototype of the actual plant, or (3) a mathematical model simulation on a computer. Usually the most economical of the experimental methods is the simulation of the system on a mathematical model. Most successful design schemes utilize a combination of simulation and analytically based techniques. Of course, in order to carry out such schemes, it is necessary to develop reasonably accurate mathematical models of the process to be controlled.

A system model, P_M, is developed primarily from an understanding and mathematical formulation of the fundamental physical processes taking place in the actual system. Of course, the extent of this understanding and the subsequent mathematical formulation varies considerably depending on the processes involved. For example, in the aerospace area the basic processes are usually amenable to analysis by Newtonian mechanics and are well understood. Contrasted to this is the Fourdrinier process of papermaking in which the underlying physical processes are extremely complicated and knowledge is primarily of an empirical nature [1]*. However, regardless of the state of knowledge

*Square brackets are used to denote the references listed in the bibliography at the end of the paper.

concerning the process, boundaries of the system model must eventually be drawn somewhere and hence, in general, it is not possible to model all aspects of the system. Therefore, modeling uncertainties arise in addition to the uncertainties which are present due to the action of the real world environment on the system. These unmodeled forces and real world action on the plant are primarily unpredictable and unmeasurable in the real system and therefore must be modeled through some statistical means. This random character of the signals which affect the system will lead to a performance measure which is a random variable depending on the particular sample path taken by the stochastic processes. Thus, in this stochastic setting, it is necessary to use some statistically based measure of the performance. This is usually accomplished by using the mathematical expectation or variance of the original performance measure. Therefore, after the modeling phase has been completed, the control problem takes the general form shown in Figure 1.2.

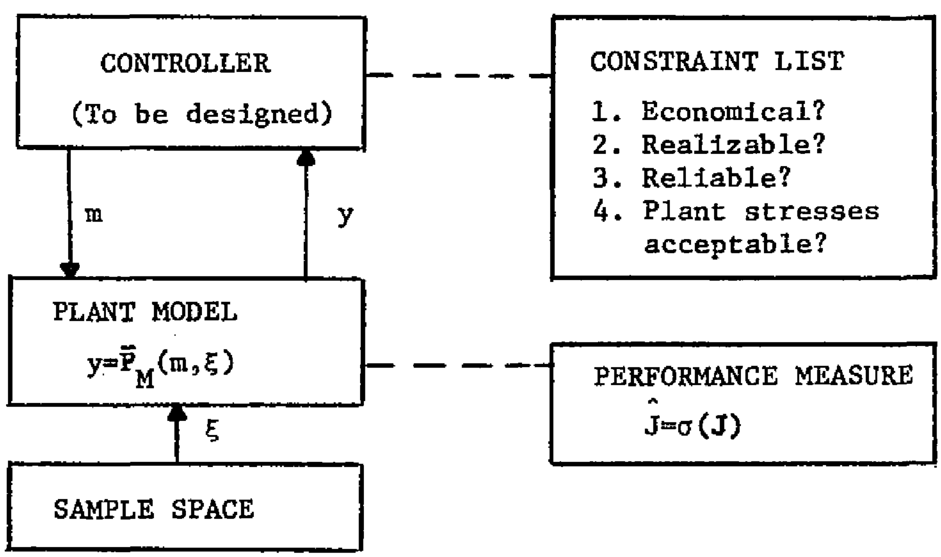

Fig. 1.2 General control problem after modeling.

Fundamental knowledge of physical processes is primarily embodied in the differential equations derived from limiting arguments based on a macroscopic view of the process and in the various conservation laws of physics [2-3]. A large class of processes may be represented via elements from the class of smooth dynamical systems [4]. For systems of this type, it is well known that a basic unifying concept is that of state [4]. After choosing a representation of the state, the system operator, $\bar{P}_M$, may be decomposed into a dynamical operation followed by a memoryless functional relationship. That is, letting x represent the state of P_M and $t \in [t_o, t_f]$ the system operation can be described by

$$x(t) = X(m[t_o,t],\ \xi[t_o,t],\ x(t_o),\ t) \qquad (1.1)$$

$$y(t) = Y(x(t),\ m(t),\ \xi(t),\ t) \qquad (1.2)$$

wherein

$$m[t_o,t] \overset{\Delta}{=} \{(m(\tau),\tau):\ \tau \in [t_o,t]\}$$

and similarly for ξ.

Roughly speaking, the state of the system is the information required in order to compute the system output at time t_2 based on knowledge of all inputs over the interval $[t_1,t_2]$ and the state at time t_1. Interest in this paper centers essentially on systems in which the state space is finite dimensional, although some of the results are potentially useful in suitable discretizations of infinite dimensional systems. Thus, the systems of interest here are those described by ordinary differential equations.

As noted in the above the constraints of the control problem play an important role and can not be overlooked if the control system

is to be useful in the real world. There are a number of constraints which may be present in a typical design problem. The primary focus in this research is on constraints which reflect the fact that in a practical system there may be limits on the amount of information which can be transmitted within the system. That is, we are interested in systems which are informationally decentralized in the sense that there is no central agency which is capable of processing all of the observations obtained from the system. This research area appears to be one of the most interesting and at the same time potentially useful of the research areas in the control of interconnected large-scale systems.

During the past several years, problems related to the control of large-scale systems have begun to attract a considerable amount of theoretical interest (see e.g. [5-8]). This interest and the concomitant classification of system control problems on the bases of system size and complexity stems primarily from the fact that the implementation of a controller derived from a straightforward application of existing stochastic optimal control theory often requires a prohibitive amount of data handling and computational capability.

Much of the research to date in large-scale system control has focused attention on problems arising in what might be called the "planning phase" of controller implementation. For example, the concept of ε-coupling [6] apparently arose out of the desire to approximate the solution of the requisite Riccati equations by a sequence of solutions to decoupled equations of lower dimensionality. Of course, such a technique is primarily computational in that the resulting controller structure is the same as that dictated by standard stochastic optimal control theory. Thus, the problem of managing the on-line

information flows remains. Controller design subject to constraints on the allowable information flow is that aspect of large-scale system control which is of interest in this paper.

The concept of decentralization [9] affords a technique for alleviating the on-line data handling requirements and has been the subject of extensive discussion in the literature of management science, econometrics, and organization theory (see e.g. [9-16]). Recently, as is evidenced by [17-21], interest in decentralized control and filtering structures has been shown in both the practical and theoretical control literature.

From a practical point of view this interest stems from the fact that many industrial processes are inherently large-scale and the large amount of data processing required by conventional stochastic control make these techniques impractical. In addition, improvements made in minicomputer technology have provided process control with an attractive alternative to completely centralized control structures. These improvements which have been launched by advances in MOS/LSI electronic technology have made the small computer very attractive for localized filtering and control operation under the supervision of a large central computer [22-23].

Theoretical interest in decentralized structures is sparked by the basic objective of this decentralization; viz., the efficient distribution of the information processing and computational effort involved in controlling the system leads to non-standard control and filtering problems. That is, an effective decentralization results in a structure involving multiple controllers each with different information on the system. Using the method of classification proposed in

Ho's survey of generalized control theory [24], design of decentralized control structures falls most naturally in the team theory class. In other words, the multiple controllers in the decentralized structure have the common goal of extremizing a common overall performance index.

Specifically, in this paper consideration is given to plants which can be effectively modeled as a given* collection $\{S_i: i=1,2,..N\}$ of N interacting dynamical subsystems, S_i. On the given time interval $[t_o,t_f]$ each subsystem is described by the Ito equation model [25-27]

$$S_i: \quad dx_i(t,\omega) = f_i(x_i(t,\omega),\ u_i(t,\omega),\ m_i(t,\omega),\ t)dt + dw_i \quad (1.3)$$

$$u_i(t,\omega) = L_i(x(t,\omega)) \quad (1.4)$$

$$dy_i(t,\omega) = H_i(z_i(t,\omega))dt + d\eta_i \quad (1.5)$$

$$z_i(t,\omega) = [x_i'(t,\omega) \quad u_i'(t,\omega)]' \quad (1.6)$$

wherein for each $t \in [t_o,t_f]$ and each ω in the underlying probability space, Ω, $x_i(t,\omega)$ is the n_i-vector valued state of S_i, $u_i(t,\omega)$ is the p_i-vector valued interaction input to S_i derived from the other subsystems, $y_i(t,\omega)$ is the q_i-vector valued observation available at S_i, $m_i(t,\omega)$ is the r_i-vector valued local control input to subsystem S_i, w_i is the local plant disturbance noise and η_i the local observation noise at S_i. The initial state, $x_i(t_o, \cdot)$, is assumed to be a gaussian random vector with mean x_i^o, and covariance Σ_i^o. The noise processes are assumed to be gaussian white processes with zero means and

*One can envision situations in which it is of interest to determine optimal decompositions of a system into subsystems. However, throughout this paper it is assumed that the decomposition is given.

covariances W_i and N_i respectively. Various independence assumptions between the stochastic variables will be needed. Unless otherwise stated to the contrary it will be assumed that the noise processes in S_i are stochastically independent from those in S_j for all $i \neq j$ and that for $i=j$ the plant disturbance noise and observation noise are independent. All noise processes are assumed to be independent of the initial state, and the initial state of S_i is assumed to be independent of the initial state of S_j for all $i \neq j$. Finally, without loss of generality, it is assumed that the observation noise processes are nondegenerate and thus the N_i $i=1,2,\ldots N$ are positive definite.

From equations (1.5-1.6) it is seen that in general we assume that the local observations at subsystem S_i consist of partial observations on the local state vector and the interaction input to S_i. In addition, from (1.3) and (1.4) it should be noted that we are assuming that the control signals are localized. That is, the control input to S_i influences the state of S_j $j \neq i$ only through its influence on the state of S_i and the subsequent effect of x_i on x_j. This would appear to be a valid assumption for a reasonably large class of systems.

It is often necessary to consider the n-vector valued composite state vector, $x = [x_1' \;\; x_2' \;\ldots\; x_N']'$, $n \triangleq \sum_{i=1}^{N} n_i$, and to use the projections* $P_{X,i}: R^n \to R^{n_i}$ defined by the rule

$$P_{X,i}(x) = x_i \;.$$

---

*R^n and $\prod_{i=1}^{N} R^{n_i}$ are different objects but are trivially isomorphic and hence no distinction is made between them. For brevity we usually write R^n in place of $\prod_{i=1}^{N} R^{n_i}$.

Clearly, the matrix representation of $P_{X,i}$ on the canonical bases for R^n and R^{n_i} is given by

$$[0 \quad 0 \ldots 0 \; I_{n_i} \; 0 \ldots 0 \quad 0]$$

wherein I_{n_i} is the canonical representation of the identity on R^{n_i} and the 0's are appropriately dimensioned null matrices. Since no confusion is likely to result, no distinction is made between $P_{X,i}$ and the matrix representation given above. Another map which is useful notationally is the map from R^{n_i} to R^n which takes x_i in R^{n_i} into $(0 \; 0 \ldots x_i \; 0 \ldots 0 \; 0)$ in $\prod_{i=1}^{N} R^{n_i}$ where the nonzero entry appears in the i-th position. It is readily seen that the canonical representation of this map is $P_{X,i}{}'$.

Similarly it is often necessary to consider the vector valued composite observation $y = [y_1' \quad y_2' \quad \ldots \; y_N']' \epsilon \prod_{i=1}^{N} R^{q_i}$, and the projection $P_{Y,i}$: $R^q(=\prod_{i=1}^{N} R^{q_i}) \rightarrow R^{q_i}$. A corresponding convention is followed for the composite control vector, $m \epsilon R^r(=\prod_{i=1}^{N} R^{r_i})$, and the projection $P_{M,i}$.

A schematic representation delineating the structure of the systems considered in this paper is given in Figure 1.3 in which only the plant and the probability space are shown and the stochastic signal ξ_i is used to represent $(w_i, \eta_i, x_i(t_o, \cdot))$. Note that essentially no loss of generality is entailed in the assumption that the interconnection grid is static--if this is not the case the interconnection dynamics may be incorporated into a fictitious (uncontrolled) subsystem.

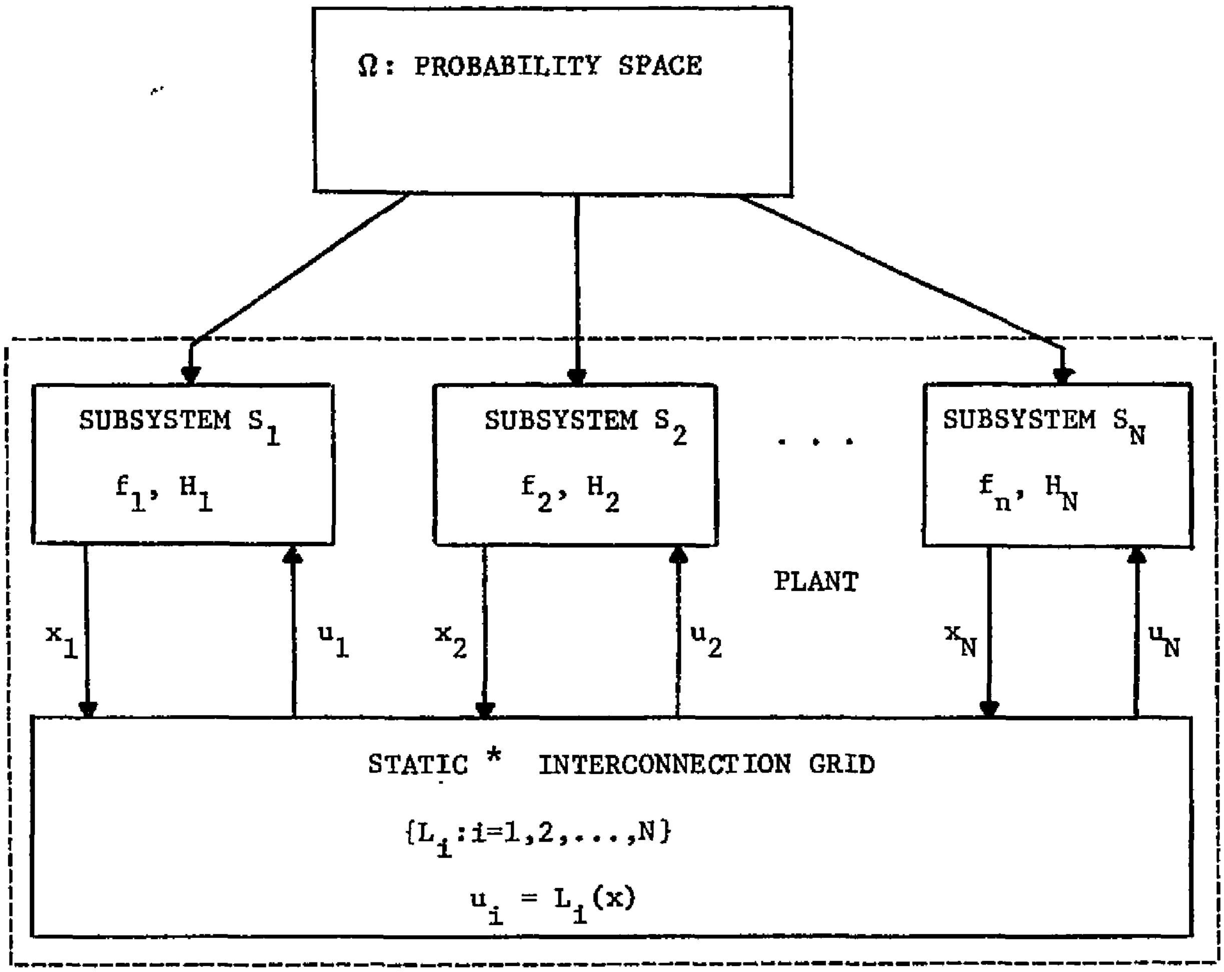

Fig. 1.3 A large-scale stochastic system.

The information describing a system of the above form is most naturally composed of three classes: (1) subsystem model data $I_{MD_i} = \{f_i, L_i, H_i\}$, (2) statistical parameters representing knowledge concerning the stochastic processes $I_{PD_i} = \{x_i^o, \Sigma_i^o, W_i, N_i\}$, and (3) on-line data obtained from observations on the operating system, $I_{OD_i}(t) = \{y_i(t), m_i(t)\}$. Data from classes (1) and (2) can be used in the

*Static is used here in the sense of nondynamic. This, of course, does not eliminate the case wherein the L_i are time-varying.

"planning" or pre-operational phase of controller implementation; whereas data from class (3) can be used only in the on-line phase. One can, of course, consider an adaptive situation in which on-line measurements are utilized to improve the controller performance. However, the decision rule or strategy used in implementing this adaptive procedure must be designed in the planning phase using only information from classes (1) and (2).

The viewpoint taken in this paper and the constraints imposed on the various controller structures are motivated by problems from the areas of electric energy systems, transportation systems, and process control systems. In each of these areas problems having the structure outlined above can occur quite naturally. The electric utility industry with its more than 3000 interconnected subsystems, its numerous power pools, and wide diversity of owners truly constitutes a large-scale system. Many proposed urban transportation systems incorporate a large number of privately occupied vehicles driven at relatively high speed under computer control on a single automated guideway. If one formulates the control problem for such a transportation system as that of maintaining a prespecified pattern or spacing between the vehicles, a system having essentially the above structure results. Of course, in the process control area the various processing stages are usually carried out in unit operations wherein each unit or subsystem has its own local controller. Thus process control problems fit quite naturally into the above structure.

In the majority of system monitoring and control techniques, it is necessary to estimate the state of the system from noisy output signals. Chapter II focuses attention on this filtering problem and

it is shown that under certain natural complexity constraints the decentralized filter can be realized via a computationally simple modification of the standard Kalman filtering Riccati equation computations. In addition, it is noted that this filter is equivalent to the Kalman filter subject to appropriate "local modeling" of the interactions as white noise processes. Of course, the optimal decentralized filter is suboptimal relative to the filter utilizing centralized information and this motivates the discussion of methods for evaluating this suboptimality. It is shown that for time-invariant systems if a certain observability condition is satisfied, then the degradation of the filter performance due to decentralized information approaches a bounded steady state value.

In Chapter III the control problem is considered. As in the case of the decentralized filter, complexity constraints are imposed on the controller structure and a variational approach is used to derive necessary conditions for the optimal parameters. A simple and apparently new derivation of the separation principle for centralized information patterns is given and the applicability of a separation principle in the decentralized case is discussed.

In an engineering approach to problems of decentralized control, one often takes the approach of modeling the interaction. Then the problem of controller design is localized--for example, the infinite-bus analysis of electric power systems. This approach is considered in Chapter IV and it is shown that a controller having a particularly attractive structure results if the interaction models are taken from a certain class. The structural properties and a design procedure for controllers based on this technique are considered. In Chapter V

conclusions and suggestions for further research are given.

In addition to some important intermediate results, the appendices give performance data for the application of the decentralized filter to a two-area power system.

CHAPTER II

DECENTRALIZED FILTERING VIA CONSTRAINED FILTERS

1. Introduction

One of the fundamental problems confronting any decision maker is that of processing information obtained from observations which are corrupted by unknown and unmeasurable signals over which he has no control. This very old problem in systems science can be traced to Gauss in his astronomical studies [28-29] and continues to be the subject of a large amount of research--see [30] for an execellent bibliography on this subject.

The modern approach to state estimation received its initial impetus from Wiener [31] and Kolmogorov [32] in their studies during the early 1940's. With the advent of the state concept in the early 1960's Kalman [33] and Kalman and Bucy [34] extended the Wiener-Kolmogorov results and devised efficient algorithms for the digital implementation of the estimator.

From the control viewpoint, interest in state estimation stems from the fact that in order to accurately control a system it is often necessary that accurate estimates of the system state be available. Thus, estimation is usually a first step in the implementation of a control strategy. Moreover, it has been shown [35-37] that in the case of linear systems, gaussian noise processes, quadratic performance functionals, and centralized information patterns the optimal controller structure consists of a conditional mean estimator in cascade with the

optimal deterministic controller. Recently, Brooks [38-39] has formulated the stochastic control problem as a minimum norm problem in Hilbert space, and has extended the separation theorem to the case in which the only requirements on the stochastic processes are that they have finite Hilbert norms and independent increments. However, a centralized information pattern is assumed. The case of nonclassical information patterns has been studied primarily by Witsenhausen [40] who has shown that in general the separation theorem does not hold in these cases. Nevertheless, in this chapter the filtering problem is considered separately from any control strategy. Justification for such an approach in the light of Witsenhausen's results are: (1) in a number of applications one is interested only in tracking the state of the system and thus the control problem does not enter into consideration, and (2) even though the separation theorem may lead to suboptimal controllers, it may be the best feasible approach; viz., a controller structure in which the controller and filter sections can not be designed separately may be too difficult to implement.

In this chapter as well as the majority of the remainder of the paper variational techniques are employed to determine the optimal element in a class of structurally constrained filters. The original stochastic optimization problem is reformulated into a deterministic optimal control problem via the following steps: (1) choose a class of implementable filter structures, (2) parameterize this class of structures, (3) reformulate the cost functional and system dynamics in terms of the parameters, and (4) regarding the parameters as control variables, apply techniques from deterministic control theory to derive necessary conditions for optimality.

This approach has been used by Athans and Tse [41] to give a direct derivation of the optimal linear centralized filter. An obvious advantage of such an approach is that the resulting filter structure is implementable. As a matter of fact this approach is in the same spirit as that of Johansen [42] and Sims [43] and allows one to place complexity constraints on the resulting filters. An obvious disadvantage to such a technique is the fact that unless the optimal unconstrained filter lies in the chosen class, the resulting filter will be sub-optimal relative to this unconstrained filter. In addition, even though the structural class is implementable, the optimal element may not exist or if it does exist the conditions specifying this element may be too complex to use as design tool.

For completeness, ease of reference, and to illustrate the variational technique, the chapter begins with a direct derivation of the optimal filter under the centralized information pattern in which the information

$$I_c(t) \triangleq \bigcup_{i=1}^{N} \{I_{MD_i} \cup I_{PD_i} \cup I_{OD_i}(t)\}$$

is available at a central location. Next, a natural class of complexity constrained filters is introduced for the case of completely decentralized information and it is shown that the optimal filter in this class can be obtained from a computationally simple modification of the standard Kalman filtering Riccati equation computations. Evaluation of the performance of this decentralized filter is then considered and it is shown that for time invariant systems a certain observability condition guarantees that the performance loss due to decentralization approaches a bounded steady state value.

II. The Centralized Filter [41]

Consider the classical case in which the plant is linear, the cost functional quadratic, and the information pattern centralized. That is, for each $i=1,2,\ldots,N$

$$f_i(x_i, u_i, m_i, t) = A_i x_i + L_{ii} u_i + B_i m_i \tag{2.1}$$

$$L_i(x) = \sum_{\substack{j=1 \\ j\neq i}}^{N} L_{ij} x_j \tag{2.2}$$

$$H_i(z_i) = H_i z_i \tag{2.3}$$

wherein A_i, B_i, H_i, and L_{ij} $i,j=1,2,\ldots,N$ are appropriately dimensioned matrices with real-valued entries which are continuous on the interval $[t_o,t_f]$. The loss function which is to measure the filter performance is defined by

$$J(t) = \mathscr{E}\{||e(t, \cdot)||_Q^2\} \tag{2.4}$$

wherein the estimation error, $e(t, \cdot)$, is the random variable defined by

$$e(t,\omega) = x(t,\omega) - \hat{x}(t,\omega)$$

$\hat{x}(t, \cdot)$ is the filter output, $\mathscr{E}$ is the a priori expectation operator, and Q is an nxn positive definite matrix. The centralized information pattern implies that there exists a single filter which has information $I_c(t)$ as defined in the previous section.

Using the above subsystem dynamics and the fact that the composite state vector can be written

$$x = \sum_{i=1}^{N} P'_{X,i} x_i$$

it is readily shown that the dynamics for the overall process are given by

$$dx(t,\omega) = [A\ x(t,\omega) + \tilde{m}(t,\omega)]dt + dw \qquad (2.5)$$

wherein

$$A = \sum_{i=1}^{N} P'_{X,i}(A_i P_{X,i} + L_{ii} L_i)$$

$$\tilde{m} = \sum_{i=1}^{N} P'_{X,i} B_i m_i$$

and

$$w = [w_1'\ w_2'\ \ldots\ w_N']'$$

Note that since the information is centralized, the filter has the necessary data from which to construct A and $\tilde{m}$. Similarly the equation describing the observation process, (2.3), may be written

$$dy(t,\omega) = H\ x(t,\omega)\ dt + d\eta \qquad (2.6)$$

wherein

$$H = \sum_{i=1}^{N} P'_{Y,i} H_i [P'_{X,i}\quad L_i']'$$

and

$$\eta = [\eta_1'\ \eta_2'\ \ldots\ \eta_N']'$$

As before, the centralized information pattern implies that the on-line data, $I_{OD_i}(t)$, and the model parameters are available to the filter and enable it to form the above composite observation vector.

Now let the class of implementable filters be those that are linear, smooth n-th order dynamic systems which give unbiased estimates. That is, the class of filters is specified by

$$\mathscr{F}: \quad \dot{\hat{x}}(t,\omega) = F\,\hat{x}(t,\omega) + C\,\tilde{m}(t,\omega) + G\,y(t,\omega) \qquad (2.7)$$

wherein F, G, C are continuous real-valued matrices on $[t_o,t_f]$, and

$$\mathscr{E}\{e(t,\cdot)\} = 0 \qquad \forall\ t\ \varepsilon[t_o,t_f] \qquad (2.8).$$

Using (2.5) and (2.7) it is easily shown that the estimation error satisfies

$$\dot{e}(t,\omega) = Fe(t,\omega) + (A-F-GH)x(t,\omega) + (I-C)\tilde{m}(t,\omega) + w - G\eta \qquad (2.9)$$

Recall that the expectation operation represents an integral and under certain reasonable conditions (see [78], p. 46) commutes with the operation of differentiation with respect to the parameter t. Thus

$$\mathscr{E}\{\dot{e}(t,\cdot)\} = \dot{\mathscr{E}}\{e(t,\cdot)\} \qquad (2.10)$$

Now, if $\mathscr{E}\{e(t,\cdot)\} = 0$ for each $t\ \varepsilon[t_o,t_f]$ then for each t it follows that $\dot{\mathscr{E}}\{e(t,\cdot)\} = 0$ and hence from (2.9), (2.10), and the fact that w and η are zero mean processes it follows that a necessary condition for unbiased estimates is

$$0 = (A-F-GH)\mathscr{E}\{x(t,\cdot)\} + (I-C)\ \mathscr{E}\{\tilde{m}(t,\cdot)\}$$

Therefore, in general, necessary conditions for unbiased estimates are

$$\begin{aligned} F &= A - G\,H \\ C &= I \end{aligned} \tag{2.11}$$

In addition, letting $t=t_o$ in (2.8) results in the condition

$$\hat{x}(t_o,\omega) = x^o \qquad \forall\ \omega\varepsilon\Omega \tag{2.12}$$

On the other hand using conditions (2.11) and (2.12) in (2.9) it follows that

$$\dot{\mathscr{E}}\{e(t,\ \cdot)\} = F\mathscr{E}\{e(t,\cdot)\}$$

$$\mathscr{E}\{e(t,\ \cdot)\}(t_o) = 0.$$

Therefore, conditions (2.11) and (2.12) are also sufficient for unbiased estimates. Thus the class, $\mathscr{F}$, of filters over which the optimization is to be performed is characterized by the stochastic differential equation

$$\begin{aligned} \dot{\hat{x}}(t,\omega) &= A\ \hat{x}(t,\omega) + G\left\{y(t,\omega) - H\ \hat{x}(t,\omega)\right\} + \tilde{m}(t,\omega) \\ \hat{x}(t_o,\omega) &= x^o \qquad\qquad \forall\ \omega\varepsilon\Omega \end{aligned} \tag{2.13}$$

wherein the filter gain, G, remains to be chosen in an optimal manner relative to the loss functional defined in (2.4). To this end, let V(t) denote the covariance of the estimation error. Since the estimates are unbiased, V(t) is given by

$$V(t) = \mathscr{E}\{e(t,\cdot)e'(t,\cdot)\} \tag{2.14}$$

Using the well-known fact that x'Ax = Trace(Axx') we have for each $t\ \varepsilon\ [t_o,t_f]$

$$\mathscr{E}\{||e(t_1, \cdot)||_Q^2\} = \mathrm{tr}\{Q\,\mathscr{E}\{e(t_1,\cdot)e'(t_1,\cdot)\}\}$$

and therefore

$$J(t_1) = \mathrm{tr}\{QV(t_1)\} \tag{2.15}$$

In order to formulate the optimization problem in a control-theoretic setting it is useful to derive a dynamical equation which reflects the effect of the parameter G on the error covariance. Note that

$$\dot{V}(t) = \mathscr{E}\{\dot{e}(t,\cdot)e'(t,\cdot) + e(t,\cdot)\dot{e}(t,\cdot)\} \tag{2.16}$$

and that for the case of unbiased estimates (2.9) becomes

$$\dot{e}(t,\omega) = (A-GH)e(t,\omega) + w - G\eta \tag{2.17}$$

Therefore using the fact that the expectation is a linear operator, the first term in (2.16) can be written

$$\mathscr{E}\{\dot{e}(t,\cdot)e'(t,\cdot)\} = (A-GH)V(t) + \mathscr{E}\{(w-G\eta)e'(t,\cdot)\} \tag{2.18}$$

Proceeding formally,solutions to (2.17) can be written in terms of the state transition matrix, Φ, as

$$e(t,\omega) = \Phi(t,t_o)(x(t_o;\omega)- x^o) + \int_{t_o}^{t} \Phi(t,\tau)(w-G\eta)(\tau)d\tau \tag{2.19}$$

Using (2.19) and the fact that w and η have independent increments and are stochastically independent from the initial state it follows that*

$$\mathscr{E}\{(w-G\eta)e'(t,\cdot)\} = [W + G\bar{N}G']/2 \tag{2.20}$$

---

*Recall that w and η are assumed to be stochastically independent.

The second term in (2.16) is the transpose of the first and thus collecting the results of (2.18) and (2.20) in (2.16) yields the desired dynamical equation

$$\dot{V}(t) = (A-GH)V(t) + V(t)(A-GH)' + W + G\bar{N}G' \qquad (2.21)$$

The boundary condition for (2.21) may be obtained from (2.14) and (2.12)

$$\begin{aligned} V(t_o) &= \mathcal{E}\{e(t_o,\cdot)e'(t_o,\cdot)\} \\ &= \mathcal{E}\{(x(t_o,\cdot)-x^o)(x(t_o,\cdot)-x^o)'\} \end{aligned}$$

or using the covariance of the initial state of each S_i and the fact that these initial states are gaussian and independent* there results

$$V(t_o) = \text{Block Diag. } \{\Sigma_1^o, \Sigma_2^o, \ldots, \Sigma_N^o\} \triangleq \Sigma^o \qquad (2.22)$$

Therefore the problem of optimizing the filter gain G may be recast as that of minimizing the cost, J, given in (2.15) subject to the dynamical equation (2.21) -(2.22) with G being regarded as a control input. This technique therefore allows one to convert the original stochastic optimization problem on the n-dimensional state space, R^n, to a deterministic optimization problem in the n^2-dimensional state space, $R^n x R^n$, and is similar in spirit to earlier work of the author [44-48] relative to the control of nonlinear stochastic dynamical

---

*The assumption of independence between the initial state of S_i and S_j for $i \neq j$ is, of course, not necessary here but will be useful in the decentralized case.

systems+ (see also [49]). In summary, the reformulated problem takes the form of a standard control problem in which there is only a penalty for the terminal ($t=t_1$) state: viz,

$\mathscr{P}$:

$$\text{Minimize} \quad J(t_1) = \text{trace}\left\{QV(t_1)\right\} \tag{2.15}$$

Subject to:

$$\dot{V}(t) = (A-GH)V(t) + V(t)(A-GH)' + W + G\bar{N}G' \tag{2.21}$$

$$V(t_o) = \Sigma^o \tag{2.22}$$

by choosing G from the set of continuous real-valued nxq matrices on $[t_o, t_f]$.

At this point one could write the coordinate-wise equations contained in (2.21) and apply Pontryagin's principle [50] to obtain necessary conditions for the optimal filter gain, G^*. However, Athans [51] has provided a notationally useful matrix minimum principle which makes it possible to apply Pontryagin's results directly to problems in which the state is most conveniently regarded as a matrix.

To apply this formalism to the above problem let P denote the nxn costate and Ψ the Hamiltonian for the above problem. Thus we have

$$\begin{aligned}\Psi &= \text{trace} \quad \{V(t)P'(t)\} \\ &= \text{tr}\{(A-GH)VP' + V(A-GH)'P' + WP' + G\bar{N}G'P'\}\end{aligned} \tag{2.23}$$

$$\dot{P} = \frac{-\partial\Psi}{\partial V} \quad , \quad P(t_1) = \frac{\partial\, \text{tr}\{QV(t_1)\}}{\partial\, V(t_1)} \tag{2.24}$$

+The primary difference is that this reformulation applied to nonlinear systems leads to deterministic optimization problems on an infinite dimensional state space.

$$\dot{V} = \frac{\partial \Psi}{\partial P} \quad , \quad V(t_o) = \Sigma^o \tag{2.25}$$

and

$$\left.\frac{\partial \Psi}{\partial G}\right|_{G=G^*} = 0 \tag{2.26}$$

First consider the differential equation for the costate. From (2.23) and the gradient matrix results of [51] it follows that

$$\frac{\partial \Psi}{\partial V} = (A-GH)P + P(A-GH)'$$

and

$$\frac{\partial}{\partial V(t_1)} \operatorname{tr}\{QV(t_1)\} = Q \tag{2.27}$$

Therefore from (2.24) the costate is specified by the differential equation

$$\dot{P}(t) = -(A-GH)P(t) - P(t)(A-GH)'$$

$$P(t_1) = Q$$

Clearly, if P is a solution to (2.27) then P' is also* and therefore by uniqueness of solutions to (2.27) [52] it follows that P is a symmetric matrix. Moreover, since Q is positive definite it follows that P is also (see Appendix A).

To obtain the necessary conditions for G^* note that the Hamiltonian can be written

$$\Psi = \Psi_o + \operatorname{tr}\{G\bar{N}G'P' - GHVP' - VH'G'P'\}$$

wherein Ψ_o does not depend explicitly on G. Applying the gradient matrix results of [51] and using the fact that P and V are symmetric

---

*Recall that Q is symmetric.

yields

$$\frac{\partial \Psi}{\partial G} = 2PG\bar{N} - 2PVH' \tag{2.28}$$

Thus from (2.26) and (2.28) a necessary condition for G to be optimal is

$$P(G^*\bar{N} - VH') = 0 \tag{2.29}$$

and since P and $\bar{N}$ are positive definite the necessary condition (2.29) is equivalent to

$$G^* = VH'\bar{N}^{-1} \tag{2.30}$$

Finally, using the above G^* in the dynamical equation (2.21) it follows that V evaluated along the optimal trajectory is given by

$$\dot{V} = AV + VA' - VH'\bar{N}^{-1}HV + W \quad , \quad V(t_o) = \Sigma^o \tag{2.31}$$

Equations (2.30) and (2.31) together with (2.13) are, of course, the well-known equations specifying the Kalman filter for centralized information patterns.

III. A Decentralized Filter

The optimality of the Kalman filter derived above when the class of implementable filters is expanded to include any for which the output is a measurable function with respect to the σ-algebra generated by the past data is well known (see e.g. [53]). This fact is, of course, not evident in the above derivation since the class of filters over which the optimization is carried out is restricted a priori.

However, there are certain disadvantages that place some constraint on the implementation of the filter. For example, suppose that

a decision maker, DM_i, located at subsystem S_i desires an estimate of the local state. Two important factors involved in determining the capability of DM_i to obtain accurate estimates of the local state are: (1) the information, $I_{DM_i}(t)$, available at S_i and (2) the computational resources available at S_i. Concerning (1), a multitude of cases exist relative to the classes of information introduced in Chapter I. That is, a number of relations between $I_{DM_i}(t)$ and $\{I_{DM_j}, I_{OD_j}, I_{PD_j}\}_{j=1}^{N}$ are possible. A natural assumption is that

$$I_{DM_i}(t) \supset I_{MD_i} \cup I_{PD_i} \cup I_{OD_i}(t)$$

i.e. DM_i understands (has a model for) and observes his local subsystem.

Relative to (2) consider the case in which the information pattern has the form

$$I_{DM_i}(t) = I_{MD} \cup I_{PD} \cup I_M(t) \cup I_{OD_i}(t) \tag{2.32}$$

where

$$I_{MD} = \bigcup_{j=1}^{N} I_{MD_j}$$

$$I_{PD} = \bigcup_{j=1}^{N} I_{PD_j}$$

and

$$I_M(t) = \bigcup_{j=1}^{N} m_j(t)$$

In other words, DM_i has complete knowledge of the dynamics of the overall system, the control inputs, and statistical parameters describing the stochastic processes--$I_{DM_i}(t)$ is deficient relative to

centralized information only because of the absence of some of the on-line data. Assuming that DM_i has unlimited computational resources, the best estimate of the local state by DM_i is obtained as the projection

$$\hat{x}_i = P_{X,i}\hat{x}$$

of $\hat{x}$ onto the i-th state space, R^{n_i}, with $\hat{x}$ being generated by the filter of section I with H in (2.12), (2.30-2.31) replaced by $\tilde{H}_i = H_i[P'_{X,i}\ L'_i]'$. Computationally this filter is essentially* as difficult to implement as the centralized filter since DM_i must solve a Riccati equation of the same dimensionality as in the centralized case and must implement a dynamical system (2.12) of the same dimensionality as the overall system.

In addition to the computational difficulties discussed above, the information pattern in a number of applications is much smaller than that given in (2.32). In this section we begin by placing explicit constraints on the filter structure to reflect factors (1) and (2). Then the variational technique is employed to derive the optimal filter in the chosen class. The information pattern chosen is the basic "localized" pattern

$$I_{DM_i}(t) = I_{MD_i} \cup I_{PD_i} \cup I_{OD_i}(t) \qquad (2.33)$$

and the computational resources available are assumed to be sufficient to implement a dynamical system of order n_i, the order of the local

---

*For a comprehensive study of the computational effort involved in implementing the centralized filter see [54]. Some results are available on reducing the computational effort [55].

system. Thus, in the absence of any interconnections, DM_i is capable of performing optimal filtering on S_i. The collection of local filters is regarded as a decentralized filter for the overall system and hence the cost functional is defined by

$$J = \mathscr{E}\{||e(t,\cdot)||_Q^2\}$$

wherein the fact that each DM_i $i=1,2,\ldots N$ is only interested in estimating his local state is made explicit by assuming

$$Q = Q_1 \dotplus Q_2 \dotplus \ldots \dotplus Q_N$$

with each Q_i an $n_i x n_i$ symmetric positive definite matrix and $\dotplus$ denoting the direct sum of the Q_i's. An alternate expression for Q is

$$Q = \sum_{i=1}^{N} P'_{X,i} Q_i P_{X,i}$$

and thus using the linearity of the expectation operation the overall cost, J, can be written

$$J = \sum_{i=1}^{N} J_i \tag{2.34}$$

where

$$J_i = \mathscr{E}\{||e_i(t,\cdot)||_{Q_i}^2\}$$

and the estimation error, $e_i(t,\cdot)$ is the random variable defined by

$$e_i(t,\cdot) = x_i(t,\cdot) - \hat{x}_i(t,\cdot)$$

with $\hat{x}_i(t,\cdot)$ the local filter output at S_i.

Now the class of implementable local filters is chosen to be those that are linear, smooth n_i-th order dynamic systems which give

locally unbiased estimates. That is, the local filter class is specified by*

$$\mathscr{F}_i: \qquad \dot{\hat{x}}_i(t) = F_i\hat{x}_i(t) + C_i m_i(t) + G_i y_i(t) \tag{2.35}$$

wherein F_i, G_i, C_i are continuous real-valued matrices on $[t_o,t_f]$, and

$$\mathscr{E}\{e_i(t)\} = 0 \qquad \forall\ t \in [t_o,t_f] \tag{2.36}$$

The optimization problem is then to find an element from this class of filters which is optimal relative to the loss function in (2.34).

First consider the characterization of the locally unbiased filter. From (2.1), (2.35) and the definitions of e_i and $\tilde{H}_i$ it follows that

$$\begin{aligned} \dot{e}_i = F_i e_i &+ (A_i P_{X,i} + L_{ii}L_i - F_i P_{X,i} - G_i\tilde{H}_i)x \\ &+ (B_i - C_i)m_i + w_i - G_i\eta_i \end{aligned} \tag{2.37}$$

Therefore using an argument similar to that leading to (2.11) and (2.12) the necessary conditions for unbiased estimates are+

$$\begin{aligned} (A_i - F_i)P_{X,i} + L_{ii}L_i - G_i\tilde{H}_i &= 0 \\ C_i &= B_i \\ \hat{x}_i(t_o,\omega) &= x_i^o \qquad \forall\, \omega\varepsilon\Omega \end{aligned} \tag{2.38}$$

---

*For ease of writing, explicit notation indicating sample path dependence will be dropped in the sequel.

+Concerning the special assumption that the local control for S_j does not feed directly to S_i ($j \neq i$), note that if this assumption does not hold then there will, in general, not exist a locally unbiased filter under the information pattern of (2.33). If one increases the information set to include the control inputs that feed directly to S_i, then it may be possible to obtain unbiased estimates.

On the other hand, using conditions (2.38) in (2.37) it follows that these conditions are also sufficient for locally unbiased estimates.

The first condition in (2.38) can be considered in more detail by partitioning H_i conformably with $[P'_{X,i} \mid L'_i]'$. This partition is of the form

$$H_i = [\bar{H}_{i1} \mid \bar{H}_{i2}]$$

wherein $\bar{H}_{i1}$ and $\bar{H}_{i2}$ are $q_i x n_i$ and $q_i x p_i$ matrices respectively. Using this partitioning and the detailed structure, (2.2) of L_i, the first condition in (2.38) becomes

$$F_i = A_i - G_i\bar{H}_{i1}$$

and

$$(L_{ii} - G_i\bar{H}_{i2})L_{ij} = 0 \qquad j=1,2,\ldots N \;\; j\neq i \tag{2.39}$$

The questions of existence and uniqueness hinge on the second condition in (2.39). Relative to the question of uniqueness it is advantageous to define the class of <u>surely locally unbiased</u> (S.L.U.) filters characterized by the subset, Γ_i^o, of $n_i x q_i$ matrices defined by

$$\Gamma_i^o = \{G_i\colon L_{ii} - G_i\bar{H}_{i2} = 0\}$$

Now define the subspace $\mathscr{L}_i$ of $n_i x p_i$ matrices by

$$\mathscr{L}_i = \{\Theta\colon \Theta L_{ij} = 0, \;\; \forall j=1,2,\ldots N/\{i\}\}$$

and the subset Γ_i of $n_i x\, q_i$ matrices by

$$\Gamma_i = \{G_i\colon \exists\, \Theta\varepsilon\mathscr{L}_i \ni L_{ii} - G_i\bar{H}_{i2} + \Theta = 0\}$$

Note that the second condition of (2.39) is equivalent to Γ_i. Also since $0 \in \mathcal{L}_i$ it follows that $\Gamma_i^o \subset \Gamma_i$. Thus, if DM_i has perfect knowledge of the interaction maps, $\{L_{ij}: j=1,2,\ldots,N/\{i\}\}$, for S_i he can (potentially , at least) use this information to do better than the S.L.U. class of filters. However if DM_i does not have enough confidence in the L_{ij} data and at the same time places great importance on obtaining unbiased estimates he may choose to optimize over the class of surely locally unbiased filters (this class may, of course, be empty). In the following, attention is focused on the S.L.U. case. Regarding the existence of a member of the S.L.U. class, a special assumption is now made which will be sufficient to guarantee this existence. Unless explicitly stated to the contrary, this assumption will be used throughout the remainder of the paper.

$\bar{H}_{i2}$ ASSUMPTION: For each $i=1,2,\ldots,N$ assume that $q_i > p_i$ and that $\bar{H}_{i2}$ has full rank, p_i.

Comments: Recall that q_i is the dimensionality of the output space of the on-line information available to DM_i and that this output consists of information on the local state and the interaction input to S_i. $\bar{H}_{i2}$ satisfying the above assumption means intuitively that DM_i can observe all of the interactions to S_i. These interaction observations may be mixed with observations on the local state through $\bar{H}_{i1}$. This assumption allows us to assume without further loss of generality that $\bar{H}_{i2}$ has the form

$$\begin{bmatrix} I_{p_i} \\ \hline 0 \end{bmatrix} \updownarrow q_i - p_i \tag{2.40}$$

since if this is not the case there exists a nonsingular $p_i x p_i$ matrix Z_i and a nonsingular $q_i x q_i$ matrix Z_2 such that

$$H^o_{i2} = Z_2^{-1}\bar{H}_{i2}Z_1$$

wherein H^o_{i2} is of the above specified form ([56], Thm. 4.1, p.52). Then the replacements $L_{ii} \to L_{ii}Z_1$, $L_i \to Z_1^{-1} L_i$, $\eta_i \to Z_2^{-1} \eta_i$, $\bar{H}_{i1} \to Z_2^{-1} \bar{H}_{i1}$, and $G_i \to G_i Z_2$ convert the original problem into an equivalent problem in which H^o_{i2} has the structure given in (2.40). Of course once $G_i Z_2$ has been computed it is a simple operation to compute the desired G_i.

Now partition G_i to conform with the structure in (2.40)

$$G_i = [G_{i1} \mid G_{i2}]$$

and $\bar{H}_{i1}$ to conform with the structure of G_i. That is

$$\bar{H}_{i1} = \begin{bmatrix} H_{i1} \\ \hline H_{i3} \end{bmatrix}$$

Then from (2.38), (2.39) and the definition of the S.L.U. class of filters it follows that this class is characterized by the relations

$$F_i = A_i - L_{ii} H_{i1} - G_{i2} H_{i3}$$

$$C_i = B_i \qquad \text{(S.L.U)}$$

$$\hat{x}_i(t_o,\omega) = x_i^o \qquad \forall \omega \in \Omega$$

The filter gain, G_{i2}, remains to be chosen in an optimal manner relative to the loss functional in (2.34). The development here proceeds along the same lines as those for the centralized case. First,

note from (2.37) that in the case of S.L.U. the error e_i evolves according to

$$\dot{e}_i(t) = F_i e_i(t) + w_i - G_i \eta_i$$

and therefore the overall error, $e \triangleq \sum_{i=1}^{N} P'_{X,i} e_i$, evolves according to

$$\dot{e}(t) = Fe(t) + \hat{w} \tag{2.41}$$

wherein

$$F = \sum_{i=1}^{N} P'_{X,i} F_i P_{X,i}$$

$$\hat{w} = \sum_{i=1}^{N} P'_{X,i}(w_i - G_i \eta_i)$$

Note that since $\hat{w}$ is a linear transformation on the composite process, $[w' \;\; \eta']'$, it inherits all of the statistical properties of this process. Namely, $\hat{w}$ is a zero mean white gaussian process which is stochastically independent of the initial state. Moreover, letting $\hat{w}_i$ denote the components of $\hat{w}$ (in $\prod_{i=1}^{N} R^{n_i}$) it follows that $\hat{w}_i$ and $\hat{w}_j$ are stochastically independent for $i \neq j$. Therefore, proceeding in a manner similar to that used in the development leading to (2.21), it is readily shown that the error covariance satisfies the differential equation

$$\dot{V}(t) = FV(t) + V(t)F' + \hat{W} \tag{2.42}$$

$$V(t_o) = \Sigma^o = \text{Block diagonal } \{\Sigma_i^o : i=1,2,\ldots,N\}$$

$$= \sum_{i=1}^{N} P'_{X,i} \, \Sigma_i^o \, P_{X,i}$$

with

$$\hat{W} = \mathscr{E}\{\hat{w}\hat{w}'\} = \sum_{i=1}^{N} P'_{X,i}(W_i + G_i N_i G'_i) P_{X,i} \quad .$$

Clearly, as can be verified by direct substitution, the solution to (2.42) can be written as

$$V(t) = \sum_{i=1}^{N} P'_{X,i} V_i(t) P_{X,i}$$

where each $V_i(t)$ is a solution to

$$\dot{V}_i(t) = F_i V_i(t) + V_i(t) F'_i + W_i + G_i N_i G'_i \qquad (2.43)$$

$$V_i(t_o) = \Sigma_i^o$$

Note that the V_i in (2.43) is the error covariance for the i-th local filter and thus

$$J_i(G_{i2},t) = \mathrm{tr}\{Q_i V_i(t)\}$$

Therefore since the error for the decentralized filter is the sum (2.34) of these local costs with each $J_i > 0$ and independent of G_{j2} for $j \neq i$, it follows that the original problem is equivalent to N local problems of the form

$\mathscr{P}_i$: Find G^*_{i2} such that for each $t_1 \in [t_o, t_f]$ $J_i(G^*_{i2}, t_1) \leq J_i(G_{i2}, t_1)$ for every G_{i2} in the set of continuous real-valued $n_i x(q_i - p_i)$ matrices, subject to the constraint (2.43) with

$$G_i = [L_{i1} \quad G_{i2}]$$

$$F_i = A_i - L_{i1}H_{i1} - G_{i2}H_{i3} = \bar{A}_i - G_{i2}H_{i3}.$$

To progress toward the variational solution to $\mathscr{P}_i$, partition the observation noise vector, η_i, to conform to the partitioning of G_i. Corresponding to this partitioning the covariance matrix, N_i, takes the

form

$$N_i = \left[\begin{array}{c|c} N_{i1} & N_{i3} \\ \hline N_{i3}' & N_{i2} \end{array}\right]$$

and thus the differential equation (2.43) for V_i can be written

$$\dot{V}_i = (\bar{A}_i - G_{i2}H_{i3})V_i + V_i(\bar{A}_i - G_{i2}H_{i3})' + \bar{W}_i +$$
$$G_{i2}N_{i2}G_{i2}' + G_{i2}N_{i3}'L_{ii}' + L_{ii}N_{i3}G_{i2}' \qquad (2.44)$$

wherein $\bar{W}_i = W_i + L_{ii}N_{i1}L_{ii}'$.

Note that except for the terms involving N_{i3}, (2.44) has the same form as (2.21). Therefore if $N_{i3}=0$, then the solution to the local filtering problem can be obtained from (2.30) and (2.31) by making the replacements $A \to \bar{A}_i$, $G \to G_{i2}^*$, $H \to H_{i3}$, $N \to N_{i2}$, $W \to \bar{W}_i$, $V \to V_i$, $\Sigma^o \to \Sigma_i^o$, and $Q \to Q_i$.

Now consider the case in which $N_{i3} \neq 0$. Since this development proceeds along the same lines as that leading to (2.30) and (2.31), only the highlights will be given and the similarities between (2.44) and (2.21) will be exploited. Forming the Hamiltonian as in (2.23) note that the only difference is in the terms involving N_{i3}. Since these terms are not explicit functions of V_i, the equation for the co-state will be the same as in the $N_{i3}=0$. However, as noted above in the discussion of the $N_{i3}=0$ case this costate equation is the same as (2.27) with the above replacements. Therefore the discussion following (2.27) is applicable to this case. Next the gradient matrix results can be applied to yield the necessary conditions for G_{i2}^*. It is readily seen that

$$\frac{\partial \Psi}{\partial G_{i2}} = 2P(G_{i2}N_{i2} + L_{ii}N_{i3} - V_iH_{i3}')$$

from which it follows that G_{i2}^* is given by

$$G_{i2}^* = (V_iH_{i3}' - L_{ii}N_{i3})N_{i2}^{-1} \tag{2.45}$$

Substituting (2.45) into (2.44) and applying some algebra it can be shown that V_i is given as the solution of the Riccati equation

$$\dot{V}_i = A_i^*V_i + V_iA_i^{*'} - V_iH_{i3}'N_{i2}^{-1}H_{i3}V_i + W_i^* \tag{2.46}$$

$$V_i(t_o) = \Sigma_i^o$$

wherein

$$A_i^* = A_i + L_{ii}(N_{i3}N_{i2}^{-1}H_{i3} - H_{i1})$$

and

$$W_i^* = W_i + L_{ii}(N_{i1} - N_{i3}N_{i2}^{-1}N_{i3}')L_{ii}' \; .$$

Therefore the local filter from the surely locally unbiased class is specified by (2.35), (2.45), and (2.46). Writing (2.35) in detail yields

$$\dot{\hat{x}}_i = A_i\hat{x}_i + L_{ii}(y_{i1} - H_{i1}\hat{x}_i) + B_im_i + G_{i2}^*(y_{i2} - H_{i3}\hat{x}_i)$$

$$\hat{x}_i(t_o) = x_i^o \tag{2.47}$$

A schematic representation of this local filter, $\mathscr{F}_i$, together with S_i is shown in Figure 2.1. For convenience, it is assumed in the construction of this figure that the measurement matrix, $\bar{H}_{i2}$, has the canonical structure given in (2.40).

If the system measurement matrix $\bar{H}_{i2}$ does not have the canonical

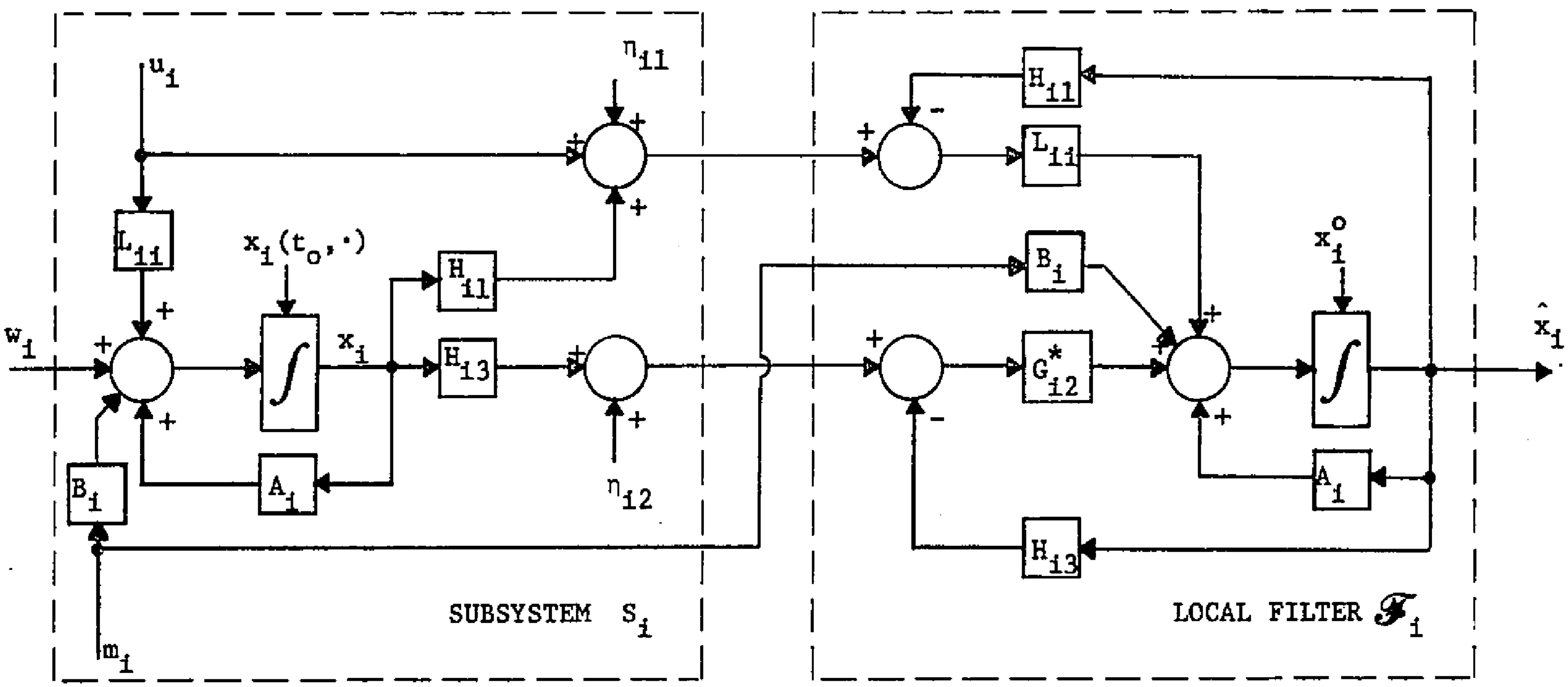

Fig. 2.1 Subsystem S_i and its Local Filter $\mathscr{F}_i$

structure, then the following procedure must be implemented in the optimal gain calculations, and subsequent construction of the filter.

Step 1. Find the Z_1 and Z_2 such that $Z_2^{-1}\,\bar{H}_{i2}\,Z_1$ has the structure given in (2.40).

Step 2. Make the replacements $L_{i1} \rightarrow L_{i1}Z_1$, $L_i \rightarrow Z_1^{-1}L_i$, $N_i \rightarrow Z_2^{-1}N_i(Z_2^{-1})'$, $\bar{H}_{i1} \rightarrow Z^{-1}\bar{H}_{i1}$.

Step 3. Identify H_{i1}, H_{i3}, N_{i1}, N_{i2}, N_{i3} from the following+

$$\bar{H}_{i1} = \begin{array}{c} \left[\begin{array}{c} H_{i1} \\ \hline H_{i3} \end{array}\right] \end{array} \begin{array}{l} p_i \\ q_i - p_i \end{array} \qquad N_i = \left[\begin{array}{c|c} N_{i1} & N_{i3} \\ \hline N_{i3}' & N_{i2} \end{array}\right] \begin{array}{l} p_i \\ q_i - p_i \end{array}$$

Step 4. Solve for G_i^o using the above replacements in (2.45) and (2.46).

Step 5. Compute $G_i^* = [L_{i1} \quad G_i^o]Z_2^{-1}$

Step 6. Implement the dynamical system++

$$\dot{\hat{x}}_i = A_i\hat{x}_i + B_i m_i + G_i^*(y_i - \bar{H}_{i1}\hat{x}_i)$$

$$\hat{x}_i(t_o) = x_i^o$$

It is interesting to consider some special cases of the above results. For convenience consideration is limited to systems in which $\bar{H}_{i2}$ has the canonical structure. First, as a partial check on these

---

+Of course, these identifications are made on the $\bar{H}_{i1}$ and N_i which result after the implementation of Step 2.

++The original $\bar{H}_{i1}$.

results, suppose that $H_{i1} = 0$ and $\eta_{i1} = 0$. That is, complete noise-free observations of the interactions are available. In this case $N_{i1} = 0$, $N_{i3} = 0$, $A_i^* = A_i$, $W_i^* = W_i$, and $G_{i2}^* = V_i H_{i3}' N_{i2}^{-1}$. Therefore the filter is the "standard" filter for S_i obtained by regarding the interaction input as a known control input. This is, of course, as it should be.

Secondly consider the more interesting case in which $H_{i1} = 0$ and $N_{i3} = 0$. That is, complete noisy observations of the interactions are available and the noise corrupting the interaction measurements is independent of the noise corrupting the partial measurements of the local state. In this case $A_i^* = A_i$, $W_i^* = W_i + L_{ii} N_{i1} L_{ii}'$, and $G_{i2}^* = V_i H_{i3}' N_{i2}^{-1}$. Therefore the gain, G_{i2}^*, in this case is the "standard" gain that would be obtained by ignoring the fact that observations of the interactions are available and modeling the interaction as a zero mean white gaussian process with the same statistics (N_{i1}) as the observation noise, η_{i1}, corrupting the interaction measurements. Once this gain has been calculated the interaction measurements are treated as if they were noise-free and are fed directly to the filter. This observation indicates to some extent the nature of the approximations inherent in this class of complexity constrained filters. Since the interaction is a linear transformation on the composite state, x, which evolves according to (2.5), it follows that the interaction input to S_i is actually a colored gaussian process. However, as noted above, the filter effectively "thinks" that the interaction input is a white noise process and therefore expects the local state to be more rapidly changing. This will lead to higher gains than those obtained for the local filter which takes the dynamics of the other subsystems into account (the filter

discussed at the beginning of this section). In the next section a more analytically based discussion of the filter performance relative to the centralized filter is given.

IV. Performance of the Decentralized Filter

This section is motivated by the fact that although the above filter is optimal within the surely locally unbiased class of filters, it is suboptimal relative to less restricted classes. This suboptimality stems primarily from the decentralization of two basic sources of information. Namely, (1) each DM_i uses only the local model data, I_{MD_i}, and local process data, I_{PD_i}, in designing the local filter, and (2) each DM_i processes only the local observations with the information relative to the remainder of the system limited to the on-line interaction observations. The trade-offs here are that (1) reduces the complexity of the filter that DM_i must implement and (2) reduces the complexity of the communication system connecting the DM_i--indeed, the only communication between the DM's is the "naturally" occurring communication in the underlying system.

In view of the above there exist several filters with which it is natural to compare the decentralized filter in evaluating the trade-offs. Relative to (1) it is natural to consider the local filters individually and compare them with the local filters designed under the information pattern $I_{DM_i} = I_{MD} \cup I_{PD} \cup I_{OD_i}(t)$, and concerning (2) to regard the collection of local filters as a single decentralized filter for the overall system and compare it with the filter designed under the information pattern $I_c = I_{MD} \cup I_{PD} \cup I_{OD}(t)$.

For both of the above comparisons it is possible to use the Riccati equations involved in computing the error covariance. Consider, for example, the case in which the collection of local filters is regarded as a single decentralized filter for the overall process. In measuring the performance loss engendered by decentralization, it is natural to define the cost increase by

$$\Delta J(t) \overset{\Delta}{=} J_D(t) - J_C(t)$$

wherein J_D and J_C are respectively the cost associated with the decentralized and centralized filters. In terms of the associated error covariance matrices, V_D and V_C the cost increase may be written

$$\Delta J(t) = \mathrm{tr}\ \{Q\Delta V(t)\}$$

with

$$\Delta V(t) = V_D(t) - V_C(t).$$

The associated Riccati equations for the error covariances are+

$$\dot{V}_i = A_i^* V_i + V_i A_i^{*\prime} - V_i H_{i3}' N_{i2}^{-1} H_{i3} V_i + \overset{*}{W}_i \tag{2.46}$$

with $$V_i(t_o) = \Sigma_i^o$$

$$V_D = \sum_{i=1}^{N} P_{X,i}' V_i P_{X,i}$$

and

$$\dot{V}_c = AV_c + V_c A' - V_c H' \bar{N}^{-1} H V_c + W \tag{2.31}$$

with $$V_c(t_o) = \sum_{i=1}^{N} P_{X,i}' \Sigma_i^o P_{X,i}$$

+It is assumed throughout the present discussion that the measurement matrices have been put into canonical form.

For the case in which the system is time-invariant it is of interest to determine conditions under which the increase in cost approaches a steady state value and does not grow without bound. In order for this cost increase to be bounded it suffices for the error covariance Riccati equations (2.31) and (2.46) to possess steady state solutions. It is well known [34] that sufficient conditions+ for an equation of the form (2.31) to possess a steady state solution can be phrased in terms of the fundamental concept of observability [57]. Namely, if the pair $[A,H]$ is observable then (2.31) possesses a steady state solution. Since the structure of (2.46) is the same as that of (2.31) it follows that (2.46) has a steady state solution provided the pair $[A_i^*, H_{i3}]$ is observable. Now, it can also be shown that $[A,H]$ is observable if and only if $[A+KH, H]$ is observable for every matrix K. Thus, recalling that (see (2.46), p.36) A_i^* is defined in terms of the fundamental system matrices as

$$A_i^* = A_i - L_{ii}H_{i1} + L_{ii}N_{i3}N_{i2}^{-1}H_{i3}$$

it follows that a sufficient condition for the existence of a steady state solution to (2.46) is that $[A_i - L_{ii}H_{i1}, H_{i3}]$ be an observable pair.

In order to collect and summarize the main results of this chapter the following theorem is presented.

+Somewhat weaker sufficient conditions have been established. In this regard see [58-60].

THEOREM 2.1: Suppose that the following assumptions are true

(A1): SYSTEM MODEL: The system is modeled by (2.1)-(2.3) and for each subsystem $q_i \geq p_i$ and $\bar{H}_{i2}$ has full rank, p_i.

(A2): INITIAL STATE STATISTICS[+]: For each $i=1,2,\ldots,N$ $x_i(t_o,\cdot)$ is gaussian with mean x_i^o and covariance Σ_i^o and for $i \neq j$ $x_i(t_o,\cdot) \perp x_j(t_o,\cdot)$, and $x(t_o) \perp [w' \ \eta']' \ (t \ )$.

(A3): NOISE STATISTICS: For each $i=1,2,\ldots,N$ η_i and w_i are zero mean white gaussian processes with covariance matrices N_i and W_i respectively and for each $t \in [t_o,t_f]$, $w_i(t,\cdot) \perp \eta_i(t,\cdot)$, and $\eta_i(t,\cdot) \perp \eta_j(t,\cdot)$, $w_i(t,\cdot) \perp w_j(t,\cdot)$ for all $i \neq j$.

(A4): COST FUNCTIONAL: The filter performance is measured by

$$J(t) = \mathscr{E}\{||x(t,\cdot) - x(t,\cdot)||_Q^2\}$$

wherein $Q = Q_1 \dotplus Q_2 \dotplus \ldots \dotplus Q_N$ with each Q_i positive definite.

Then

(a) The class of surely locally unbiased filters is nonempty and the optimum element in this class may be constructed by the implementation of steps 1-6 of the design procedure on page 38.

(b) After steps 1-3 have been implemented, the filter performance may be assessed by solving the Riccati equation (2.46).

(c) If in addition to (A1)-(A4) the system is time invariant and when the system is reduced to the canonical form (via steps 1-3) the resulting pair $[A_i - L_{ii}H_{i1}, H_{i3}]$ is observable, then there

---

[+] For brevity, $x \perp y$ is used to denote that the random variables x and y are stochastically independent.

exists a steady state value for the error covariance and in this case the performance degradation due to decentralization approaches a bounded steady state value.

V. Extensions of the Filter

As in the case of the centralized information pattern, the filter derived in this chapter may be extended in at least two directions. In the case of nonlinear systems one can utilize the decentralized filter by linearizing about the best estimate of the local state.

The other extension involves systems for which in addition to the white noise disturbances there are "semi-stochastic" disturbance inputs present which represent additional uncertainty. Typical of this class is the one which can be modeled as the solution of an unforced system subject to random initial conditions. That is, disturbances of the form

$$\dot{\gamma}_i = \Theta_i \gamma_i \tag{2.48}$$

with $\gamma_i(t_o)$ a gaussian random vector with mean γ_i^o and covariance $\Sigma_{\gamma_i}^o$. Consider an extension of the model of Chapter I of the form

$$\dot{x}_i = A_i x_i + L_{ii} u_i + B_i m_i + \Gamma_i \gamma_i + w_i \tag{2.49}$$

with γ_i given by (2.48) and for all $i,j=1,2,\dots,N$ $\gamma_i(t_o) \perp w_j(t)$, $\gamma_i(t_o) \perp n_j(t)$, $\gamma_i(t_o) \perp x_j(t_o)$, and for $i \neq j$ $\gamma_i(t_o) \perp \gamma_j(t_o)$.

In most cases of practical interest the disturbance, γ_i, can not be measured and hence the observation equation remains as

$$y_i = \bar{H}_{i1} x_i + \bar{H}_{i2} u_i + n_i \tag{2.50}$$

The decentralized filter can be extended to handle this case. Furthermore, certain structural and computational simplifications which are important from an implementation standpoint can be shown to exist. In particular, one can show (see Appendix E) that the filter can be reduced to the form shown in Figure 2.2 in which

$$G^o_{i2} = (\tilde{V}_{i1}H'_{i3} - L_{ii}N_{i3})N^{-1}_{i2}$$

$$\alpha_i = \Theta_i - V'_{i3}H'_{i3}N^{-1}_{i2}H_{i3}S_i$$

$$V_{i3} = Q_{x_i}M_iQ'_{\gamma_i}$$

$$\dot{Q}_{\gamma_i} = \Theta_iQ_{\gamma_i} \qquad Q_{\gamma_i}(t_o) = I$$

$$\dot{Q}_{x_i} = (A^*_i - \tilde{V}_{i1}H'_{i3}N^{-1}_{i2}H_{i3})Q_{x_i} + \Gamma_iQ_{\gamma_i} \;, \quad Q_{x_i}(t_o) = 0$$

$$\dot{M}_i = -M_iQ'_{x_i}H'_{i3}N^{-1}_{i2}H_{i3}Q_{x_i}M_i \;, \quad M_i(t_o) = \Sigma^o_{\gamma_i}$$

$$S_i = Q_{x_i}Q^{-1}_{\gamma_i}$$

with $\tilde{V}_{i1}$ and A^*_i given in equation (2.46). Note that the upper section in Figure 2.2 is the local filter for the case in which there are no γ_i disturbances.

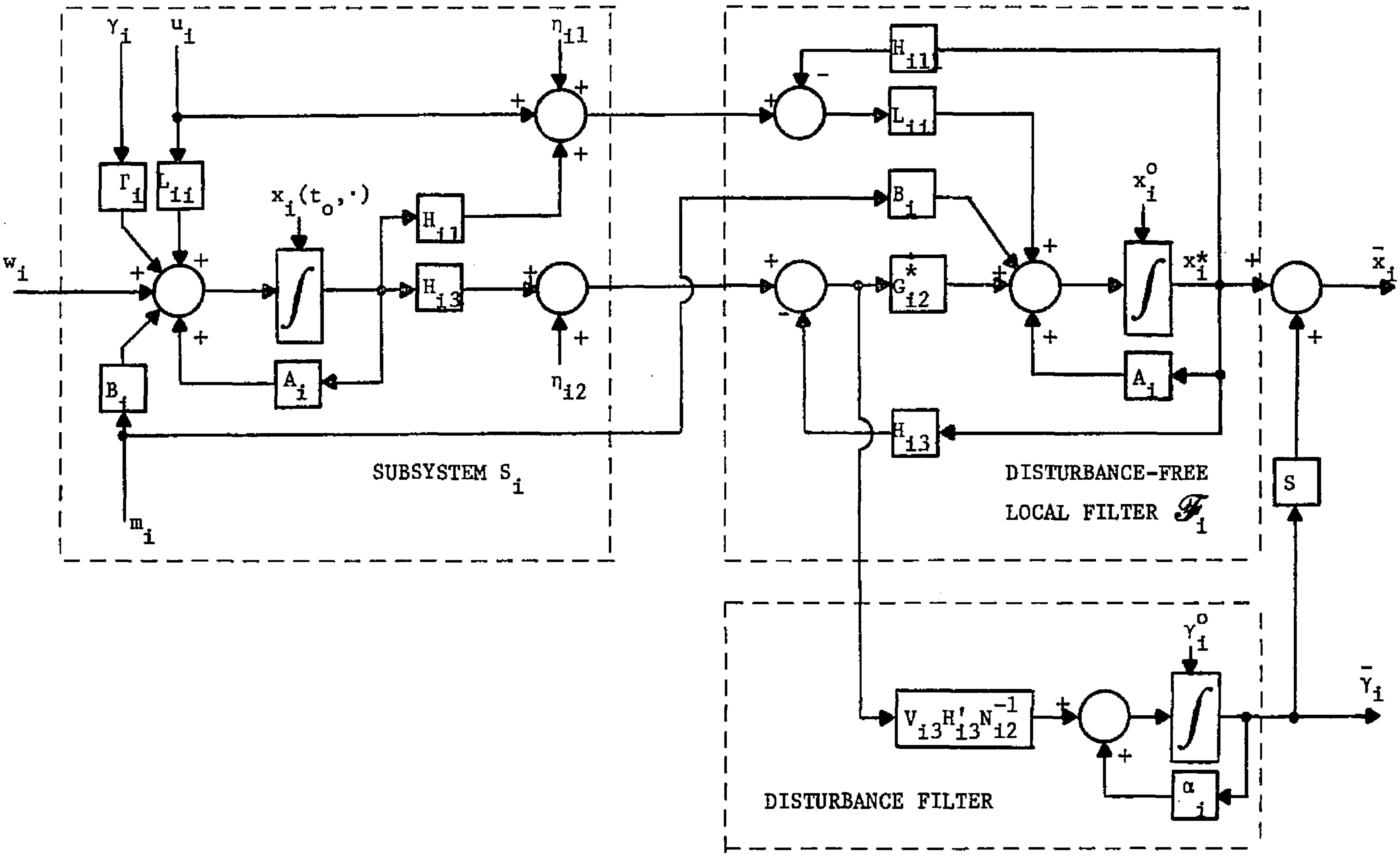

Fig. 2.2 Subsystem S_i and its local disturbance filter.

CHAPTER III

DECENTRALIZED CONTROL VIA CONSTRAINED CONTROLLERS

I. Introduction

As noted previously, one of the fundamental problems of system science is that of controlling a given system in such a way that it gives acceptable performance even in the presence of unmeasurable and unpredictable disturbances. A lunar spacecraft's booster stage is required to inject the spacecraft into a precise orbit around the Earth even in the presence of random wind gusts and other meteorological phenomena. Similarly, an electrical utility is expected to meet its load in the presence of an essentially random load demand.

In this chapter attention is focused on the problem of decentralized control of large-scale interconnected systems subject to a single quadratic performance functional. Thus, the problem falls in the area of team theory. The system structure is the same as that considered earlier in which the overall system consists of a collection of N interacting dynamical systems each having its own local controller which has a subset of the total system information set from which to determine its control signal.

For completeness and in order to compare the variational derivations of the centralized and decentralized controllers, the chapter begins with a brief derivation of the centralized controller via the variational approach. A nonlinear two point boundary value problem is derived which specifies the optimal system parameters, and it is shown that the optimal parameters can be obtained by solving two single point boundary value problems. That is, a form of the separation theorem is developed

using variational techniques. Following this discussion, the decentralized case is considered by means of the same technique and it is shown that the optimal controller within the chosen implementable class can again be obtained through the solution of a two point boundary value problem. The approach taken is similar to earlier work of Chong and Athans [61]. However in [61] the detailed structure of the overall system was not considered and in addition the complexity constraints did not play the role that they do in this paper.

The necessary conditions which must be satisfied by the optimal system parameters are utilized to investigate the applicability of the separation principle in the case of decentralized information. Finally a suboptimal controller based on the decentralized filter discussed previously and the decentralized output feedback controller of Appendix B is proposed and discussed.

II. The Centralized Controller

In this section consideration is given to the classical stochastic control problem in which the plant is linear, the cost quadratic, the noise processes gaussian, and the information centralized. Thus, combining the subsystem dynamics as before there result the equations describing the composite dynamics and observations.

$$\dot{x} = Ax + Bm + w \tag{3.1}$$

$$y = Hx + \eta \tag{3.2}$$

wherein A is given in (2.5), H in (2.6), and B is defined by

$$B \triangleq \sum_{i=1}^{N} P_{x,i}' \, B_i \, P_{M,i} \quad .$$

The controller performance is to be judged via a performance functional of the form

$$J = \mathscr{E}\{||x(t_f)||^2_{Q_f} + \int_{t_o}^{t_f} ||x(t)||^2_Q + ||m(t)||^2_M \, dt\} \qquad (3.3)$$

wherein Q_f, Q, and M are given matrices with Q_f and Q positive semidefinite and M positive definite. The stochastic optimal control problem is to find an element in the class of implementable control laws which minimzes the performance functional, J.

The class of implementable control laws is chosen to be the class consisting of an n-th order dynamic estimator followed by a memoryless linear transformation. Thus, using previous results on unbiased estimators, it follows that the class of implementable controllers is characterized by the equations

$$\begin{aligned} m(t) &= D\,\hat{x}\,(t) \\ \dot{\hat{x}} &= A\hat{x} + G\{y-H\hat{x}\} + Bm(t) \\ \hat{x}(t_o) &= x^o \end{aligned} \qquad (3.4)$$

wherein D and G are continuous real-valued matrices on $[t_o, t_f]$. These matrices are to be chosen in an optimal manner relative to the performance functional defined in (3.3). A schematic representation of this structure is given in Figure 3.1.

To reformulate the above stochastic optimization problem on the n-dimensional state space into a deterministic optimization on the finite dimensional state space $R^{2n} \times R^{2n}$, it is useful to define the state vector $v' \triangleq [x' \; \hat{x}']$. From (3.1), (3.2), and (3.4) it follows that this vector evolves according to the stochastic differential equation

$$\dot{v} = \bar{A}v + \bar{w} \qquad (3.5)$$

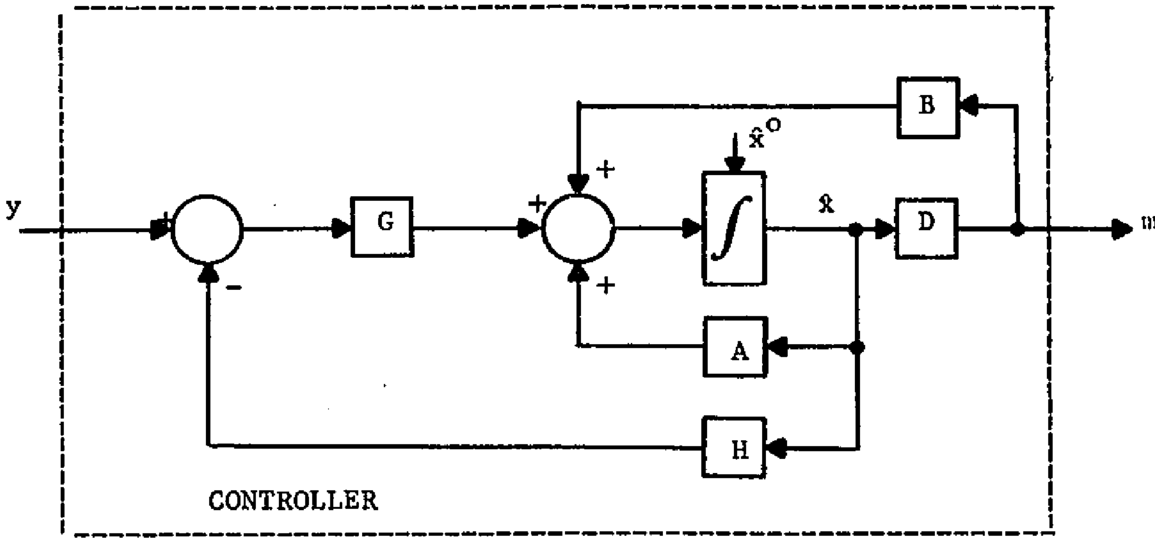

Fig. 3.1 Structure of the constrained controller.

with

$$\bar{A} \triangleq \left[\begin{array}{c|c} A & BD \\ \hline GH & A+BD-GH \end{array}\right]$$

and

$$\bar{w} = \left[\begin{array}{c} w \\ \hline G\eta \end{array}\right]$$

For notational purposes it is convenient to utilize the projections P_1 and P_2 from $R^n \times R^n$ to R^n defined by

$$P_1(x) = x_1 \ , \quad P_2(x) = x_2$$

where $x = [x_1' \ \ x_2']' \in R^n \times R^n$. It then follows that the cost functional of (3.3) may be written as

$$J = \mathcal{E}\{v'(t_f)\ \bar{Q}_f v(t_f) + \int_{t_o}^{t_f} v'(t)\bar{Q}v(t)dt\} \tag{3.6}$$

wherein

$$\bar{Q}_f = P_1'Q_fP_1$$

and $\qquad \bar{Q} = P_1'QP_1 + P_2'D'MDP_2$. Finally, letting the 2n×2n matrix V(t) be defined for each $t \in [t_o, t_f]$ and each (G,D) by

$$V(t) = \mathscr{E}\{v(t)\ v'(t)\} \tag{3.7}$$

it follows that the cost can be written

$$J = \mathrm{tr}\{\bar{Q}_f V(t_f) + \int_{t_o}^{t_f} \bar{Q}V(t)\ dt\} \tag{3.8}$$

and that the dynamical equation having V as its solution is

$$\dot{V} = \bar{A}V + V\bar{A}' + \bar{W}$$

where

$$\bar{W} = \left[\begin{array}{c|c} W & 0 \\ \hline 0 & G\bar{N}G' \end{array}\right]$$

and

$$V(t_o) = \left[\begin{array}{c|c} \sum^o + x^o x^{o'} & x^o x^{o'} \\ \hline x^o x^{o'} & x^o x^{o'} \end{array}\right] \triangleq V^o$$

Thus (3.8) and (3.9) can be regarded as a reformulation of the original problem into the deterministic optimization problem

$\mathscr{P}$: Minimize J as given by (3.8) with $\bar{Q}_f$ and $\bar{Q}$ as given in (3.6) subject to the dynamical constraint (3.9).

Necessary conditions for the optimal G and D gain matrices can be obtained via the matrix minimum principle. In the application of this principle it is useful to note that $\bar{A}$ and $\bar{W}$ can be written

$$\bar{A} = \bar{A}_o + (P_1+P_2)' BDP_2 + P_2'GH(P_1-P_2) \tag{3.10}$$

and

$$\bar{W} = \bar{W}_o + P_2'G\bar{N}G'P_2 \tag{3.11}$$

wherein $\bar{A}_o$ and $\bar{W}_o$ are independent of the gain matrices, G and D, and the P_i are the matrix representations of the projections defined above.

To apply the minimum principle to the control problem, $\mathscr{P}$, let the 2n×2n matrix, Z, represent the costate and ψ the Hamiltonian for $\mathscr{P}$. Thus,

$$\begin{aligned}\psi &= tr\{\dot{V}Z' + \bar{Q}V\}\\ &= tr\{(\bar{A}V + V\bar{A}' + \bar{W})Z' + \bar{Q}V\}\end{aligned} \tag{3.12}$$

$$\dot{Z} = -\frac{\partial\psi}{\partial V},\ Z(t_f) = \frac{\partial}{\partial V(t_f)}tr\{\bar{Q}_f V(t_f)\} \tag{3.13}$$

$$\dot{V} = \frac{\partial\psi}{\partial Z},\quad V(t_o) = V^o \tag{3.14}$$

and

$$\left.\frac{\partial\psi}{\partial G}\right|_{G=G*} = 0 \qquad\qquad \left.\frac{\partial\psi}{\partial D}\right|_{D=D*} = 0 \tag{3.15}$$

First consider the differential equation specifying the costate. From (3.12) and the gradient matrix results [51] it follows that

$$\frac{\partial\psi}{\partial V} = \bar{Q} + \bar{A}'Z + Z\bar{A}$$

and $\frac{\partial}{\partial V(t_f)}tr\{\bar{Q}_f V(t_r)\} = \bar{Q}_f$. Therefore from (3.13) it follows that the costate is given as the solution to the matrix differential equation

$$\begin{aligned}\dot{Z} &= -\bar{A}Z' - Z\bar{A} - \bar{A}\\ Z(t_f) &= \bar{Q}_f\end{aligned} \tag{3.16}$$

Clearly, for each $\bar{A}$ the solution of (3.16) is a symmetric matrix since if Z is a solution to (3.16) then so is Z'.

Now using (3.12), the definitions of $\bar{Q}$, $\bar{A}$, $\bar{W}$, and the gradient matrix results [51] it is readily shown that

$$\frac{\partial\psi}{\partial D} = 2B'(P_1+P_2)ZVP_2' + 2MDP_2VP_2'$$

$$\frac{\partial\psi}{\partial G} = 2P_2ZP_2'G\bar{N} + 2P_2ZV(P_1-P_2)'H'.$$

Thus, the conditions for the optimal G and D can be written

$$D(P_2VP_2') = -M^{-1}B'(P_1+P_2)ZVP_2' \quad (3.17)$$

$$(P_2ZP_2')G = P_2ZV\ (P_2-P_1)'H'N^{-1} \quad (3.18)$$

wherein V and Z are solutions to

$$\dot{V} = \bar{A}V + V\bar{A}' + \bar{W},\ \ V(t_o) = V^o \quad (3.9)$$

$$\dot{Z} = -Z\bar{A} - \bar{A}'Z - \bar{Q},\ Z(t_o) = \bar{Q}_f \quad (3.16)$$

Equations (3.9) and (3.16-3.18) consitute a two-point boundary value problem which when solved yields the optimal parameters, G^* and D^*.

III. A Separation Principle

In general the solution of multidimensional nonlinear boundary value problems of the above type is a difficult computational problem. However, for the above problem it is possible to determine the optimal parameters by solving two single point boundary value problems of the Riccati type. This, of course, is the essence of the separation principle of stochastic control theory. This principle has been studied by a number of researchers [35-40] and is included here primarily for completeness and to indicate an avenue of approach for the decentralized case. The derivation given is simple, straight-forward rigorous, and

to the author's knowledge has not been published in the literature. The result is limited, of course, by the fact that the controller is restricted a priori to be linear and hence the fact that the controller obtained is optimal over a larger class cannot be claimed using this development. The derivation is essentially one of verifying that a certain pair satisfies the above necessary condition.

To this end consider the pair (G_o, D_o) defined by the following

$$G_o = K_o H' \bar{N}^{-1} \tag{3.19a}$$

with K_o given as the solution of the Riccati equation

$$\dot{K}_o = AK_o + K_o A' - K_o H' \bar{N}^{-1} HK_o + W \qquad , K_o(t_o) = \textstyle\sum^o \tag{3.19b}$$

and

$$D_o = -M^{-1} B' \theta_o \tag{3.20a}$$

with θ_o given as the solution to

$$\dot{\theta}_o = -A' \theta_o - \theta_o A + \theta_o BM^{-1} B' \theta_o - Q \tag{3.20b}$$

$$\theta_o(t_f) = Q_f$$

Note that G_o and D_o are continuous.

It will now be shown that the pair (G_o, D_o) satisfies the necessary conditions (3.9) and (3.16)-(3.18). To aid in the discussion it is useful to define the sets

$S_1 \triangleq \{V\colon (G,D) \text{ continuous and } V \text{ satisfies } (3.9)\}$

$S_2 \triangleq \{Z\colon (G,D) \text{ continuous and } Z \text{ satisfies } (3.16)\}$

and the maps Γ_1, and Γ_2 taking $2n{\times}2n$ matrices into $n{\times}n$ matrices and defined by the rules

$$\Gamma_1(V) = V_{11} - V_{12} - V_{21} + V_{22}$$

$$\Gamma_2(V) = V_{11} + V_{12} + V_{21} + V_{22}$$

where

$$V = \left[\begin{array}{c|c} V_{11} & V_{12} \\ \hline V_{21} & V_{22} \end{array}\right]$$

Using the above nomenclature we have the

<u>Lemma 1</u>:

(a) If $V \in S_1$ then $\Gamma_1(V)$ is a solution to

$$\Gamma_1(V)(t_o) = \Sigma^o$$

$$\dot{\Gamma}_1(V) = (A - GH)\Gamma_1(V) + \Gamma_1(V)(A-GH)' + W + G\bar{N}G'$$

(b) If $Z \in S_2$ then $\Gamma_2(Z)$ is a solution to

$$\Gamma_2(Z)(t_f) = Q_f$$

$$\dot{\Gamma}_2(Z) = -\Gamma_2(Z)(A+BD) - (A+BD)'\,\Gamma_2(Z) - Q - DMD'$$

Proof: Part (a) follows immediately by writing (3.9) in detail and using the definition of Γ_1. Part (b) follows in a similar fashion using (3.16) and the definition of Γ_2.

The essence of the above lemma is that the composition map of Γ_1 with the solution operator of (3.9) is invariant with D and the composition map of Γ_2 with the solution operator of (3.16) is invariant with G. Now use the pair (G_o, D_o) defined in (3.19) and (3.20) to define the "sections" of S_1 and S_2 by

$$S_1(G_o) \triangleq \{V \in S_1 : G = G_o\}$$

$$S_2(D_o) \triangleq \{Z\varepsilon S_2 : D = D_o\}$$

and

$$\bar{S}_1(G_o) = \{V\varepsilon S_1(G_o) : \Gamma_1(V) = K_o\}$$

$$\bar{S}_2(D_o) = \{Z\varepsilon S_2(D_o) : \Gamma_2(Z) = \theta_o\}$$

From the above definitions it follows that

<u>Lemma 2</u>:

(a) $\bar{S}_1(G_o) = S_1(G_o)$

(b) $\bar{S}_2(D_o) = S_2(D_o)$

Proof:

(a) Clearly $\bar{S}_1(G_o) \subset S_1(G_o)$ and therefore assume that $V\varepsilon S_1(G_o)$. Then V satisfies (3.9) with $G = G_o$ and some continuous* D. Therefore by lemma 1, $\Gamma_1(V)$ satisfies

$$\Gamma_1(V)(t_o) = \Sigma^o$$

$$\dot{\Gamma}_1(V) = (A-G_oH)\Gamma_1(V) + P_1(V)(A-G_oH)' + W + G_o\bar{N}G_o'$$

and using the definition of G_o we have $(\Gamma_1(V) = \Gamma_1)$

$$\dot{\Gamma}_1 = A\Gamma_1 + \Gamma_1 A + W - K_oH'\bar{N}^{-1}H\Gamma_1 - \Gamma_1 H'\bar{N}^{-1}HK_o$$
$$+ K_oH'\bar{N}^{-1}HK_o$$

From the definition of K_o(3.19b) it follows that $\Gamma_1 - K_o$ satisfies

$$(\dot{\Gamma}_1 - \dot{K}_o) = A(\Gamma_1-K_o) + (\Gamma_1-K_o)A'$$
$$+ (K_o-\Gamma_1)H'\bar{N}^{-1}HK_o + K_oH'\bar{N}^{-1}H(K_o-\Gamma_1)$$

*For brevity, in the sequel we denote "D is continuous" by $D\varepsilon\mathcal{D}$, and "G is continuous" by $G\varepsilon\mathcal{G}$.

with the initial condition $(\Gamma_1-K_o)(t_o) = 0$. Hence $(\Gamma_1-K_o)(t) = 0$ for all $t \in [t_o,t_f]$ and thus $V \in \bar{S}_1(G_o)$. This establishes (a). Part (b) follows in a similar fashion.

From lemma 2 and the definitions of G_o and D_o we have the following useful

Lemma 3:

(a) If $(G,D) \in \{G_o\} \times \mathcal{D}$, then $G_o = \Gamma_1(V(D))H'\bar{N}^{-1}$

(b) If $(G,D) \in \mathcal{G} \times \{D_o\}$, then $D_o = M^{-1}B'\Gamma_2(Z(G))$

Proof: Let $D \in \mathcal{D}$ and $V(D)$ the corresponding element in $S_1(G_o)$. By lemma 2, $\Gamma_1(V(D)) = K_o$ from which (a) follows by the definition of G_o. Part (b) follows from the corresponding (b) of lemma 2.

The above simple results can be used to establish an important result leading to the separation theorem. That is,

Lemma 4:

(a) If $(G,D) \in \{G_o\}\times\mathcal{D}$ then $V_{12} = V_{22}$

(b) If $(G,D) \in \mathcal{G}\times \{D_o\}$ then $Z_{12} = -Z_{22}$.

Proof:

(a) Clearly $V_{12}(t_o) = V_{22}(t_o)$. Writing (3.9) in detail, applying some algebra, and the definition of $\Gamma_1(V)$ yields

$$(V_{22}\dot{-}V_{12}) = (V_{22}-V_{12})(A+BD-G_oH)'+(A-G_oH)(V_{22}-V_{12})$$
$$+ (V_{22}-V_{12})(G_oH)' + G_o\bar{N}G_o' - \Gamma_1(V)H'G_o'$$

Now by lemma 3, $G_o = \Gamma_1(V)H'\bar{N}^{-1}$, and thus $G_o\bar{N}G_o' - \Gamma_1(V)H'G_o' = 0$, which establishes (a).

Part (b) follows in a similar fashion.

Now it is an easy computation to establish the fact that (G_o, D_o) satisfies the necessary conditions (3.9), (3.16)-(3.18). Choose $(G,D) = (G_o, D_o)$ and let V^o, Z^o be the corresponding solutions to (3.9) and (3.16). From lemma 4 it follows that

$$V^o_{12} = V^o_{22} \;, \quad Z^o_{12} = -Z^o_{22}$$

and

$$\Gamma_1(V^o) = V^o_{11} - V^o_{21}$$

$$\Gamma_2(Z^o) = Z^o_{11} + Z^o_{21}$$

Therefore from lemma 3

$$G_o = (V^o_{11} - V^o_{21})\, H'\bar{N}^{-1}$$

$$D_o = -M^{-1}B'(Z^o_{11} + Z^o_{21})$$

Now use the above results to write

$$\begin{aligned} -M^{-1}B'(P_1+P_2)Z^oV^oP_2' &= -M^{-1}B'(Z^o_{11}+Z^o_{21}+Z^o_{12}+Z^o_{22})\; V^o_{22} \\ &= -M^{-1}B'(Z^o_{11}+Z^o_{21})V^o_{22} \\ &= D_oV^o_{22} \end{aligned}$$

and

$$\begin{aligned} P_2Z^oV^o(P_2-P_1)'H'\bar{N}^{-1} &= Z^o_{22}(V^o_{11} - V^o_{12})H'\bar{N}^{-1} \\ &= Z^o_{22}\; G_o \end{aligned}$$

and hence (G_o, D_o) satisfies the necessary conditions for optimality (3.9), (3.16)-(3.19).

The importance of the above result lies in the separation principle which it provides. That is, in designing the controller-filter pair the controller section may be designed by solving (3.20b)

and the filter section by solving (3.19b). Note that the optimal controller gain, D_o, depends only on the parameters describing the deterministic aspects of the control problem and is independent of the statistical parameters describing the noise processes and the initial state uncertainty. As is well known the controller gain is that gain which solves the deterministic control problem relative to the cost functional formed by removing the expectation in (3.3). In a similar vein, the filter gain, G_o, is independent of the particular cost functional matrices and depends only on the statistical parameters describing the noise processes, the initial state uncertainty, and the plant dynamics. It can be noted from a comparison of (3.19) and (2.30-31) that the filter section is merely the filter section developed in Chapter II.

IV. A Decentralized Controller

If there are no complexity constraints imposed, the above controller is optimal-even when the class over which the optimization is performed is not restricted a priori. However, in a large number of cases there are overriding technical and economical considerations which preclude the use of the above centralized controller. Therefore, in this section, we investigate the design of controllers for large-scale systems subject to "natural" information flow and complexity constraints.

Specifically, it is assumed that the information available to each controller consists of the local parameter and on-line data. That is

$$I_{DM_i}(t) = I_{MD_i} \cup I_{PD_i} \cup I_{OD_i}(t)$$

The complexity constraints placed on the controller are that it consists

of a collection of N local filter-controller pairs wherein the filter section is allowed to be an n_i-dimensional dynamical system. In addition, the requirement is imposed that the estimates are surely locally unbiased. Furthermore, it is assumed throughout that the subsystem measurement matrices have been put into canonical form+. Finally, the local control input, m_i, to S_i is constrained to be a linear transformation of the locally estimated state. Thus, in effect, it is assumed that each DM_i has enough resources to implement the locally optimal controller when the interactions from other subsystems are known to be zero.

From the above constraints and the results of Chapter II, it follows that each local filter-controller pair has the form given in Figure 3.2.

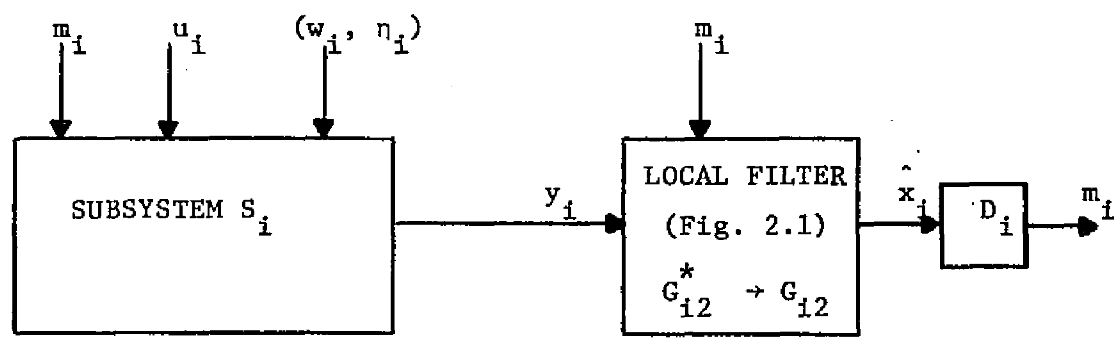

Fig. 3.2 Local filter-controller structure.

+A noted in Chapter II this assumption entails no further loss of generality after the $\bar{H}_{12}$ assumption has been made.

For the structure of subsystem S_i and the local filter section, consult Figure 2.1 and replace G_{i2}^* by G_{i2} -- that is we are not assuming that the filter gain may be designed independently of the controller gain, D_i.

Using the equations describing the subsystems and their local filter-controller pairs, the following equations describing the total system state can be derived.

$$\dot{x} = A_{11}x + A_{12}\hat{x} + w \tag{3.21a}$$

$$\dot{\hat{x}} = A_{21}x + A_{22}\hat{x} + \bar{\eta} \tag{3.21b}$$

where

$$A_{11} \triangleq \sum_{i=1}^{N} P'_{x,i}(A_i P_{x,i} + L_{ii}L_i) \tag{3.21c}$$

$$A_{12} \triangleq \sum_{i=1}^{N} P'_{x,i} B_i D_i P_{x,i} \tag{3.21d}$$

$$A_{21} \triangleq \sum_{i=1}^{N} P'_{x,i}(L_{ii}L_i + L_{ii}H_{i1}P_{x,i} + G_{i2}H_{i3}P_{x,i}) \tag{3.21e}$$

$$A_{22} \triangleq \sum_{i=1}^{N} P'_{x,i}(A_i - L_{ii}H_{i1} + B_i D_i - G_{i2}H_{i3})\, P_{x,i} \tag{3.21f}$$

$$\bar{\eta} \triangleq \sum_{i=1}^{N} P'_{x,i}(L_{ii}\eta_{i1} + G_{i2}\eta_{i2}) \tag{3.21g}$$

The parameters (G_{i2}, D_i) $i = 1,2,\ldots,N$ are to be chosen to minimize the cost functional of (3.3) wherein the cost matrices Q_f, Q, and M have the special structure

$$Q_f = Q_{f,1} \mathbin{\dot{+}} Q_{f,2} \mathbin{\dot{+}} \cdots + Q_{f,N}$$

$$Q = Q_1 \mathbin{\dot{+}} Q_2 \mathbin{\dot{+}} \cdots \mathbin{\dot{+}} Q_N$$

$$M = M_1 \mathbin{\dot{+}} M_2 \mathbin{\dot{+}} \cdots \mathbin{\dot{+}} M_N$$

Proceeding as in the centralized case it follows that the cost may be written

$$J = \mathrm{tr}\{\bar{Q}_f V(t_f) + \int_{t_o}^{t_f} \bar{Q}\, V(t)dt\} \tag{3.22}$$

with

$$\bar{Q}_f = P_1' Q_f P$$

$$\bar{Q} = P_1' Q P_1 + P_2' \sum_{i=1}^{N} P_{x,i}' D_i' M_i D_i P_{x,i} P_2 \ .$$

In (3.22) the 2n×2n matrix state, V, evolves according to

$$\dot{V} = \bar{A}V + V\bar{A}' + \bar{W} \tag{3.23}$$

with $V(t_o) = V^o$ (as in (3.9) and $\bar{A},\bar{W}$ given by

$$\bar{A} = \left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right], \quad \tilde{W} = \left[\begin{array}{c|c} W & 0 \\ \hline 0 & \bar{N} \end{array}\right]$$

and

$$\bar{N}(t) = \mathscr{E}\{\bar{\eta}(t)\bar{\eta}'(t)\}$$

$$= \sum_{i=1}^{N} P_{x,i}' \, (L_{ii}N_{ii}L_{ii}' + L_{ii}N_{i3}G_{i2}' + G_{i2}\, N_{i3}' L_{ii}'$$

$$+ G_{i2}\, N_{i2}\, G_{i2}')\, P_{x,i} \ .$$

Now equations (3.22) and (3.23) are of the same form as their counterparts, (3.8) and (3.9), in the centralized case. Thus the application of the minimum principle in the decentralized case will follow the same lines as for the centralized case. The results of this application are

$$\dot{V} = \bar{A}V + V\bar{A}' + \bar{W} \tag{3.25}$$

$$V(t_o) = V^o$$

$$\dot{Z} = -Z\bar{A} - \bar{A}'Z = \bar{Q} \tag{3.26}$$

$$Z(t_f) = \bar{Q}_f$$

$$D_i(P_{x,i}P_2VP_2'P_{x,i}') = -M_i^{-1}B_iP_{x,i}(P_1+P_2)ZVP_2'P_{x,i}' \tag{3.27}$$

$$(P_{x,i}P_2ZP_2'P_{x,i}')G_{i2} = P_{x,i}P_2ZV(P_2-P_1)'P_{x,i}'H_{i3}'N_{i2}^{-1} - P_{x,i}P_2ZP_2'P_{x,i}'L_{ii}N_{i3}N_{i2}^{-1} \tag{3.28}$$

Equations (3.25)-(3.28) consitute a two point boundary value problem whose solution leads to the optimal parameters, (G_{i2}^*, D_i^*) $i = 1,2,\ldots,N$. These equations are similar in form to the corresponding equations (3.9), (3.16)-(3.18) for the centralized case. However, in contrast to the centralized case, the computations inherent in the above equations apparently can not be separated. A logical conjecture in this regard would be that the optimal filter gains are those of the optimal decentralized filter of Chapter II. In an attempt to verify this conjecture using an argument similar to that of Section III, part (a) of lemmas 1-3 go through with only slight modifications. It is at lemma 4 that one encounters difficulty. For this choice of filter gains it is apparently not the case that $V_{12} = V_{22}$.

In the absence of a separation principle for the above problem, one is faced with the difficulty of solving (3.25) - (3.28) via an iterative technique such as the successive sweep method [62]. That

is, after an initial approximation+, (G_{12}^o, D_i^o) has been obtained, equations (3.25) and (3.26) are integrated to obtain V^o, Z^o. From V^o and Z^o one then uses (3.27) and (3.28) to compute the new approximation (G_{12}^1, D_i^1) and the process is repeated until a satisfactory approximation is obtained.

V. A Decentralized Controller for Time-Invariant Systems

The solution of the above two point boundary value problem represents a considerable computational effort and hence a controller based on this approach will, in general, be difficult to design. In the interest of obtaining controllers for which the computational requirements are more modest, one can combine the decentralized filter of Chapter II with the decentralized output feedback controller of Appendix B to obtain a local controller of the structural class indicated in Figure 3.2. This corresponds to setting G_{12} to the value computed using (2.45) and then designing the D_i gains using the technique of Appendix B. There are two basic approaches to the computation of the D_i.

Approach 1: The controller gains are computed using $D_i = - F_i^*$ wherein F_i^* is given by (B.31-B.33) after making the replacements

$$A \rightarrow \sum_{i=1}^{N} P_{x,i}' (A_i P_{x,i} + L_{ii} L_i)$$

+As an example of an initial approximation, one might choose the optimal decentralized controller based on the assumption that the subsystems are decoupled (i.e. assume $L_i = 0$, $i = 1,L,\ldots,N$). If the system is lightly coupled, then this initial approximation should be close to the optimal values. In addition this approximation is relatively easy to compute since it can be determined from N separate solutions of local optimizations.

$$B_i \rightarrow P_{x,i}' B_i$$

$$H_i \rightarrow P_{x,i}$$

$$X_o \rightarrow \Sigma^o + x^o x^{o'}$$

Approach 2: The controller gains are computed using $D_i = - F_i^*$ wherein the F_i^* is given by (B.31-B.33) after the following replacements are made

$$A^* \rightarrow \left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right] \tag{3.21 c-f}$$

$$B_i \rightarrow \left[\begin{array}{c} P_{x,i}' \; B_i \\ \hline P_{x,i}' \; B_i \end{array}\right]$$

$$H_i \rightarrow \left[\begin{array}{c|c} 0 & P_{x,i} \end{array}\right]$$

$$X_o \rightarrow \left[\begin{array}{c|c} \Sigma^o + x^o x^{o'} & x^o x^{o'} \\ \hline x^o x^{o'} & x^o x^{o'} \end{array}\right]$$

$$Q \rightarrow \left[\begin{array}{c|c} Q & 0 \\ \hline 0 & 0 \end{array}\right]$$

The first approach can be considered to be a result of applying the separation principle. That is, a filter is designed to give an estimate of the state and in the subsequent design of the controller gains it is assumed that the state is available--equivalently the filter dynamics are not taken into account in the design of the controller gains. Since the second approach does account for the filter dynamics, it is potentially more accurate but at the same time involves a larger computational effort since the effective order of the system has been increased from n to 2n.

The performance of the controllers designed using the above techniques can be evaluated using (3.22) and (3.23). The mean square performance of these controllers will, of course, be suboptimal. However, certain important advantages are present which should be considered. Among these are, (1) the filter gains and controller gains are computed separately with the only iterative part being the computation of the D_i gains, (2) the iterative process in computing the D_i involves the solution of algebraic equations rather than differential equations, (3) the resulting controller gains are constant, and (4) stochastic effects are accounted for in the filter design while the system coupling is accounted for in the computation of the controller gains. The application of the above design techniques to the development of a load frequency controller for a two-area power system is currently under study [63]. The results of this study will be reported elsewhere [63-64].

CHAPTER IV

DECENTRALIZED FILTERING AND CONTROL VIA INTERACTION MODELING

I. Introduction

The technique proposed in Chapter III provides a method for the centralized design of a decentralized controller. This approach has the advantage that once the class of implementable structures has been chosen the solution of the requisite nonlinear two-point boundary value problem leads to a controller which is optimal within the chosen class. Thus, once the implementable class has been sufficiently delineated, the design is (conceptually) straightforward. However, the techniques of Chapter III have certain disadvantages. Among these are: (1) the controller is decentralized only in the on-line phase, (2) efficient distribution of the computations required in the design of the local controllers is difficult to obtain, (3) the computational burden on the central agency is of considerable* magnitude, and (4) the only coordination between the decision makers is in the planning (gain computation) phase of controller implementation with no transfer of information allowed during the on-line phase.

The above disadvantages stem primarily from the nature of the coordination between the DM_i and the resulting task assigned to each DM_i. Namely, if the DM_i agree to allow the design to be carried out by a

*As a point in fact, the computational effort involved is larger than that required in designing the gain matrices for the centralized controller. This is not an unexpected result since more stringent constraints have been imposed on the controller in the decentralized case.

central agency, then they effectively relegate themselves to the task of merely implementing the controller gain matrices. Thus, while the approach given in Chapter III succeeds relative to the objective of distributing the on-line control effort, it does not achieve an efficient distribution of the decision making.

A computationally effective method of distributing the decision making effort is to assign a local control problem to each DM_i. However, due to the localized nature of the basic information pattern

$$I_{DM_i} = I_{PD_i} \cup I_{MD_i} \cup I_{OD_i}(t) \quad ,$$

each DM_i does not possess the necessary information with which to carry out the solution of a local control problem. The deficiency in the basic information pattern stems from the fact that it does not contain a characterization of the interaction process. This is representative of what Mesarovic [65] calls internal system uncertainty.

Now from a central vantage point it is clear that if all local controllers use linear control laws, then the interaction process is a colored gaussian process which is dependent on the local state. If the information available to DM_i is enlarged to include all the dynamics and control laws of the other subsystems then DM_i can characterize the interaction input to S_i and utilize centralized control theory in the design of the local controller. However, such an approach will lead to extremely complicated local controllers in addition to requiring an extensive communications complex to implement this enlarged information pattern. In addition, to implement such an approach some ordering must be placed on the times at which each DM makes a decision. For a rather

complete discussion of these team decision problems the interested reader should consult the work of Chu and Ho [66-67]. In this chapter we take a decidedly engineering approach and do not consider the team-decision-theoretic aspects of the design problem.

II. Interaction Modeling

In this chapter each DM_i is allowed to characterize his uncertainty regarding the interaction process by choosing a model from a particular class of interaction models. Two general classes of uncertain processes are (1) processes with a set membership constraint and (2) processes generated by stochastic differential equations. In (1) knowledge of the uncertain process consists only of the information that it lies in some known (usually convex) set and satisfies certain regularity conditions. The control of systems having set constrained disturbances has been recently reported by Glover and Schweppe[68] and Bertsekas and Rhodes [69].

In this chapter we consider interaction models from class (2) for which the information requirements lie somewhere between those of class (1) and the true stochastic interaction process. The models chosen are motivated by the "infinite-bus" concept which is often utilized in the analysis and design of load frequency controllers for interconnected electric power systems*. If one regards the frequency deviation of the other areas as the interaction variable, then the

*For a detailed account of the ramifications of this assumption in the design of a load frequency controller for the deterministic case the interested reader should see [70].

infinite-bus assumption is equivalent to choosing an interaction model of the form

$$\dot{u}_i(t) = 0$$

To reflect the fact that the deviation may be an unknown constant, the initial conditions for the above model can be assumed to be gaussian with appropriate mean and covariance.

Generalizing the above situation, we assume in this chapter that the interaction model consists of a differential equation with random initial conditions and possibly driven by white gaussian noise processes. Specifically, the interaction models are taken from the class of processes of the form

$$\dot{b}_i(t) = F_i b_i(t) + \xi_i$$

$$u_i(t) = T_i b_i(t)$$

with $b_i(t_o,\cdot)$ a gaussian random variable with mean b_i^o and covariance, Σ_i^o, and ξ_i a white gaussian noise process with zero mean and covariance, C_i. It is assumed that the noise process is stochastically independent of the initial state of S_i and the initial state of the interaction model.

Utilizing this model for the interactions, DM_i's image of the overall process becomes

$$S_i\colon \quad \dot{x}_i = A_i x_i + L_{ii} u_i + B_i m_i + w_i$$

$$\dot{b}_i = F_i b_i + \xi_i$$

$$u_i = T_i b_i$$

$$y_i = \bar{H}_{i1} x_i + \bar{H}_{i2} u_i + \eta_i$$

Taking the observation model into account and substituting into the above equations for u_i, DM_i's image of the overall process becomes

$$\tilde{S}_i: \quad \begin{aligned} \dot{x}_i &= A_i x_i + L_{i1} T_i b_i + B_i m_i + w_i \\ \dot{b}_i &= F_i b_i + \xi_i \\ y_i &= \bar{H}_{i1} x_i + \bar{H}_{i2} T_i b_i + n_i \end{aligned} \tag{4.1}$$

To further specify the local control problem for DM_i, we assume that the local cost functional is that part of the overall cost functional which depends explicitly on x_i. That is,

$$J_i = \mathscr{E}\{||x_i(t_f)||^2_{Q_{f,i}} + \int_{t_o}^{t_f} ||x_i(t)||^2_{Q_i} + ||m_i(t)||^2_{M_i} dt\} \tag{4.2}$$

Therefore, from the viewpoint of DM_i, the task is to determine a control law which minimizes J_i subject to the constraint $\tilde{S}_i$.

III. Local Controller Structure and Design Aspects

In this section attention is focused on the local control problem as viewed by DM_i through an interaction model of the form given in Section II. In addition to developing a design procedure for the local controller, it is shown that the local controller has a particularly simple structure.

Now, from the viewpoint of DM_i the problem as specified by (4.1) and (4.2) is a standard problem in centralized stochastic control theory. Thus the results of Chapter II imply that the separation principle is applicable to this problem,, and hence that the controller consists of a Kalman filter followed by the optimal controller for the deterministic system. In the following we investigate the detailed structure of each

of these subsystems of the local controller.

A. Controller Subsystem

The controller subsystem is the optimal controller for the deterministic system and is therefore the solution to the deterministic optimization problem

$$\mathscr{P}_c : \text{Minimize} \quad J_i = \; ||x_i(t_f)||^2_{Q_{f,i}} + \int_{t_o}^{t_f} ||x_i(t)||^2_{Q_i} \quad ||m_i(t)\,||^2_{M_i} dt \tag{4.3}$$

subject to

$$\dot{x}_i = A_i x_i + L_{ii} T_i b_i + B_i m_i \tag{4.4}$$

$$\dot{b}_i = F_i b_i \tag{4.5}$$

From the results of Section III , Chapter III, the solution is given by

$$m_i(t) = -M_i^{-1} \bar{B}_i' P \; \bar{x}_i(t) \tag{4.6}$$

with P the solution to the Riccati equation

$$\dot{P} = -\bar{A}_i' P - P\bar{A}_i + P\bar{B}_i M_i^{-1} B_i' P - \bar{Q}_i \tag{4.7}$$

$$P(t_f) = \bar{Q}_{f,i}$$

and

$$\bar{A}_i = \left[\begin{array}{c|c} A_i & L_{ii}T_i \\ \hline 0 & F_i \end{array}\right] \quad , \quad \bar{B}_i = \left[\begin{array}{c} B_i \\ \hline 0 \end{array}\right] \tag{4.8}$$

$$\bar{Q}_i = \left[\begin{array}{c|c} Q_i & 0 \\ \hline 0 & 0 \end{array}\right] \quad , \quad \bar{Q}_{f,i} = \left[\begin{array}{c|c} Q_{f,i} & 0 \\ \hline 0 & 0 \end{array}\right] \tag{4.9}$$

$$\bar{x}_i = [\; x_i' \quad b_i' \;]'$$

Now, because of the special structure of the matrices $\bar{A}_i$, $\bar{B}_i$, $\bar{Q}_i$, and $\bar{Q}_{f,i}$ in the above, the computations for P can be simplified considerably, [71]. Partition P to conform with the partitioning of $\bar{A}_i$ and note that the special structure of $\bar{B}_i$ leads to

$$m_i^* = -M^{-1}B_i'(P_{11}x_i + P_{12}b_i) \tag{4.10}$$

and thus P_{22} is not explicitly needed in the formation of the control law. Thus if P_{11} and P_{12} are not coupled to P_{22} in (4.7), then P_{22} need not be computed. Examining (4.7) in detail it is seen that P_{11} and P_{12} can be computed as solutions to the differential equations

$$\dot{P}_{11} = -A_i'P_{11} - P_{11}A_i + P_{11}B_iM_i^{-1}B_i'P_{11} - Q_i \tag{4.11}$$

$$\dot{P}_{12} = (P_{11}B_iM_i^{-1}B_i' - A_i')P_{12} - P_{12}F_i - P_{11}L_{ii}T_i \tag{4.12}$$

with $$P_{11}(t_f) = Q_{f,i} \quad , \quad P_{12}(t_f) = 0$$

It is interesting to note that the feedback gain, $-M_i^{-1}B_i'P_{11}$, which controls the amount of local state feedback is the same gain as would result if one neglected the interactions. Also note that P_{11} may be computed independently of P_{12} and then used in the computation of P_{12}. This simplifies the controller design considerably. Therefore the controller subsystem of the local controller has the structure shown in Figure 4.1 wherein P_{11} and P_{12} are obtained as the solutions of (4.11) and (4.12) and $[\ \hat{x}_i' \quad \hat{b}_i'\ ]'$ is the estimate provided by the filter to be discussed in the next subsection.

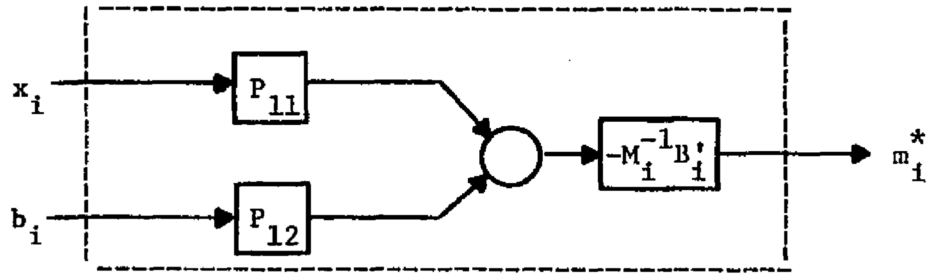

Fig. 4.1 Structure of the controller section.

B. Filter Subsystem

The filter subsystem consists of a Kalman filter for the stochastic system model

$$
\begin{aligned}
\dot{x}_i &= A_i x_i + L_{ii} T_i b_i + B_i m_i + w_i \\
\dot{b}_i &= F_i b_i + \xi_i \\
y_i &= \bar{H}_{i1} x_i + \bar{H}_{i2} T_i b_i + \eta_i
\end{aligned}
\tag{4.13}
$$

wherein m_i is regarded as a known deterministic signal.

First consider the case in which the driving noise, ξ_i, to the interaction model is zero. For this case it can be shown [72-73] that the optimal estimates can be generated by the stochastic differential equations

$$
\begin{aligned}
\dot{\tilde{x}}_i &= A_i \tilde{x}_i + G_i\{y_i - \bar{H}_{i1}\tilde{x}_i\} + B_i m_i \\
\dot{\hat{b}}_i &= \alpha_i \hat{b}_i + \beta_i\{y_i - \bar{H}_{i1}\tilde{x}\}
\end{aligned}
\tag{4.14}
$$

with $\quad \hat{x}_i = \tilde{x}_i + S\hat{b}_i$

and $\quad \tilde{x}_i(t_o) = x_i(t_o) \quad , \quad \hat{b}_i(t_o) = b_i^o \, .$

From (4.14) it can be seen that the optimal state estimates are generated by a system having the structure shown in Figure 4.2.

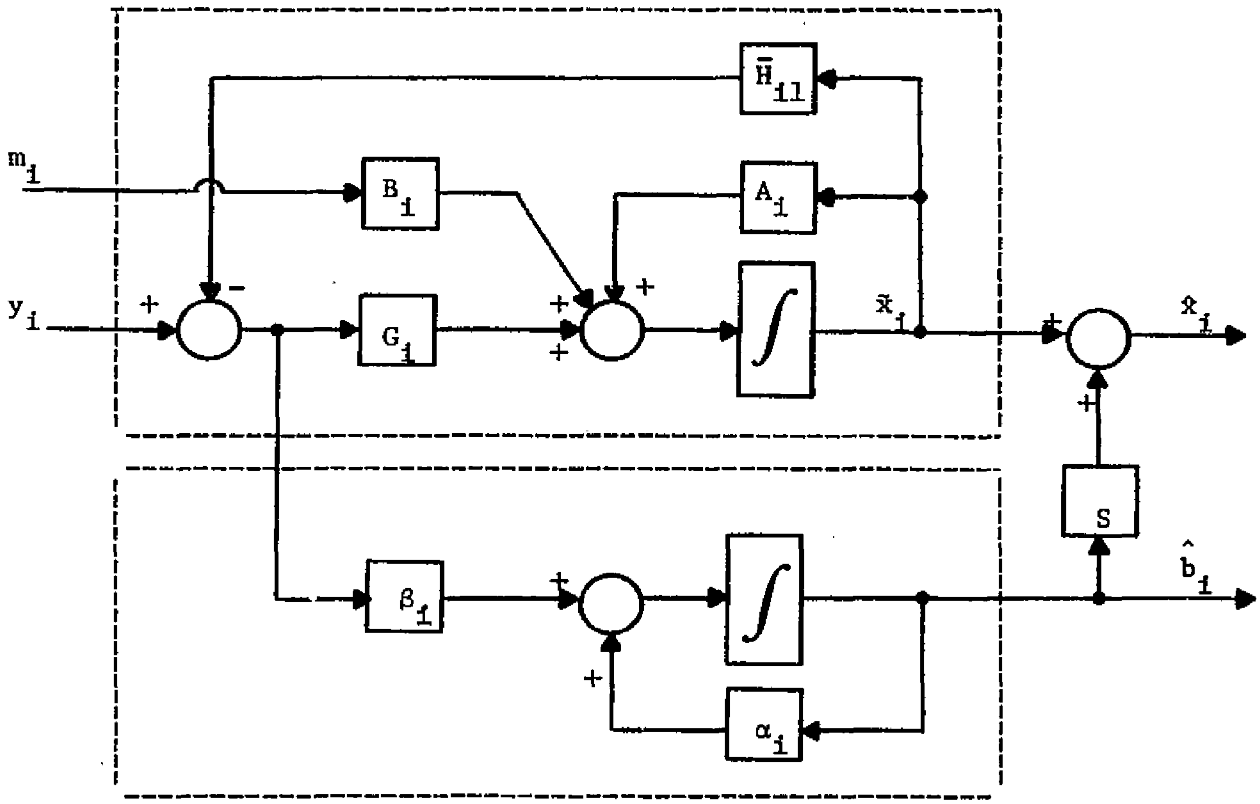

Fig. 4.2 Structure of the filter section.

The parameters in the above structure are given by [72-73]

$$G_i = P_x \bar{H}'_{i1} N_i^{-1} \qquad S = V_x V_b^{-1}$$

$$\beta_i = (P'_{xb} \bar{H}'_{i1} + P_b \bar{H}'_{i2}) N^{-1} \quad , \quad \alpha_i = F_i - \beta_i(\bar{H}_{i2} + \bar{H}_{i1} S) \tag{4.15}$$

with

$$\dot{P}_x = A_i P_x + P_x A'_i - P_x \bar{H}'_{i1} N_i^{-1} \bar{H}_{i1} P_x + W_i \tag{4.16}$$

$$\dot{V}_b = F_i V_b \tag{4.17}$$

$$\dot{V}_x = (A_i - P_x \bar{H}'_{i1} N_i^{-1} \bar{H}_{i1}) V_x + (L_{i1} T_i - P_x \bar{H}'_{i1} N^{-1} \bar{H}_{i1}) V_b \tag{4.18}$$

$$\dot{M} = -M(V'_x \bar{H}'_{i1} + V'_b \bar{H}'_{i2}) N_i^{-1} (\bar{H}_{i1} V_x + \bar{H}_{i2} V_b) M \tag{4.19}$$

$$P_{xb} = V_x M V'_b \quad , \quad P_b = V_b M V'_b \tag{4.20}$$

subject to the initial conditions

$$P_x(t_o) = \Sigma_i^o \quad , \qquad V_b(t_o) = I$$
$$V_x(t_o) = 0 \qquad\qquad M(t_o) = \Sigma^o_{b_i} \qquad (4.21)$$

Comparing the specification for G_i in (4.15) and P_x in (4.16) with equations (2.30) and (2.31) it should be noted that the upper subsystem in Figure 4.2 is the interaction-free filter for S_i. That is, it is the filter designed for S_i based on the assumption that u_i=0. In summary, for the case in which ξ_i=0, the local controller structure has the form shown in Figure 4.3. The design procedure for the subsystems in this figure is contained in equations (4.11), (4.12),and (4.15)-(4.21) and may be organized as follows.

DESIGN PROCEDURE

Step 1: Design the interaction-free section using (4.11), (4.16) with

$$P_x(t_o) = \Sigma_i^o \text{ , and } \quad G_i = P_x\bar{H}_{i1}'N_i^{-1}$$

Step 2: Determine an interaction model of the form

$$\dot{b}_i = F_i b_i \quad , \quad u_i = T_i b_i$$

with $b_i(t_o)$ gaussian with mean b_i^o and covariance $\Sigma^o_{b_i}$.

Step 3: Design interaction section by covering the computational lattice shown in Figure 4.4.

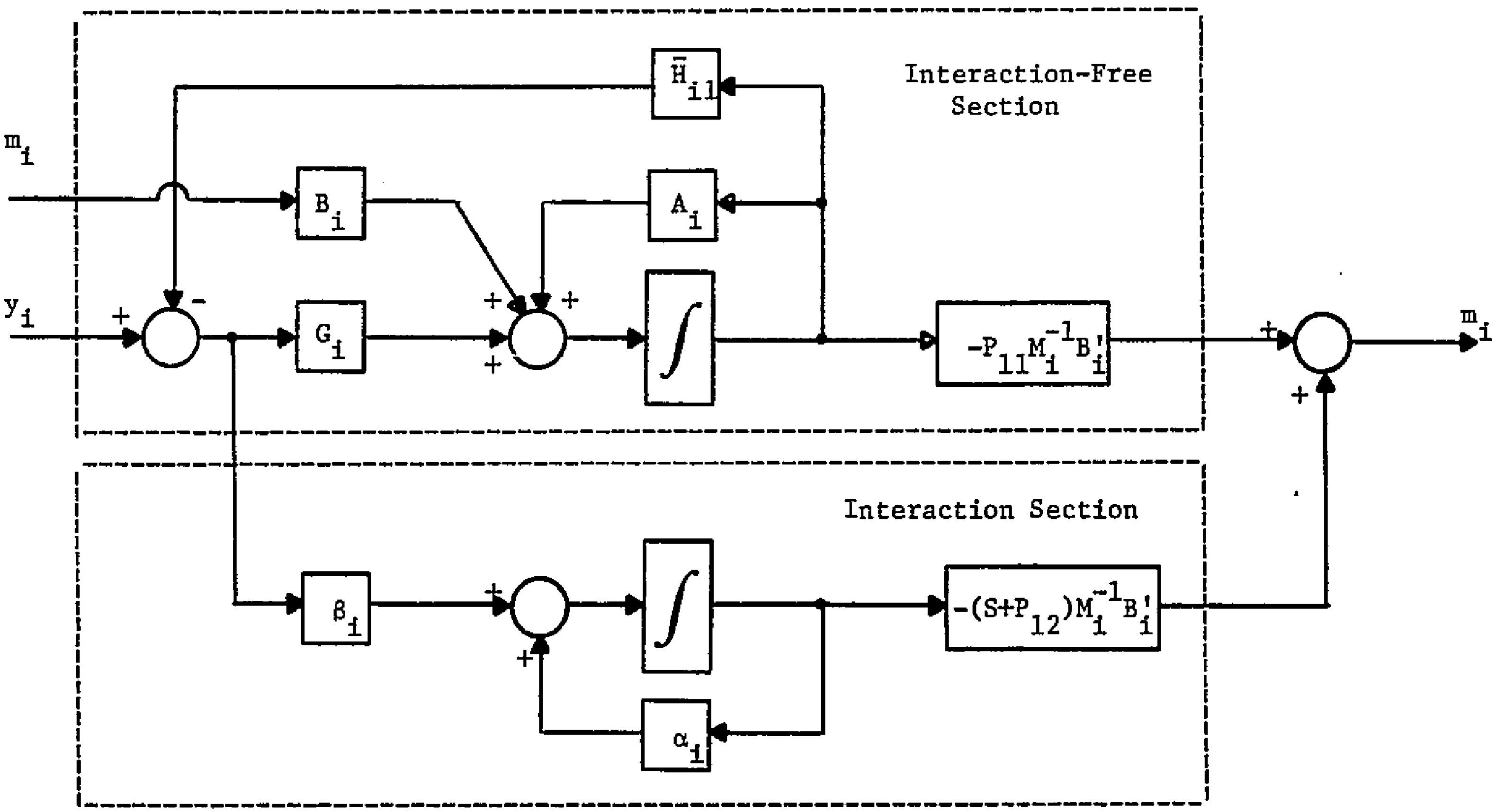

Fig. 4.3 Structure of the local controller.

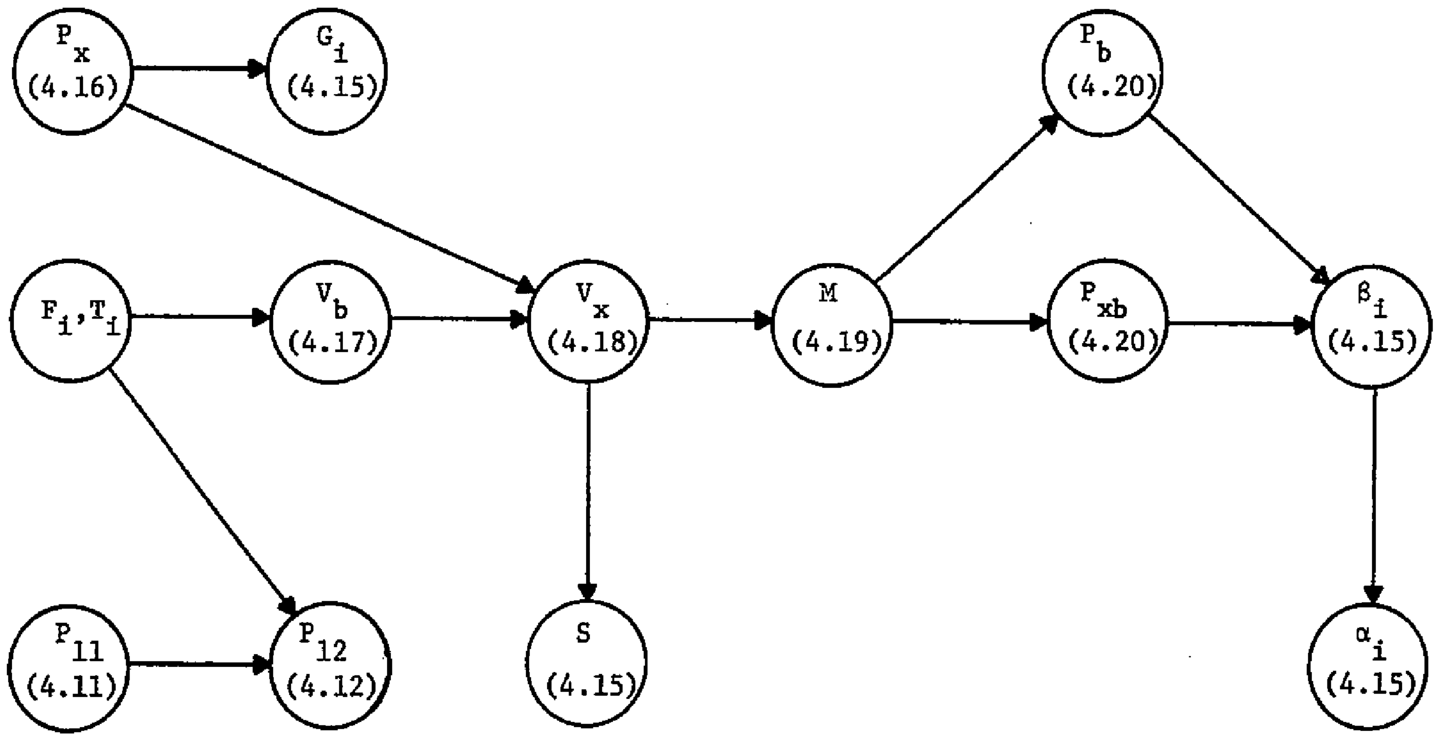

Fig. 4.4 Lattice of design computations.

In the case presented above it was assumed that the noise driving the interaction model is identically zero. It is possible to remove this restriction but apparently [72] at the expense of losing some of the simplicity of the filter structure. That is, if the noise, ξ_i, is not identically zero then apparently one cannot separate the filter into subsystems of the form shown in Figure 4.3--note that the two sections shown in this figure are not only dynamically separated but are also separated relative to the gain computations. However even for the case in which $\xi_i \neq 0$ it is possible, using the transformation discussed in Appendix C to obtain a structure having a form similar to that in Figure 4.3.

To treat the case $\xi_i \neq 0$, return to (4.13) and define

$$x = \left[\begin{array}{c} x_i \\ \hline b_i \end{array}\right], \quad A = \left[\begin{array}{c|c} A_i & L_{ii}T_i \\ \hline 0 & F_i \end{array}\right], \quad H = \left[\bar{H}_{i1} \ \bar{H}_{i2}\right]$$

$$w = \begin{bmatrix} w_i \\ \xi_i \end{bmatrix}, \quad \sum^o = \left[\begin{array}{c|c} \sum_i^o & 0 \\ \hline 0 & \sum_{b_i}^o \end{array}\right], \quad \eta = \eta_i$$

$$m = \left[\begin{array}{c} B_i m_i \\ \hline 0 \end{array}\right], \quad y = y_i$$

and note that (4.13) can be written as

$$\begin{aligned} \dot{x} &= Ax + m(t) + w \\ y &= Hx + \eta \end{aligned} \qquad (4.13')$$

Comparing (4.13') with (2.5) and (2.6) it is noted that the filter for the process described by (4.13) is given by (2.30), (2.31) and (2.13). Therefore the filter equations can be written

$$\dot{\hat{x}}_i = A_i\hat{x}_i + (L_{1i}-G_1\bar{H}_{12})T_i\hat{b}_i + G_1(y_i-\bar{H}_{11}\hat{x}_i) + B_i m_i \tag{4.22}$$

$$\dot{\hat{b}}_i = (F_i-G_2\bar{H}_{12}T_i)\hat{b}_i + G_2(y_i - H_i,\hat{x}_i)$$

Now define

$$\hat{A} = \left[\begin{array}{cc} A_i & (L_{1i}-G_1\bar{H}_{12})T_i \\ \hline 0 & F_i - G_2\bar{H}_{22}T_i \end{array}\right]$$

and make the $\hat{A}$-based transformation of Appendix C

$$v = T\,\left|\hat{x}_i' \;\; \hat{b}_i'\right|'$$

to obtain (see example in Appendix C)

$$\begin{aligned} \dot{v}_1 &= A_i v_1 + B_i m_i + (G_1 + T_{12}G_2)(y_i-\bar{H}_{11}(v_1-T_{12}v_2)) \\ \dot{v}_2 &= \alpha^o v_2 + G_2(y_i - \bar{H}_{11}(v_i-T_{12}v_2)) \end{aligned} \tag{4.23}$$

with

$$\alpha^o = F_i - G_2\bar{H}_{12}T_i$$

$$\hat{x}_i = v_1 - T_{12}v_2$$

$$\hat{b}_i = v_2$$

and

$$\begin{aligned} \dot{T}_{12} &= A_i T_{12} - T_{12}(F_i-G_2\bar{H}_{12}T_i) - (L_{1i}-G_1\bar{H}_{12})T_i \\ T_{12}(t_o) &= 0 \end{aligned} \tag{4.26}$$

The filter described by (4.23-4.26) has the structure shown in Figure 4.5.

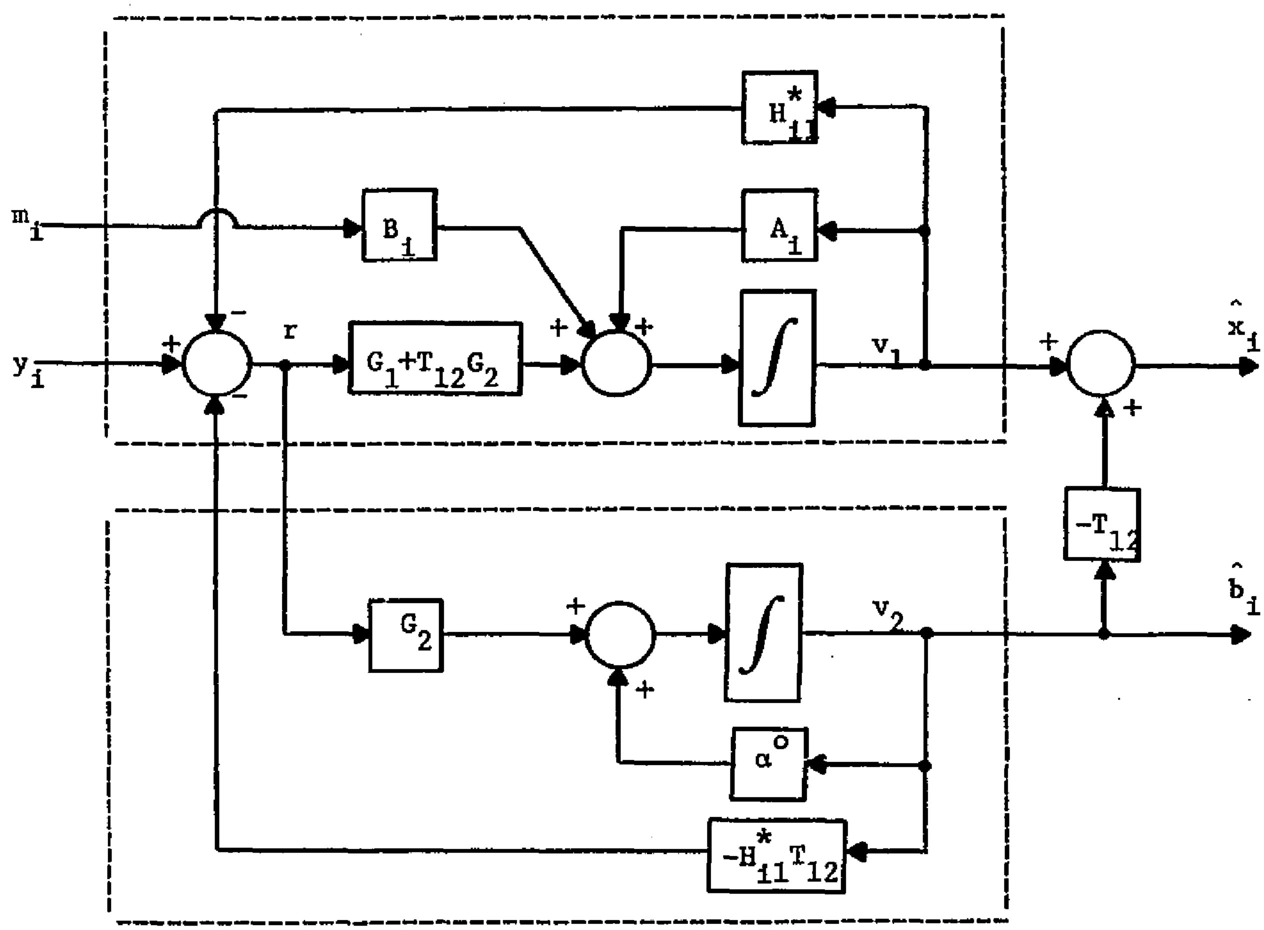

Fig. 4.5 A partially decoupled filter for $\xi_i \neq 0$.

Two basic differences should be noted between Figures 4.2 and 4.5: (1) the upper section of Figure 4.5 is not the interaction-free section of Figure 4.2 since the gain $G_1 + T_{12}G_2$ is not the interaction-free gain, and (2) the residuals (r) depend on both $\hat{x}_i$ and $\hat{b}_i$ in 4.5 whereas in 4.2 the residuals are independent of $\hat{b}_i$. Thus, in computing the residuals, the filter of Figure 4.5 requires a bidirectional channel between the two sections whereas the filter of Figure 4.2 requires only a unidirectional channel.

These additional communication requirements coupled with the fact that the gain computations require the solution of a higher dimensional

Riccati equation make the filter for the case $\xi_1 \neq 0$ more difficult to implement. However, there is a potential tradeoff present here which should be considered in the application of these results. Namely, if ξ_1 is assumed to be zero, the interaction section of the filter will eventually begin to ignore the residuals and may therefore diverge [74]. More specifically, note from (4.16) and the definition of β that if the interaction model is stable then $V_b \to 0$ as $t \to \infty$ and thus $\beta \to 0$ as $t \to \infty$. In contrast, if $\xi_1 \neq 0$, then the noise covariance driving the term W in the Riccati equation will cause the gain, G_2, in the interaction section to reach a steady-state* nonzero value and hence the filter will continue to "look at" the interactions. For systems in which the interaction model is of low dimension (e.g. infinite bus models in power systems), the increased complexity engendered by assuming $\xi_1 \neq 0$ will probably not be large and hence for this class of systems it will probably be best to include a noise driving term in the interaction model.

---

*Assuming that the system model is time invariant and observable.

CHAPTER V

CONCLUSIONS AND SUGGESTIONS FOR FURTHER RESEARCH

The research initiated in this paper has provided the following specific results:

(1) A formulation of the decentralized control and filtering problems for large-scale stochastic systems with particular emphasis on the information flow constraints which are often present in such systems has been developed.

(2) A derivation of the optimal decentralized filter within a certain class has been given together with a delineation of the conditions under which such a filter exists. The structural and design aspects of this filter have been interpreted in terms of the standard centralized Kalman filter. In particular it has been shown that subject to some computationally simple modifications, algorithms available for the design of the standard filter may be utilized in the decentralized case. A qualitative discussion of the filter performance has been given together with an observability condition which will insure that the error covariance is asymptotically stable. In addition to these qualitative results, quantitative performance data have been generated for the case of state estimation in a two-area power system.

(3) A simple and apparently new derivation of the separation theorem of centralized stochastic control theory has been provided. This result, which is of interest in its own right, has been used to

investigate the applicability of the separation principle for the class of controllers introduced in Chapter III.

(4) A modeling approach to decentralized controller design has been introduced via a particular class of interaction models. The structural and design aspects of the resulting localized controller have been investigated in considerable detail.

This author believes that the developing area of decentralized filtering and control structures offers a number of possibilities for further research. In relation to the work presented here some of these are:

(1) A quantitative study of the performance of the decentralized filter for particular system classes. A study of this type is currently in progress for the case of a two-area power system. Preliminary results from this study are given in Appendix D and will also be presented in [75]. Work is also in progress in investigating the performance of an extended version of this filter when applied to a nonlinear system [79].

(2) The assumption regarding the rank of $\bar{H}_{12}$ is somewhat restrictive and thus it would therefore be useful to study the case in which the surely locally unbiased class is empty. A study of the amount of improvement obtained by allowing filters that are outside the S.L.U. class would also be of interest.

(3) In the more difficult area of decentralized control a numerical study of the controllers based on the interaction modeling technique would certainly be a suitable project at the master's level. This study for the particular case of interconnected power systems

could use the results of Chapter IV to extend earlier work in [70] to the stochastic case.

(4) While the interaction modeling technique results in controllers having a simple structure, the large-scale/decision theoretic aspects of this approach have not been investigated in this paper. Relative to choosing the interaction models it would be of interest to determine conditions under which there exist optimal (in an appropriate sense) interaction models. For example, under what conditions does the class of interaction models admit a Nash, Pareto, etc. equilibrium.

(5) The control and filtering structures investigated in this research are very close to being completely decentralized. Thus an important aspect of decentralized structures has been omitted. Namely, we have not considered the possibility of allowing the local controllers to communicate with each other. It should be noted that any investigation of such structures must include some cost of communication in the performance functional. This aspect of such a study would probably increase the difficulty of this investigation by a considerable amount--one of the major problems appears to be the determination of a cost functional which accounts for the communication costs and at the same time is amenable to analytical work. The author is not aware of any previous research in this particular area of decentralized structures*.

---

*A possible exception to this is the very recent work of Aoki [20]. How- no communication costs are accounted for in this work.

REFERENCES

[1] K.J. Astrom, "Control Problems in Papermaking," Proc. IBM Scientific Computing Symposium Control Theory and Applications, pp. 135-167, Yorktown Heights, N.Y., 1964.

[2] R.E. Mortensen, "Mathematical Problems of Modeling Stochastic Nonlinear Dynamic Systems," J. of Statistical Physics, Vol. 1, No. 2, pp. 271-296, 1969.

[3] A. Blanc-Lapierre, and R. Fortet, Theory of Random Functions, Trans. by J. Gani, Gordon and Breach, New York, 1965.

[4] R.E. Kalman, P.L. Falb, M.A. Arbib, Topics in Mathematical System Theory, McGraw-Hill, New York, N.Y., 1969.

[5] D.A. Wismer (Ed.), Optimization Methods for Large-Scale Systems--with Applications, McGraw-Hill, New York, N.Y., 1971.

[6] P.V. Kokotovic, W.R. Perkins, J.B. Cruz, G. D'Ans, "ε-Coupling Method for Near Optimum Design of Large-Scale Linear Systems," Proc. IEE, Vol. 116, No. 5, pp. 889-892, 1969.

[7] J.D. Pearson, S. Reich, "Decomposition of Large Optimal Control Problems," Proc. IEE, Vol. 114, No. 6, pp. 845-851, 1967.

[8] M. Aoki, "Control of Large-Scale Dynamic Systems by Aggregation," IEEE Trans. Auto. Control, Vol. AC-13, No. 3, pp. 246-253, 1968.

[9] T. Marschak, "Centralization and Decentralization in Economic Organizations," Econometrica, Vol. 27, No. 3, pp. 399-430, 1959.

[10] H.J. Kruisinga (Ed.), The Balance Between Centralization and Decentralization in Managerial Control, H.E. Stenfert Kroese, Leiden, 1954.

[11] K.J. Arrow, "Optimization, Decentralization, and Internal Pricing in Business Firms," in Contributions to Scientific Research in Management, Univ. California Press, 1959.

[12] K.J. Arrow, L. Hurwicz, "Decentralization and Computation in Resource Allocation," in R.W. Pfouts (Ed.), Essays on Economics and Econometrics, Univ. North Carolina Press, 1960.

[13] R. Radner, "Team Decision Problems," Annals of Mathematical Statistics, Vol. 33, pp. 857-881, 1962.

[14] R. Radner, "The Evaluation of Information in Organizations," Proc. of the Fourth Berkeley Symposium on Probability and Mathematical Statistics, pp. 491-530, Univ. California Press, Berkeley, California, 1961.

[15] C.H. Kriebel, "Quadratic Teams, Information Economics, and Aggregate Planning Decisions," Econometrica, Vol. 36, No. 3-4, pp. 530-543, 1968.

[16] C.H. Kriebel, "Information Processing and Programmed Decision Systems," Management Science, Vol. 16, No. 3, pp. 149-164, 1969.

[17] J.D. Schoeffler, L.S. Lasdon, "Decentralized Plant Control," ISA Trans. Vol. 5, No. 2, pp. 175-183, 1966.

[18] R. Lau, R.C.M. Persiano, P.P. Varaiya, "Decentralized Information and Control: A Network Flow Example," IEEE Trans. Auto. Control, Vol. AC-17, No. 4, pp. 466-473, 1972.

[19] M. Aoki, "On Feedback Stabilizability of Decentralized Dynamic Systems," Automatica, Vol. 8, pp. 163-173, 1972.

[20] M. Aoki, "On Decentralized Linear Stochastic Control Problems with Quadratic Cost", IEEE Trans. Auto. Control, Vol. AC-18, No. 3, pp. 243-250, 1973.

[21] C.W. Sanders, Jr., E.C. Tacker, T.D. Linton, "Decentralized Stochastic Control of Large-Scale Systems via Constrained Controllers," Proc. 1973 SWIEEECO, pp. 1-3, 1973.

[22] R.J. Mouly, "Systems Engineering in the Glass Industry," IEEE Trans. System Science and Cybernetics, Vol. SSC-5, pp. 300-312, Oct. 1969.

[23] R.G. Kaenel, "Minicomputers--A Profile of Tomorrow's Component", IEEE Trans. Audio and Electroacoustics, Vol. AU-18, pp. 354-379, Dec. 1970.

[24] Y.C. Ho, "Differential Games, Dynamic Optimization and Generalized Control Theory," J. Optimization Theory and Applications, Vol. 6, No. 3, pp. 179-209, 1970.

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

[26] I.I. Gikhman, and A.V. Skorokhod, Introduction to the Theory of Random Processes, W.B. Saunders, Phila., Pa., 1969.

[27] J.L. Doob, Stochastic Processes, John Wiley, New York, 1953.

[28] K.F. Gauss, Theory of Motion of the Heavenly Bodies, Dover, New York, New York, 1963.

[29] H.W. Sorenson, "Least Squares Estimation: from Gauss to Kalman," IEEE Spectrum, pp. 63-68, July, 1970.

[30] J.M. Mendel, and D.L. Gieseking, "Bibliography on the Linear-Quadratic-Gaussian Problem," IEEE Trans. Auto Control, Vol. AC-16, No. 6, pp. 847-869, 1971.

[31] N. Wiener, The Extrapolation, Interpolation, and Smoothing of Stationary Time Series with Engineering Applications, John Wiley, New York, 1949.

[32] A.N. Kolmogorov, "Interpolation and Extrapolation of Stationary Time Series," Bulletin Acad. Sciences USSR Math. Ser., Vol. 5, pp. 3-14, 1941.

[33] R.E. Kalman, "A New Approach to Linear Filtering and Prediction Problems," Trans. ASME, J. Basic Engineering, Vol. 82D, pp. 34-45, 1960.

[34] R.E. Kalman, "New Results in Linear Filtering and Prediciton Theory," Trans. ASME, J. Basic Engineering, Vol. 82D, pp. 95-108, 1961.

[35] T.F. Gunckel, G.F. Franklin, "A General Solution for Linear Sampled-Data Control Systems," Trans. ASME, J. Basic Engineering, Vol. 85D, pp. 197-203, 1963.

[36] J.E. Potter, "A Guidance-Navigation Separation Theorem", M.I.T. Exp. Astron. Lab., Report RE-11, 1964.

[37] W.M. Wonham, "On the Separation Theorem of Stochastic Control," SIAM J. Control, Vol. 6, pp. 312-326, 1968.

[38] R.A. Brooks, "Linear Stochastic Control: An Extended Separation Principle," J. Math. Analysis and Applications, Vol. 38, pp. 569-587, 1972.

[39] R.A. Brooks, "On the Separation Theorem of Stochastic Optimal Control," Int. J. Control, Vol. 16, No. 5, pp. 985-991, 1972.

[40] H.S. Witsenhausen, "Separation of Estimation and Control for Discrete-Time Systems," Proc. IEEE, Vol. 59, No. 11, pp. 1557-1566, 1971.

[41] M. Athans, and E. Tse, "A Direct Derivation of the Optimal Linear Filter Using the Maximum Principle," IEEE Trans. Auto. Control, Vol. AC-12, No. 6, pp. 690-698, 1967.

[42] D.E. Johansen, "Optimal Control of Linear Stochastic Systems with Complexity Constraints," in C.T. Leondes (Ed.), Advances in Control Systems, Vol. 4, pp. 182-278, 1966.

[43] C.S. Sims, and J.L. Melsa, "Specific Optimal Estimation," IEEE Trans. Auto. Control, Vol. AC-14, No. 2, pp. 183-186, 1969.

[44] E.C. Tacker, T.D. Linton, and C.W. Sanders, Jr., "Open-Loop Optimal Control of a Class of Continuous-Time Stochastic Systems-- A Simulation Study," Int. J. Control, (Accepted for Publication)

[45] E.C. Tacker, C.W. Sanders, Jr., and T.D. Linton, "Function-Space Derived Algorithms for the Control of Continuous-Time Stochastic Systems," 1971 IEEE Conference on Decision and Control, Miami, Florida, December 16-18, 1971.

[46] E.C. Tacker, C.W. Sanders, Jr., and T.D. Linton, "Suboptimal Feedback Control of Continuous-Time Stochastic Systems via Linear Functional Representation," Third Symposium on Nonlinear Estimation and Its Applications, San Diego, California, September, 11-13, 1972.

[47] E.C. Tacker, C.W. Sanders, Jr., and T.D. Linton, "A Priori Open-Loop Stochastic Control: The Vector Case", Fifth Annual Southeastern Symposium on System Theory, Raleigh, North Carolina, March 22-23, 1973.

[48] E.C. Tacker, C.W. Sanders, Jr., and T.D. Linton, "A Computational Algorithm for the Optimal Control of Continuous-Time Stochastic Systems," IEEE Trans. Auto. Control, Vol. AC-18, No. 3, 1973.

[49] T.D. Linton, "A Function Space Approach to the Stochastic Optimal Control Problem," Dissertation, Louisiana State University, Baton Rouge, La., 1973.

[50] L.S. Pontryagin, V.G. Boltyanski, R.V. Gamkrelidze, and E.F. Mishchenko, The Mathematical Theory of Optimal Processes, trans. by K.N. Trirogroff, Ed. by L.W. Neustadt, Interscience, John Wiley, New York, 1962.

[51] M. Athans, "The Matrix Minimum Principle," Information and Control, Vol. 11, pp. 592-606, 1967.

[52] R.W. Brockett, Finite Dimensional Linear Systems, John Wiley, New York, N.Y., 1970.

[53] E. Wong, Stochastic Processes in Information and Dynamical Systems, McGraw-Hill, New York, N.Y., 1971.

[54] J.M. Mendel, "Computational Requirements for a Discrete Kalman Filter," IEEE Trans. Auto. Control, Vol. AC-16, No. 6, pp. 748-758, 1971.

[55] V.S. Samant, and H.W. Sorenson, "On Reducing the Computational Burden in the Kalman Filter," Proc. Third Symposium on Nonlinear Estimation and Its Application, pp. 192-201, 1972.

[56] E.D. Nering, Linear Algebra and Matrix Theory, 2nd Edition, John Wiley, New York, N.Y., 1970.

[57] R.E. Kalman, "Mathematical Description of Linear Dynamical Systems," SIAM J. Control, Ser. A, Vol. 1, No. 2, pp. 152-192, 1963.

[58] R.S. Bucy, "Global Theory of the Riccati Equation", J. Computer and System Sciences, Vol. 1, No. 4, pp. 349-361, 1967.

[59] W.M. Wonham, "On a Matrix Riccati Equation of Stochastic Control," SIAM J. Control, Ser. A, Vol. 6, No. 4, pp. 681-697, 1968.

[60] D.H. Jacobson, "New Conditions for Boundedness of the Solution of a Matrix Riccati Equation," J. Differential Equations, Vol. 8, No. 2, pp. 258-263, 1970.

[61] C.Y. Chong and M. Athans, "On the Stochastic Control of Linear Systems with Different Information Sets", IEEE Trans. Auto. Control, Vol. AC-16, pp. 423-430, 1971.

[62] P. Dyer and S.R. McReynolds, The Computation and Theory of Optimal Control, Academic Press, New York, N.Y., 1970.

[63] T.D. Linton, D.R. Fischer, E.C. Tacker, and C.W. Sanders, Jr., "Decentralized Control of an Interconnected Electric Energy System Subject to Information Flow Constraints," 1973 IEEE Conference on Decision and Control, San Diego, California, December 5-7, 1973

[64] C.W. Sanders, Jr., D.R. Fischer, E.C. Tacker, and T.D. Linton, "Decentralized Stochastic Control of an Interconnected Electric Energy System," Allerton Conference on Circuits and Systems, October, 1973.

[65] M.D. Mesarovic, D. Macko and Y. Takahara, Theory of Hierarchical Multilevel Systems, Academic Press, New York, 1970.

[66] Y.C. Ho, and K.C. Chu, "Team Decision Theory and Information Structures in Optimal Control Problems--Part I," IEEE Trans. Auto. Control, Vol. AC-17, No. 1, pp. 15-22, 1972.

[67] K.C. Chu, "Team Decision Theory and Information Structures in Optimal Control Problems--Part II," IEEE Trans. Auto. Control, Vol. AC-17, No. 1, pp. 22-28, 1972.

[68] J.D. Glover, and F.C. Schweppe, "Control of Linear Dynamic Systems with Set Constrained Disturbances," IEEE Trans. Auto·Control, Vol. AC-16, No. 5, pp. 411-423, 1971.

[69] D.P. Bertsekas, and I. Rhodes, "On the Minimax Reachability of Target Sets and Target Tubes," Automatica, Vol. 7, No. 2 pp. 233-247, 1971.

[70] T.W. Reddoch, "Models and Cost Functionals for Optimal Automatic Generation Controllers," Ph.D. Dissertation, Louisiana State University, 1973.

[71] R.M. Dressler, and R.E. Larson, "Computation of Optimal Control in Partially Controlled Linear Systems," IEEE Trans. Auto. Control, Vol. AC-14, No. 5, pp. 575-578, 1969.

[72] B. Friedland, "Treatment of Bias in Recursive Filtering," IEEE Trans. Auto. Control, Vol. AC-14, No. 4, pp. 359-367, 1967.

[73] E.C. Tacker, and C.C. Lee, "Linear Filtering in the Presence of Time-Varying Bias," IEEE Trans. Auto. Control, Vol. AC-17, No. 6, pp. 828-829, 1972.

[74] R.S. Bucy, and P.D. Joseph, Filtering for Stochastic Processes with Applications to Guidance, J. Wiley, New York, N.Y., 1968.

[75] C.W. Sanders, Jr., E.C. Tacker, and T.D. Linton, "A Decentralized Filter for Interacting Dynamical Systems," Fourth Symposium on Nonlinear Estimation and Its Applications, San Diego, California, September 10-12, 1973.

[76] R.L. Kosut, "Suboptimal Control of Linear Time Invariant Systems Subject to Control Structure Constaints," IEEE Trans. Auto. Control, Vol. AC-15, No. 5, pp. 557-563, 1970.

[77] W.S. Levine and M. Athans, "On the Determination of the Optimal Constant Output Feedback Gains for Linear Multivariable Systems," IEEE Trans. Auto. Control, Vol. AC-15, No. 1, pp. 44-48, 1970.

[78] R.G. Bartle, The Elements of Integration, John Wiley, New York, 1966.

[79] E.C. Tacker, C.W. Sanders, and T.D. Linton, "Adaptive Strategies in Nonlinear Filtering," Air Force Office of Scientific Research Contract

[80] O.I.Elgerd, Electric Energy Systems Theory:An Introduction, McGraw-Hill, New York, 1971.

[81] E.C. Tacker, T.W. Reddoch, O.T. Tan, and T.D. Linton, "Automatic Generation Control of Electric Energy Systems--A Simulation Study," IEEE Trans. Systems, Man, and Cybernetics, Vol. SMC-3, No. 4, 1973.

APPENDIX A

In this appendix we establish conditions under which the costate matrix, P, used in the derivation of the filter is positive definite. Recall that P is the solution of an equation of the form ,

$$\dot{P}(t) = -\bar{A}\, P(t) - P(t)\, \bar{A}' - Q \qquad t \varepsilon [t_o, t_1]$$

$$P(t_1) = Q_1$$

where $\qquad \bar{A} = A - G\, H.$

Assuming that A, G, H, Q, Q_1 are given continuous real-valued matrices on $[t_o, t_1]$ we first establish a closed form solution for the above matrix differential equation. To this end let Z be defined by the rule

$$Z(t) = P(t_1 - t) \qquad t \varepsilon [0,\ t_1 - t_o]$$

wherein P is the solution to the above matrix differential equation. From the definition of Z it follows that

Lemma 1: Z is the solution to

$$\frac{dZ}{dt} = A^*(t)\ Z(t) + Z(t)\ A^{*\prime}(t) + Q^*, \quad t \varepsilon [0, t_1 - t_o]$$

with

$$Z(0) = Q_1$$

and $\qquad A^*(t) = \bar{A}'(t_1 - t), \quad Q^*(t) = Q(t_1 - t).$

Proof: From the definition of Z and P it follows that

$$\frac{dZ}{dt} = (-1)\left.\frac{dP}{d\lambda}\right|_{\lambda=t_1-t}$$

$$= (-1) \quad -\bar{A}\,(t_1-t)\;P(t_1-t) - P(t_1-t)\;\bar{A}'(t_1-t)$$

$$-Q(t_1-t)$$

$$= A^{*\prime}(t)\;Z(t) + Z(t)\;A^{*}(t) + Q^{*}(t)$$

and $\qquad Z(0) = P(t_1) = Q_1.$

Therefore Z(t) can be written (see [52] p. 58) as

$$Z(t) = \Phi^{*}(t,0)\;Q_1\Phi^{*'}(t,0) + \int_0^t \Phi^{*}(t,\tau)\;Q^{*}(\tau)\;\Phi^{*\prime}(t,\tau)d\tau$$

where

$$\frac{d\Phi^{*}}{dt}(t,\eta) = A^{*}(t)\;\Phi^{*}(t,\eta), \quad \Phi^{*}(\eta,\eta) = I.$$

Hence the definition of Z yields the following expression for P

$$P(t) = Z(t_1-t)$$

$$= \Phi^{*}(t_1-t,0)\;Q_1\;\Phi^{*\prime}(t_1-t,0)$$

$$+ \int_0^{t_1-t} \Phi^{*}(t_1-t,\sigma)\;Q^{*}(\sigma)\;\Phi^{*\prime}(t_1-t,\sigma)d\sigma$$

From the definition of Φ^{*}, it follows that it is the transition matrix of some system and is therefore of full rank. Hence the following result is obtained relative to the positive definiteness of P.

Lemma 2:

(1) If Q_1 is positive definite and Q positive semidefinite then P is positive definite for all $t\varepsilon[t_o,t_1]$.

(2) If Q_1 is positive semidefinite and Q positive definite then P is positive definite for all $t\varepsilon[t_o,t_1)$.

APPENDIX B

A Decentralized Output Feedback Controller for Time-Invariant Linear Systems

I. Introduction

The computational effort involved in designing the constrained controller of Chapter III is quite large and would be prohibitive for most practical systems. One of the major reasons for this difficulty is the apparent lack of a separation principle whereby the controller and filter gains can be designed separately. Thus, instead of two single point boundary value problems the design requires the solution of a single two point boundary value problem with its attendent complexities.

Of course, an engineering approach to this difficulty is to obtain at least a preliminary design by invoking the separation principle. While the results obtained in this appendix are approached from a slightly more general viewpoint they are applicable to the above engineering technique. In this appendix a method is presented for designing a decentralized controller with constant output feedback gains for linear time-invariant systems. Necessary conditions for the optimal feedback gains are obtained in terms of the correlation matrix of the initial state.

To obtain an efficient derivation of these results it is advantageous to first extend some earlier results of Levine and Athans [77]. In [77] Levine and Athans derive a set of necessary conditions for the optimal gain matrix F for the system

$$\dot{x}(t) = A\ x(t) + B\ m(t) \tag{B.1}$$

$$y(t) = H\ x(t) \tag{B.2}$$

where the control, m, is constrained to be a time-invariant transformation of the output. That is,

$$m(t) = -\ F\ y(t) \tag{B.3}$$

and the associated cost is the quadratic form

$$J = 1/2 \int_{0}^{\infty} ||x(t)||_{Q}^{2} + ||m(t)||_{M}^{2} \quad dt \tag{B.4}$$

From (B.1) and (B.2) it is noted that we are considering deterministic systems in this section. However, in order to obtain an F which is not dependent on the initial state x(0), it is useful to employ the artifice of assuming the initial state to be a random variable. Levine and Athans assumed a particular probability distribution for x(0) and then minimized the expected value of (B.4) with respect to F. In this appendix the results in [77] will be generalized in two ways:

(1) It will be shown that the necessary conditions for F may be obtained in terms of the correlation matrix, $E\{x(0)x'(0)\}$ so that the optimal feedback matrix, F^*, does not depend explicitly on the exact distribution of x(0).

(2) The results will be extended to the multicontroller case.

Together these two generalizations provide a technique for designing a decentralized controller for systems with information flow constraints.

II. Extension to Arbitrary Correlation Matrix

The cost in (B.4) can be written as

$$J = x'(0) \left[1/2 \int_0^\infty \Phi'(t,0) \ (Q+H'F'MFH)\Phi(t,0)dt\right]x(0) \qquad (B.5)$$

where Φ is the transition matrix for the undriven system

$$\dot{x}(t) = \left[A - BFH\right]x(t) \qquad (B.6)$$

Considering x(0) to be a random vector and modifying the cost to be the expected value of (B.5), we have

$$J_1 \triangleq E \ \{x'(0)Sx(0)\} \qquad (B.7)$$

where S is defined to be

$$S = 1/2 \int_0^\infty \Phi'(t,0) \ (Q + H'F'MFH) \ \Phi(t,0) \ dt$$

Using the matrix identity $x'Ax = tr\{Axx'\}$, J_1 can be written as

$$J_1 = E \ \{tr \ Sx(0)x'(0)\}$$

and since S is deterministic and the trace operator linear, J_1 can be simplified to

$$J_1 = tr \ \{SX_o\} \qquad (B.8)$$

where X_o is defined by

$$X_o = E \ \{x(0)x'(0)\} \qquad (B.9)$$

The results in [77] were derived for the cost $J_1 = tr\{S\}$ and are therefore valid for any distribution of x(0) for which $X_o = I$.

Suppose now that $X_o = W$ where W is any positive definite+ symmetric matrix. Define T by the relation++

$$T = (W^{1/2})^{-1} \tag{B.10}$$

and note that the vector

$$Z(t) \triangleq T\, x(t) \tag{B.11}$$

has initial correlation matrix

$$\begin{aligned} E\{Z(0)Z'(0)\} &= E\{Tx(0)x'(0)T\} = T\, E\{x(0)x'(0)\}T \\ &= TWT = I \end{aligned} \tag{B.12}$$

Moreover, Z satisfies the differential equation

$$\dot{Z}(t) = TAT^{-1}\, Z(t) + TB\, m(t) \tag{B.13}$$

and the output, y, of the original representation can be written in terms of Z as

$$y(t) = HT^{-1}\, Z(t) \tag{B.14}$$

Finally the original cost in (B.4) can be written as

$$J = 1/2 \int_0^\infty \left[Z'(t)\ (T^{-1}Q\ T^{-1})\ Z(t) + m'(t)\ M\ m(t)\right] dt \tag{B.15}$$

with the control still constrained by (B.3).

+It is possible to drop the constraint that X_o be positive definite, but the proof is too lengthy to present here. The derivation can be obtained for example by paralleling the original proof given in 77.

++$W^{1/2}$ is the (unique) positive definite symmetric matrix satisfying $(W^{1/2})^2 = W$.

Now because of (B.12), the results from [77] may be applied to the reformulated problem consisting of (B.13)-(B.15). Application of these results yields the following set of necessary conditions for the optimal gain matrix, F^*:

$$0 = K^*[TAT^{-1} - TBF^*HT^{-1}] + [TAT^{-1} - TBF^*HT^{-1}]' K^* + T^{-1}Q T^{-1} + T^{-1}H'F^{*'}M F^*HT^{-1} \tag{B.16}$$

$$0 = L^*[TAT^{-1} - TBF^*HT^{-1}]' + [TAT^{-1} - TBF^*HT^{-1}] L^* + I \tag{B.17}$$

$$F^* = M^{-1}B'TK^*L^*T^{-1}H' \; [HT^{-1}L^*T^{-1}H']^{-1} \tag{B.18}$$

Introducing the transformations

$$\hat{K} = TK^*T \tag{B.19}$$

$$\hat{L} = T^{-1}L^*T^{-1} \tag{B.20}$$

and using the symmetry of T and T^{-1} it is easy to verify that (B.16)-(B.18) are equivalent to

$$0 = \hat{K} \; [A - BF^*H] + [A - BF^*H]' \hat{K} + Q + H'F^{*'}MF^*C \tag{B.21}$$

$$0 = \hat{L} \; [A - BF^*H]' + [A - BF^*H] \hat{L} + W \tag{B.22}$$

$$F^* = M^{-1} B' \hat{K}\hat{L}H' \; [HLH']^{-1} \tag{B.23}$$

Thus equations (B.21)-(B.23) represent the necessary conditions for $X_o = W$, which is the desired generalization. If W = I the equations reduce to those derived in [77].

III. Extension to the Multicontroller Case

The system model to be considered here is given by

$$\dot{x}(t) = A\,x(t) + \sum_{i=1}^{N} B_i\, m_i(t) \tag{B.24}$$

$$y_i(t) = H_i x(t) \qquad i=1,2,\ldots N \tag{B.25}$$

with the constraint that each of the N controllers satisfy

$$m_i(t) = -\,F_i\; y_i(t) \qquad i=1,2,\ldots N. \tag{B.26}$$

The associated cost functional for this system is assumed to be

$$J = 1/2 \int_0^{\infty} \left[x'(t)Qx(t) + \sum_{i=1}^{N} m_i'(t)\; M_i m_i(t)\right] dt \tag{B.27}$$

The results of the previous section will now be used to obtain some necessary conditions for this more general problem.

First assuming that an optimal set of gain matrices $\{F_i^*: i=1,2,\ldots N\}$ exist, consider the system in which $\{F_i: i=1,2,\ldots N/\{j\}\}$ are held constant at optimal values and only F_j is considered to be free. In this case the system (B.24) and (B.25) can be written as

$$\dot{x}(t) = \hat{A}_j\; x(t) + B_j\; m_j(t) \tag{B.28}$$

$$y_j(t) = H_j\; x(t) \tag{B.29}$$

wherein $$\hat{A}_j \triangleq A - \sum_{\substack{i=1 \\ i\neq j}}^{N} B_i F_i^* H_i$$

and $$m_j(t) = -\,F_j\; y_j(t).$$

Furthermore, the associcted cost for this modified system can be written as

$$J_j = E\{1/2 \int_0^{\infty} \left[x'(t)\; Q_j x(t) + m_j'(t)\; M_j m_j(t)\right]\, dt\} \tag{B.30}$$

with $\hat{Q}_j$ defined by

$$\hat{Q}_j \triangleq Q + \sum_{\substack{i=1 \\ i\neq j}}^{N} H_i' F_i^{*'} M_i F_i^* H_i$$

and x(0) is assumed to be a random vector with correlation matrix X_o. Applying the results of the previous section to this problem, the necessary conditions for F_j^* are given by

$$0 = K \tilde{A} + \tilde{A}' K + \hat{Q}_j + H_j' F_j^{*'} M_j F_j^* H_j$$

$$0 = L \tilde{A}' + \tilde{A} L + X_o$$

$$F_j^* = M_j^{-1} B_j' K L H_j' \left[H_j L H_j'\right]^{-1}$$

with $\tilde{A}$ defined by

$$\tilde{A} = \hat{A}_j - B_j F_j^* H_j = A - \sum_{i=1}^{N} B_i F_i^* H_i.$$

Now since the term

$$\hat{Q}_j + H_j' F_j^{*'} M_j F_j^* H_j = Q + \sum_{i=1}^{N} H_i' F_i^{*'} M_i F_i^* H_i$$

is independent of j, it follows that the necessary conditions for the N gain matrices F_i^* are given by the N+2 equations

$$0 = K A^* + A^{*'} K + Q + \sum_{i=1}^{N} H_i' F_i^{*'} M_i F_i^* H_i \qquad (B.31)$$

$$0 = L A^{*'} + A^* L + X_o \qquad (B.32)$$

$$F_j^* = M_j^{-1} B_j' K L H_j' \left[H_j L H_j'\right]^{-1} \quad j=1,2\ldots,N \qquad (B.33)$$

with A^* defined by

$$A^* = A - \sum_{i=1}^{N} B_i F_i^* H_i$$

A generalization of the algorithm presented in [77] for solving (B.21)-(B.23) is presently available [63] for solving (B.31)-(B.33).

Questions regarding the existence of F_i^* and stability of the resulting system are open. It is clearly true that if an F_i^* exists and if the system is stabilizable by a decentralized control, then the optimal output feedback, F_i^*, will result in a stable system. Aoki [19] has obtained some results relative to the stabilizability by decentralized output feedback. However, the stabilizability question is still not completely resolved.

APPENDIX C

A Decoupling Transformation

In this appendix a transformation which is potentially useful for decoupling the dynamics of composite systems in a prespecified manner is introduced and discussed. Consider a composite system described by the following state equations

$$\begin{aligned} \dot{x}_1 &= A_{11}x_1 + A_{12}x_2 + b_1 \\ \dot{x}_2 &= A_{21}x_1 + A_{22}x_2 + b_2 \\ x_1(t_o) &= x_1^o \quad , \quad x_2(t_o) = x_2^o \end{aligned} \tag{C.1}$$

wherein for each $t \in [t_o, t_f]$, $x_i(t)$ is an n_i vector $(i=1,2)$ and the matrices A_{ij}, $i,j=1,2$ are continuous and have the appropriate dimensions.

Now consider the transformation, $T: R^{n_1} \times R^{n_2} \to R^{n_1} \times R^{n_2}$ defined by

$$T = \left[\begin{array}{c|c} I_1 & T_{12} \\ \hline T_{21} & I_2 + T_{21}T_{12} \end{array}\right] \tag{C.2}$$

where $I_i: R^{n_i} \to R^{n_i}$ is the identity for $i=1,2$, and T_{12}, T_{21} satisfy the differential equations

$$\dot{T}_{12} = A_{11}T_{12} - T_{12}A_{22} + T_{12}A_{21}T_{12} - A_{12} \tag{C.3}$$

$$\dot{T}_{21} = (A_{22} - A_{21}T_{12})T_{21} - T_{21}(A_{11} + T_{12}A_{21}) - A_{21} \tag{C.4}$$

$$T_{12}(t_o) = 0 \quad , \quad T_{21}(t_o) = 0$$

Lemma 1: Under the transformation, T, the differential equations (C.1) are equivalent to

$$\dot{y}_1 = A_{11}^* y_1 + b_1^* \quad , \quad y_1(t_o) = x_1^o$$

$$\dot{y}_2 = A_{22} y_2 + b_2^* \quad , \quad y_2(t_o) = x_2^o$$

with

$$A_{11}^* = A_{11} + T_{12}A_{21} \; , \quad A_{22}^* = A_{22} - A_{21}T_{12}$$

and

$$b^* = T\, b \; .$$

Proof: Let x be the solution to (C.1) and y=Tx. Then we have

$$\dot{y} = (\dot{T} + TA)x + Tb$$

Now from the definition of T it follows that

$$(\dot{T})_{11} = 0$$

$$(\dot{T})_{12} = \dot{T}_{12} = A_{11}T_{12} - T_{12}A_{22} + T_{12}A_{21}T_{12} - A_{12}$$

$$(\dot{T})_{21} = \dot{T}_{21} = (A_{22} - A_{21}T_{12})T_{21} - T_{21}(A_{11} + T_{12}A_{21}) - A_{21}$$

$$(\dot{T})_{22} = (A_{22} - A_{21}T_{12})T_{21}T_{12} - A_{21}T_{12}(I_2 + T_{21}T_{12}) - T_{21}(A_{12} + T_{12}A_{22})$$

and

$$(TA)_{11} = A_{11} + T_{12}A_{21} = A_{11}^*$$

$$(TA)_{12} = A_{12} + T_{12}A_{22}$$

$$(TA)_{21} = T_{21}A_{11}^* + A_{21}$$

$$(TA)_{22} = T_{21}A_{12} + (I_2 + T_{21}T_{12})A_{22}$$

Therefore

$$(\dot{T} + TA)_{11} = A_{11}^*$$

$$(\dot{T} + TA)_{12} = A_{11}^* T_{12}$$

$$(\dot{T} + TA)_{21} = A_{22}^* T_{21}$$

$$(\dot{T} + TA)_{22} = A^* \; (I_2 + T_{21}T_{12})$$

and hence

$$\dot{T} + TA = A^{*}T \quad , \quad \text{with} \quad A^{*} = \left[\begin{array}{c|c} A_{11}^{*} & 0 \\ \hline 0 & A_{22}^{*} \end{array}\right]$$

Then

$$\begin{aligned} \dot{y} &= A^{*}Tx + Tb \\ &= A^{*}y + b^{*} \end{aligned}$$

and clearly since $T(t_o) = I$ the above initial conditions hold.

Another interesting property of the T defined above is that it has a simple inverse which is easily computed once T is known.

Lemma 2: The transformation defined in (C.2)-(C.4) is nonsingular and its inverse is given by

$$T^{-1} = \left[\begin{array}{c|c} I_1 + T_{12}T_{21} & -T_{12} \\ \hline -T_{21} & I_2 \end{array}\right]$$

Proof: By direct computation we have $TT^{-1} = I$, and $T^{-1}T = I$.

Example: Consider the case in which $A_{21} = 0$. Then from (C.4) we have $T_{21}(t) = 0$ for all t and from (C.3) T_{12} is the solution of

$$\dot{T}_{12} = A_{11}T_{12} - T_{12}A_{22} - A_{12}$$

$$T_{12}(t_o) = 0$$

$$A_{11}^{*} = A_{11} \quad , \quad A_{22}^{*} = A_{22}$$

and

$$T^{-1} = \left[\begin{array}{c|c} I_1 & -T_{12} \\ \hline 0 & I_2 \end{array}\right] .$$

Appendix D

An Application of the Decentralized Filter

In Chapter II qualitative aspects of the filter performance were discussed and it was shown that for time invariant systems the performance loss of the decentralized filter attains a bounded steady state value if the pair $[A_i - L_{ii}H_{i1}, H_{i3}]$ is observable for each $i = 1,2,\ldots,N$. In this appendix quantitative performance data are generated for the case of state estimation in a two-area power system.

In this preliminary study the two-area power system has been represented by the simple model shown in Figure E.1. (See e.g. [80] or [81] for a development of this model). Here the local state vector for each area was taken to be $x_i = [\Delta f_i \ \Delta P_{Gi} \ \Delta X_{Ei}]$ and the interaction input was chosen to be the tie-line power. Consistent with standard practice in the electric power industry it was assumed that the observations available at each area consisted of noisy measurements of the local frequency deviation and the tie line power.

In this model the "synchronizing coefficient", T_{12}, is directly related to the amount of interarea coupling. Therefore in this preliminary study attention was focused on determining the effect of this parameter on the normalized performance loss

$$P.L. \triangleq \frac{J_D - J_C}{J_C} \times 100\%$$

wherein $J_D(J_C)$ is the cost of estimation using the decentralized (centralized) filter. Figure E.2 shows the results obtained from this study. Note that these results are consistent with what one would expect intuitively in the sense that as $T_{12} \to 0$ the system becomes decoupled and the optimal centralized filter should approach two local filters of the

form used in the decentralized filter. Thus one would expect T_{12} to have the effect shown in Figure E.2.

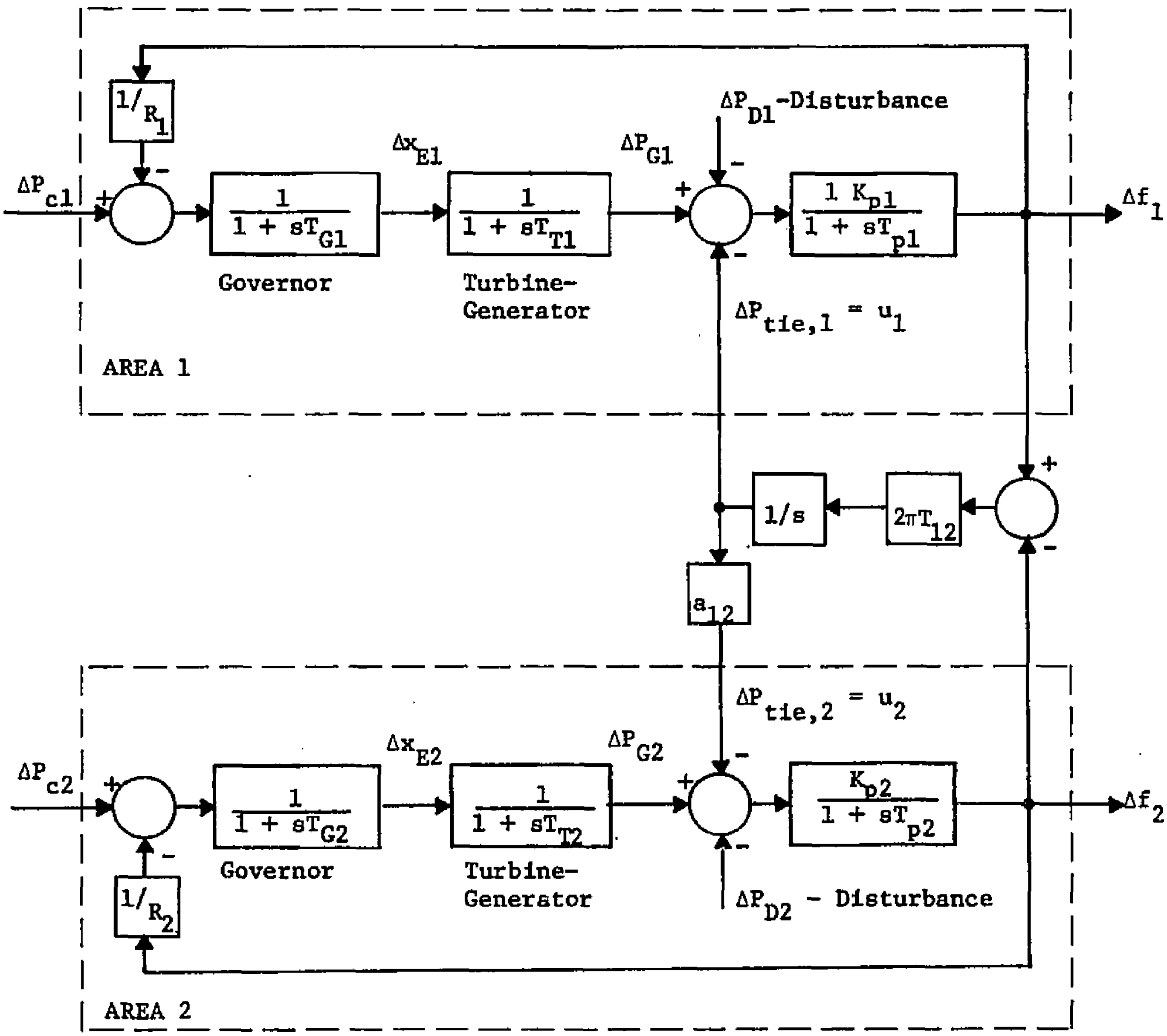

Fig. D.1 Two-area power system model.

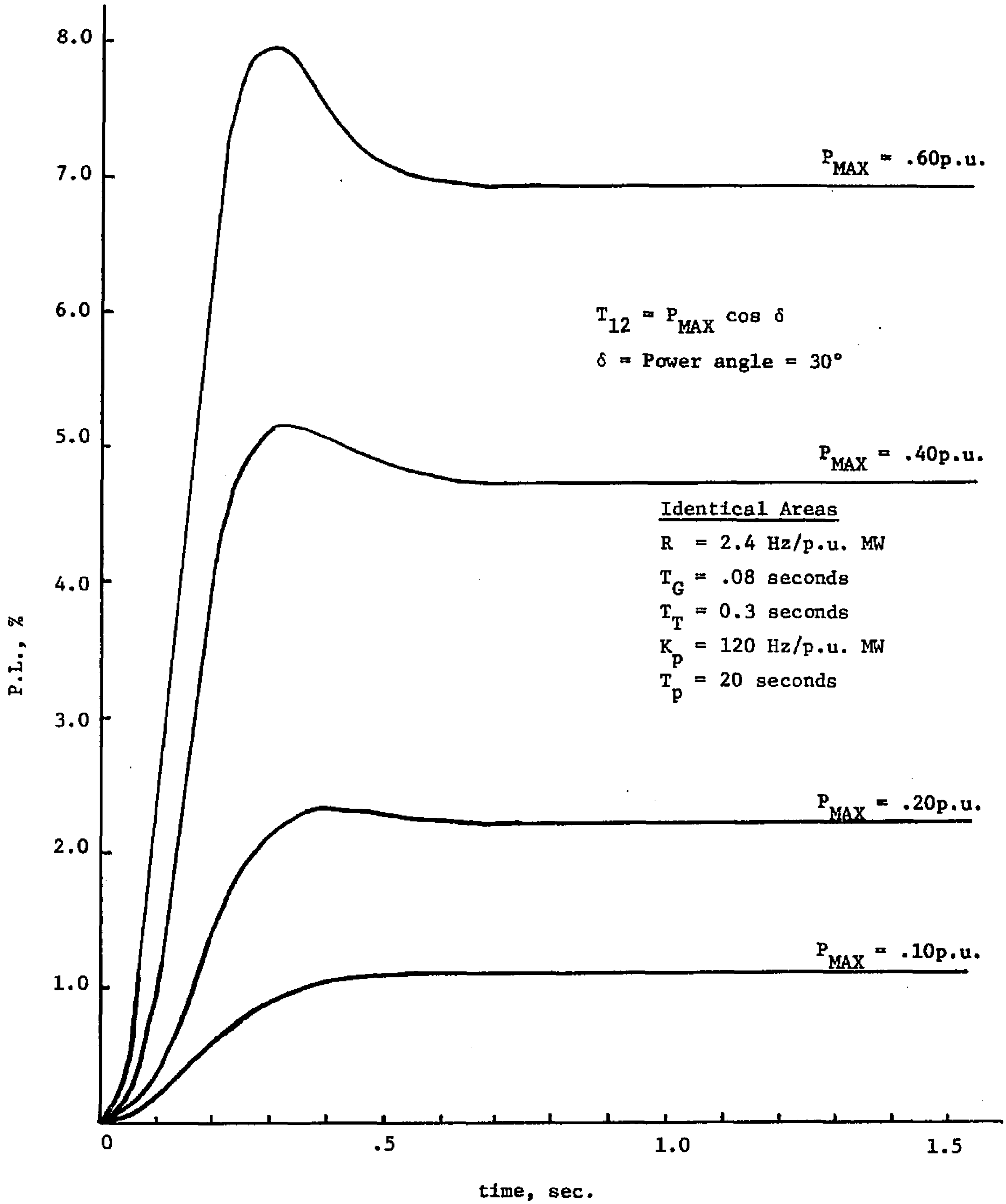

Fig.D.2 Percentage performance loss vs. system coupling.

Appendix E

In this appendix details involved in simplifying the structure of the decentralized filter for disturbance inputs of the form given in Section V, Chapter II are given. To this end consider the model in (2.48-2.49) and define $\tilde{x}_i = [x_i' \ \gamma_i']'$, $\tilde{x} = [\tilde{x}_1' \ \tilde{x}_2' \ \ldots \ \tilde{x}_N']'$. Then the subsystem model becomes

$$\tilde{S}_i: \quad \dot{\tilde{x}}_i = \begin{bmatrix} A_i & \Gamma_i \\ 0 & \theta_i \end{bmatrix} \tilde{x}_i + \begin{bmatrix} L_{ii} \\ 0 \end{bmatrix} u_i + \begin{bmatrix} B_i \\ 0 \end{bmatrix} m_i + \begin{bmatrix} w_i \\ 0 \end{bmatrix}$$

$$= \tilde{A}_i \tilde{x}_i + \tilde{L}_{ii} u_i + \tilde{B}_i m_i + \tilde{w}_i \qquad \text{(E.1)}$$

$$y_i = [\bar{H}_i \quad 0] \tilde{x}_i + \bar{H}_{i2} u_i + \eta_i$$

$$= \tilde{H}_{i1}^{*} \tilde{x}_i + \bar{H}_{i2} u_i + \eta_i \qquad \text{(E.2)}$$

$$u_i = \tilde{L}_i \tilde{x} \qquad \text{(E.3)}$$

wherein $\tilde{L}_i$ is an appropriately modified version of L_i.

Therefore the resulting augmented system (E.1)-(E.3) can be put into the form treated in Section III. Note that all of the statistical assumptions regarding $x_i(t_o,\cdot)$, w_i, and η_i remain true for $\tilde{S}$. In addition, if the system without disturbances satisfies the $\bar{H}_{i2}$ assumption, then so will $\tilde{S}$ and hence, without further loss of generality, it can be assumed that the measurement matrices are in canonical form. For ease of writing let μ_i denote the optimal S.L.U. estimate of $\tilde{x}_i$ and apply the results of Section III to obtain

$$\dot{\mu}_i = \tilde{A}_i\mu_i + \tilde{L}_{ii}(y_{i1}-\tilde{H}_{i1}\mu_i) + \tilde{B}_i m_i + G_{i2}^{*}(y_{i2}-\tilde{H}_{i3}\mu_i) \qquad \text{(E.4)}$$

wherein

$$G^*_{i2} = (V_i\tilde{H}'_{i3} - \tilde{L}_{ii}N_{i3})N^{-1}_{i2} \tag{E.5}$$

$$\dot{V}_i = \tilde{A}^*_i V_i + V_i\tilde{A}^{*'}_i - V_iH'_{i3}N^{-1}_{i2}V_i + \tilde{W}^*_i \tag{E.6}$$

$$\tilde{H}^*_{i1} = \left[\begin{array}{c|c} H_{i1} & 0 \\ \hline H_{i3} & 0 \end{array}\right] = \left[\begin{array}{c} \tilde{H}_{i1} \\ \hline \tilde{H}_{i3} \end{array}\right] \tag{E.7}$$

$$A^*_i = \tilde{A}_i + \tilde{L}_{ii}(N_{i3}\, N^{-1}_{i2}\, \tilde{H}_{i3} - \tilde{H}_{i1}) \tag{E.8}$$

and

$$\tilde{W}^*_i = \tilde{W}_i + \tilde{L}_{ii}(N_{i1} - N_{i3}N^{-1}_{i2}N'_{i3})\tilde{L}_{ii} \tag{E.9}$$

$$V_i(t_o) = \left[\begin{array}{cc} \Sigma_i & 0 \\ \hline 0 & \Sigma^o_{\gamma_i} \end{array}\right] \tag{E.10}$$

Equations (E.4 - E.10) provide a solution for the case in which disturbances of the above class are present. Moreover it is possible to simplify the filter along the same lines as in Friedland [72]. This has important implications relative to the computational aspects of the filter implementation. To obtain these simplifications note that

$$\tilde{A}^*_i = \left[\begin{array}{c|c} A^*_i & \Gamma_i \\ \hline 0 & \theta_i \end{array}\right], \quad W^*_i = \left[\begin{array}{c|c} W^*_i & 0 \\ \hline 0 & 0 \end{array}\right]$$

and let

$$V_i = \left[\begin{array}{c|c} V_{i1} & V_{i3} \\ \hline V'_{i3} & V_{i2} \end{array}\right].$$

It can be shown [72] that

$$V_i = \tilde{V}_i + QMQ'$$

where $\tilde{V}_i$ satisfies (2.56) with

$$\tilde{V}_i(t_o) = \left[\begin{array}{c|c} \Sigma_i^o & 0 \\ \hline 0 & 0 \end{array}\right]$$

and V_i satisfies (E.6) with

$$V_i(t_o) = \left[\begin{array}{c|c} \Sigma_i^o & 0 \\ \hline 0 & \Sigma_{\gamma_i}^o \end{array}\right]$$

and

$$\dot{Q} = (\tilde{A}_i^* - \tilde{V}_i\tilde{H}_{i3}'N_{i2}^{-1}\tilde{H}_{i3})Q \ , \quad Q(t_o) = \left[\frac{0}{I}\right] \tag{E.11}$$

$$\dot{M} = -MQ'\tilde{H}_{i3}'N_{i2}^{-1}\tilde{H}_{i3}'QM \ , \ M(t_o) = \Sigma_{\gamma_i}^o . \tag{E.12}$$

From the definition of $\tilde{V}_i$ it follows that $\tilde{V}_{i2} = 0$ and $\tilde{V}_{i3} = 0$. Thus, partitioning Q to conform with the partitioning of $\tilde{A}_i^*$ $(Q' = [Q_x' \ Q_\gamma'])$ it follows that Q_x, Q_γ, and M are given by

$$\dot{Q}_x = (A_i^* - \tilde{V}_{i1}H_{i3}'N_{i2}^{-1}H_{i3}) \ Q_x + \Gamma_i Q_\gamma, \ Q_x(t_o) = 0 \tag{E.13}$$

$$\dot{Q}_\gamma = \theta_i Q_\gamma \qquad\qquad Q_\gamma(t_o) = I \tag{E.14}$$

and $$\dot{M} = -MQ_x'H_{i3}'N_{i2}^{-1}H_{i3}Q_xM \qquad\qquad M(t_o) = \Sigma_{\gamma_i}^o \tag{E.15}$$

Writing (E.5) in detail it follows that G_{i2}^* is given by

$$G_{i2}^* = \left[\begin{array}{c} (V_{i1}H_{i3}' - L_{i1}N_{i3})N_{i2}^{-1} \\ \hline V_{i3}'H_{i3}'N_{i2}^{-1} \end{array}\right] \tag{E.16}$$

Thus, using (E.16) the filter dynamical equation, (E.4), can be written $(\mu_i' \triangleq [\bar{x}_i' \ \bar{\gamma}_i'])$

$$\dot{\bar{x}}_i = A_i\bar{x}_i + \Gamma_i\bar{\gamma}_i + L_{i1}(y_{i1} - H_{i1}\bar{x}_i) + B_i m_i + \tag{E.17}$$

$$(V_{i1}H_{i3}' - L_{i1}N_{i3})N_{i2}^{-1} \ (y_{i2} - H_{i3}\bar{x}_i), \ \bar{x}_i(t_o) = x_i^o$$

$$\dot{\bar{\gamma}}_i = \theta_i \bar{\gamma}_i + V_{i3}' H_{i3}' N_{i2}^{-1} (y_{i2} - H_{i3}\, \bar{x}_i), \qquad \bar{\gamma}_i(t_o) = \gamma_i^o \tag{E.18}$$

Now define x_i^* as the estimate given by the local "disturbance-free" filter,

$$\dot{x}_i^* = A_i\, x_i^* + L_{ii}(y_{i1} - H_{i1} x_i^*) + B_i m_i + (\tilde{V}_{i1} H_{i3}' - L_{ii} N_{i3}) N_{i2}^{-1} (y_{i2} - H_{i3} x_i^*)$$

$$x_i^*(t_o) = x_i^o \tag{E.19}$$

and let

$$\underline{x}_i = x_i^* + S\bar{\gamma}_i \tag{E.20}$$

with $\quad S = Q_x Q_\gamma^{-1}$ (E.21)

It is easy to show from the definition of S, Q_x, and Q_γ that

$$\dot{S} = (A_i^* - \tilde{V}_{i1} H_{i3}' N_{i2}^{-1} H_{i3}) S - S\theta_i + \Gamma_i \tag{E.22}$$

From the definition of $\underline{x}_i$, $\bar{x}_i$ and (E.22) it is straightforward to show that $e_i \triangleq \bar{x}_i - \underline{x}_i$ satisfies the stochastic differential equation

$$\dot{e}_i = [A_i - L_{ii} H_{i1} - (\tilde{V}_{i1} H_{i3}' - L_{ii} N_{i3}) N_{i2}^{-1} H_{i3}] e_i \tag{E.23}$$

$e_i(t_o) = 0$ with probability one.

and hence

$$e_i(t) = 0 \text{ (wp1)} \tag{E.24}$$

It follows from (E.20) and (E.24) that

$$\bar{x}_i = \underline{x}_i = x_i^* + S\, \bar{\gamma}_i$$

and thus the filter has the simplified structure given in Section V, Chapter II.

VITA

Charles W. Sanders, Jr., son of C. Woodford and Mary E. Sanders, was born in Jackson, Mississippi on January 10, 1943 and received his early schooling in Kentwood, Louisiana. In 1961 he entered Louisiana State University and received his BSEE and MSEE degrees in 1965 and 1968 respectively. During this time he held summer positions with Shell Oil Company and George C. Marshall Spaceflight Center and in 1967 was an instructor in Electrical Engineering. He became a member of the Technical Staff of the North American Rockwell Corporation in 1968 where he was primarily involved in the development and implementation of real-time hybrid simulations for advanced avionics systems.

From 1970 to 1972 he was a National Science Foundation Trainee and in 1973 a Dissertation Year Fellow at Louisiana State University. While studying for the Doctoral degree under the direction of Professor Edgar C. Tacker, he has contributed a number of articles to the control and filtering literature and has presented results from his research at regional and national symposia. In collaboration with Professor Tacker he has served as a consultant to the Container Corporation of America.

In August 1973, Mr. Sanders joined the faculty of the Electrical and Computer Engineering Department, University of Wisconsin, Madison, Wisconsin where he is presently teaching and conducting research in the area of mathematical system theory. His research interests lie primarily in the area of filtering and control for

large-scale systems with particular emphasis on the development of results which are applicable to electric power systems.

Mr. Sanders is a member of Eta Kappa Nu, Tau Beta Pi and Phi Kappa Phi and the Control Systems, Computer, and Power System Societies of the Institute of Electrical and Electronic Engineers.

EXAMINATION AND THESIS REPORT

Candidate: Charles Woodford Sanders, Jr.

Major Field: Electrical Engineering

Title of Thesis: Decentralized Filtering and Control in Interconnected Stochastic Systems

Approved:

Edgar C. Tacker

Major Professor and Chairman

James G. Traynham

Dean of the Graduate School

EXAMINING COMMITTEE:

J. R. Dorroh

P M Gubel

D. E. Johnson

Pen P. Tan

Date of Examination:

August 8, 1973