



A state-space approach to the design of pre- and postfilters in a sampled-data communication system
by Yern Yeh

A thesis submitted to the Graduate Faculty in partial fulfillment of the requirements for the degree of
DOCTOR OF PHILOSOPHY in Electrical Engineering
Montana State University
© Copyright by Yern Yeh (1974)

Abstract:

The basic problem considered in this thesis is how to discretize a random process observed at one end, and after transmission, to reconstruct the desired signal from the discretized data received at the other end. The discretizing process consists of sampling, quantization and a presampling operation. The system under consideration in this work can either be regarded as a data communication or a data processing system.

The main purposes of this work are: 1. To solve for the optimal realizable pre-transmission filter (prefilter) at one end, and the optimal realizable reconstruction filter (postfilter) at the other end using a state-space representation and the innovations approach.

2. To investigate the functions of the optimal realizable pre- and postfilters in comparison with that of the optimal unrealizable filters. 3. To investigate the smoothing problem as a function of allowable delay.

The major contributions of this work are: 1. The state-space and innovations approach has been successfully applied to derive the optimal realizable zero-lag pre-and postfilters. The initial step in the procedure is to find the optimal post (pre) filter with a given pre (post) filter. Joint optimization then is carried out for the special case in which the quantization noise and channel additive noise are absent. The optimal filters are specified as solutions to differential equations in the state variables. 2. In the absence of the optimal prefilter, the optimal finite-lag postfilter has been found. An error-reduction factor and a delay-efficiency factor have been defined in order to study the effect of introducing delay on reducing the error between the output of the system and the desired signal. 3. Numerical examples of the performances of various systems have been worked out, and interpretations are given.

A STATE-SPACE APPROACH TO THE DESIGN OF
PRE- AND POSTFILTERS IN A SAMPLED-DATA
COMMUNICATION SYSTEM

by

YERN YEH

A thesis submitted to the Graduate Faculty in partial
fulfillment of the requirements for the degree

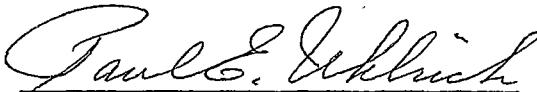
of

DOCTOR OF PHILOSOPHY

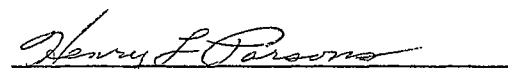
in

Electrical Engineering

Approved:


Head, Major Department


Chairman, Examining Committee


Graduate Dean

MONTANA STATE UNIVERSITY
Bozeman, Montana

June, 1974

ACKNOWLEDGMENTS

The author wishes to express his sincere appreciation to Dr. C. K. Rushforth not only for his invaluable academic guidance, but also for his moral support during the course of this research work.

Appreciation is likewise extended to Dr. D. A. Pierre, and to Dr. R. M. Johnson for their numerous suggestions.

The financial support provided by the Department of Electrical Engineering of Montana State University is gratefully acknowledged. The author wishes to thank Dr. P. E. Uhlrich, in particular, for this matter.

The author also wishes to thank Professor R. C. Seibel for his constant assistance during the author's stay in Montana State University.

Finally, the author wishes to express his gratitude to his wife Ning-Chia, and to his parents Mr. and Mrs. K. H. Yeh for their encouragements in pursuing the Ph.D. degree.

TABLE OF CONTENTS

	<u>Page</u>
VITA	ii
ACKNOWLEDGMENTS	iii
TABLE OF CONTENTS	iv
LIST OF TABLES	viii
LIST OF FIGURES	x
ABSTRACT	xv
Chapter	
1. INTRODUCTION	1
1.1 Problem Statement	1
1.2 Historical Summary	4
1.3 Outline of Contents and Statement of Results	13
1.4 Basic Concepts and Theorems	17
1.4.1 Sampling theorem and aliasing effect	17
1.4.2 The orthogonality principle in linear estimation	21
1.4.3 The innovations approach	23
1.4.4 The smoothing problem	27
1.5 Characterization of the Quantization Noise	31

Chapter	Page
2. THE OPTIMAL POSTFILTER	41
2.1 Problem Statement and Preliminary Remarks	41
2.2 Optimum Filtering and Prediction	51
2.2.1 A two-filters structure	52
2.2.2 The innovations approach	55
2.3 Optimal Smoothing	64
2.3.1 Delay of one sample	65
2.3.2 Delays of more than one sample	67
2.3.3 General solution	71
2.3.4 Delay efficiency	80
2.4 Examples	
3. THE OPTIMAL PREFILTER	110
3.1 Problem Statements and Preliminary Remarks	110
3.2 Jointly Optimized Pre- and Postfilters with Multiple Sampled-data Links	116
3.3 The Optimal Prefilter with a Single Sampled-data Link	123
3.3.1 The structure of the optimal prefilter	126
3.3.2 Examples	140

Chapter	Page
4. JOINT OPTIMIZATION OF THE PRE- AND POSTFILTERS	144
4.1 Problem Statement and Preliminary Remarks	144
4.2 Degeneration of the Joint Optimization of the Pre- and Postfilters	145
4.3 Jointly Optimized Pre- and Postfilters with $n_2(k)=0$	153
4.4 The Application of Nonlinear Programming	167
4.4.1 Outline of procedure	167
4.4.2 Examples	173
5. SYSTEM PERFORMANCE	181
5.1 Problem Statement and Preliminary Remarks	181
5.2 The Frequency-domain Approach to the Filtering Problems	182
5.2.1 Jointly optimized unrealizable pre- and postfilters	188
5.2.2 Jointly optimized realizable pre- and postfilters	190
5.3 Error Analysis	196
5.3.1 Outline of the general procedure used in working examples	196
5.3.2 Case 1: A first-order Gauss-Markov process with an uncorrelated colored noise	201

	<u>Page</u>
5.3.3 Case 2: A second-order Butterworth process with an uncorrelated white noise	221
5.3.4 Case 3: A second-order Gauss-Markov process without noise	225
5.3.5 Case 4: Second-order Gauss-Markov processes with uncorrelated colored noises	226
5.3.6 Conclusion	239
Chapter	
6. CONCLUSION, AND SUGGESTIONS FOR FURTHER WORK	245
Appendix	
A. CONTINUOUS KALMAN-BUCY STATE-VECTOR FILTER AND ITS EXTENSIONS	257
B. DISCRETE KALMAN-BUCY STATE-VECTOR FILTER AND ITS EXTENSIONS	269
C. DISCRETIZATION OF CONTINUOUS STATE EQUATION	275
D. ERROR COVARIANCE OF LINEAR ESTIMATION	282
E. CONDITIONS FOR THE OPTIMAL PREFILTER	292
REFERENCES	307

LIST OF TABLES

<u>Table</u>	<u>Page</u>
1.1 Autocorrelation functions of the input, output, and the quantization noise of the Max equal-interval quantizer	37
2.1 Effect of delay on error reduction for Example 2.1, $aT=0.01$	91
2.2 Effect of delay on error reduction for Example 2.1, $V_n/\sigma=0.05$	91
2.3 Effect of delay on error reduction for Example 2.2, $V_n/\sigma=0.1$, $aT=0.02$	100
2.4 Effect of delay on error reduction for Example 2.2, $b=7a$ and $V_n/\sigma=0.1$	100
2.5 Effect of delay on error reduction for Example 2.2, $aT=0.02$ and $b=7a$	105
2.6 Effect of delay on error reduction for Example 2.3	105
5.1 Normalized errors of the optimum systems with, and without a prefilter for Case 1, $V_n/\sigma=1.0$	204
5.2 Normalized errors of the suboptimum systems with ZOH as the postfilter, and with, and without an optimum prefilter for Case 1, $b=14a$	215
5.3 Performances of various sytems for Case 1, $V_n/\sigma=0.8$ and $b=13a$	222
5.4 Performances of various systems for Case 3, $\sigma_a=0.6\sigma_b$ and $b=10a$	228

<u>Table</u>	<u>Page</u>
5.5 Normalized errors of the optimum systems with, and without a prefilter for Case 4, $V_n/\sigma=0.5$, $\sigma_a=\sigma_b$ and $b=3a$	232
5.6 Normalized errors of the suboptimum systems with ZOH as the postfilter, and with, and without an optimum prefilter for Case 4, $\sigma_a=\sigma_b$, $b=3a$ and $V_n/\sigma=0.5$	234
5.7 Performances of various systems for Case 4, $\sigma_a=\sigma_b$, $V_n/\sigma=0.5$ and $b=3a$	237

LIST OF FIGURES

<u>Figure</u>		<u>Page</u>
1.1	A sampled-data communication system	2
1.2	Models of communication systems	6
1.3	Reconstruction of a bandlimited signal after sampling	19
1.4	Power spectrum of (a) a nonbandlimited signal; (b) its sampled output	20
1.5	(a) Optimal pre- and postfilters for the signal with power spectrum shown in Fig. 1.4 _a ; (b) the power spectrum of the resulting output	22
1.6	Geometric interpretation of the orthogon- ality principle in linear estimation	24
1.7	The innovations approach	24
1.8	The general shape of the 'Reconstruction density function' of a first-order G.M. (Gauss-Markov) process	29
1.9	Normalized reconstruction error of a first-order G.M. process after sampling	30
1.10	An equal-interval quantizer	32
1.11	Autocorrelation functions of quantization noises	36
1.12	Ratios of $R_n(0.1)$ (or $R_{n_c}(0.1)$) to $R_n(0)$ ($R_{n_c}(0)$) versus the number of quantization levels	39

<u>Figure</u>		<u>Page</u>
2.1	The design of the postfilter: (a) the message model; (b) the generation of the desired signal; (c) the system . . .	42
2.2	A two-filter structure for the optimal unrealizable postfilter	48
2.3	A two-filter structure for the optimum zero-lag postfilter	54
2.4	A piece-wise continuous function and its derivative	56
2.5	The optimum zero-lag postfilter: (a) the innovation process is generated by a separate discrete KSVF. (b) The innovation process is generated by a feedback loop	62
2.6	The optimal fixed-lag postfilter specified by Eq. (2.91)	75
2.7	The optimal fixed-lag postfilter specified by (2.110) and (2.111)	81
2.8	Normalized m.s.e. (mean-square error) versus the amount of delay allowed for Example 2.1, $aT=0.01$	86
2.9	Delay efficiency of Example 2.1, $aT=0.01$	87
2.10	Normalized m.s.e. versus the amount of delay allowed for Example 2.1, $V_n/\sigma=0.05$	88
2.11	Delay efficiency of Example 2.1, $V_n/\sigma=0.05$	89
2.12	c_t versus aT , an indication of the rate of convergence of the delay efficiency for Example 2.1	92

<u>Figure</u>		<u>Page</u>
2.13	c_t versus V_n/σ , an indication of the rate of convergence of the delay efficiency for Example 2.1	93
2.14	Normalized m.s.e. versus the amount of delay allowed for Example 2.2, $V_n/\sigma=0.1$ and $aT=0.02$	98
2.15	Delay efficiency of Example 2.2, $V_n/\sigma=0.1$ and $aT=0.02$	99
2.16	Normalized m.s.e. versus the amount of delay allowed for Example 2.2, $V_n/\sigma=0.1$ and $b=7a$	101
2.17	Delay efficiency of Example 2.2, $V_n/\sigma=0.1$ and $b=7a$	102
2.18	Normalized m.s.e. versus the amount of delay allowed for Example 2.2, $aT=0.02$ and $b=7a$	103
2.19	Delay efficiency of Example 2.2, $aT=0.02$ and $b=7a$	104
2.20	Normalized m.s.e. versus the amount of delay allowed for Example 2.3	108
2.21	Delay efficiency of Example 2.3	109
3.1	The jointly optimized pre- and postfilters with multiple sampled-data links and $n_2=0$: (a) the message model; (b) the system	122
3.2	The innovations approach to the prefiltering problem	128
3.3	The optimum prefilter with a single sampled-data link and a given postfilter	139

<u>Figure</u>		<u>Page</u>
4.1	Degeneration of the joint optimization problem: (a) generation of $x_s(t)$; (b) the degenerated system	152
4.2	Degeneration of the joint optimization problem when $n_2=0$	168
5.1	EFG and ED versus aT for Case 1	203
5.2	The aliasing of random processes with different spectrums (after sampling)	206
5.3	The error-reduction efficiency ED/EG versus aT for Case 1	207
5.4	The sampling period at which peak error reduction occurs by using the optimum prefilter for Case 1	209
5.5	The error-reduction efficiency ED/EG versus aT for Case 1	210
5.6	EFZ, EMZ, and EZ versus aT for Case 1	214
5.7	EDZ/EMZ versus aT for Case 1	217
5.8	Normalized ERRSA, ERRN, and EUNFG versus aT for Case 1	219
5.9	From bottom to top, EUNFG, EUNG, EFG, EFZ, and EG versus aT for Case 1	220
5.10	EFM and (EFG-EFM) versus T for Case 2	223
5.11	EUNFG and EFG versus T for Case 2	224
5.12	From right to left, EUNFG, EUNG, EFM, and EG (dotted line) versus aT for Case 3	227
5.13	EFM and EDM versus aT for Case 4	230

<u>Figure</u>		<u>Page</u>
5.14	The error-reduction efficiency EDM/EG versus aT for Case 4	231
5.15	EUNFG, EFM, and EG versus aT for Case 4 . .	233
5.16	EFM, EFZ, EMZ, and EZ versus aT for Case 4	236
5.17	The error-reduction efficiency EDZ/EMZ versus aT for Case 4	238
5.18	The optimum systems: (a) without a prefilter; and (b) with a prefilter, where $z(t)$ is a first-order G.M. process	240
5.19	EFGQ and EG versus aT , where the signal and n_1 are the same as those given in Case 1, and n_2 represents the quantization noise	244
C.1	The auto correlation function of a random process which can be approxi- mated by a white noise	280

ABSTRACT

The basic problem considered in this thesis is how to discretize a random process observed at one end, and after transmission, to reconstruct the desired signal from the discretized data received at the other end. The discretizing process consists of sampling, quantization and a presampling operation. The system under consideration in this work can either be regarded as a data communication or a data processing system.

The main purposes of this work are: 1. To solve for the optimal realizable pre-transmission filter (prefilter) at one end, and the optimal realizable reconstruction filter (postfilter) at the other end using a state-space representation and the innovations approach. 2. To investigate the functions of the optimal realizable pre- and postfilters in comparison with that of the optimal unrealizable filters. 3. To investigate the smoothing problem as a function of allowable delay.

The major contributions of this work are: 1. The state-space and innovations approach has been successfully applied to derive the optimal realizable zero-lag pre- and postfilters. The initial step in the procedure is to find the optimal post (pre) filter with a given pre (post) filter. Joint optimization then is carried out for the special case in which the quantization noise and channel additive noise are absent. The optimal filters are specified as solutions to differential equations in the state variables. 2. In the absence of the optimal prefilter, the optimal finite-lag postfilter has been found. An error-reduction factor and a delay-efficiency factor have been defined in order to study the effect of introducing delay on reducing the error between the output of the system and the desired signal. 3. Numerical examples of the performances of various systems have been worked out, and interpretations are given.

Chapter 1

INTRODUCTION

1.1 Problem Statement

In designing data communication or data processing systems, channel capacity is often one of the major designing factors. Whenever the amount of information exceeds the capacity of the channel being used for transmitting the information, it is advantageous, or even necessary, to compress the information before transmission. For example, one has to discretize an analog signal before transmitting it through a digital channel. After receiving the compressed information such as the discretized data at the other end, one then tries to reconstruct the original information from the compressed data according to a certain fidelity criterion.

Numerical analysis provide an example of the discretization and reconstruction process. For example, functions such as $\exp(x)$, $\cos(x)$, etc., are first tabulated and stored. One then approximates the value of $\exp(x)$ at an arbitrary point by some kind of interpolation process from a set of tabulated values.

The system under consideration in this thesis is depicted in Fig. 1.1. $n_1(t)$ in Fig. 1.1_a is the primary

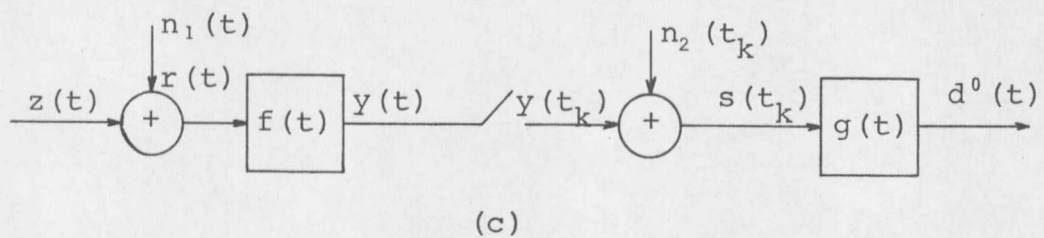
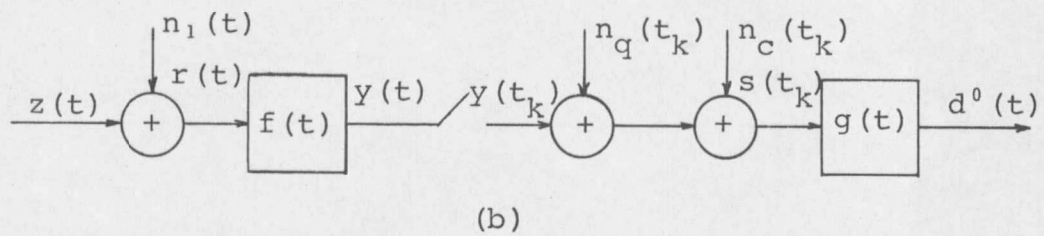
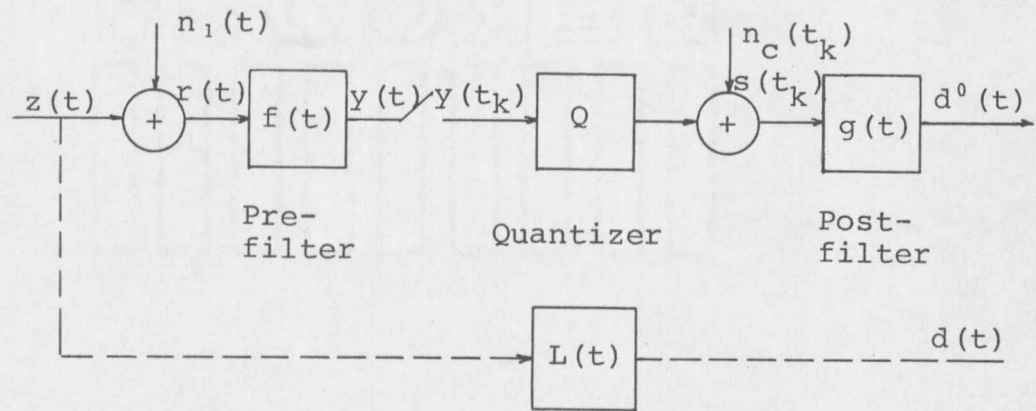


Figure 1.1 A sampled-data communication system.

observation noise. The observation process $r(t)$ is discretized before transmission. The discretizing of $r(t)$ could simply be accomplished by sampling and quantization, but in general it might be advantageous to operate on $r(t)$ prior to sampling in such a way that the overall distortion is reduced. This is one of the basic reasons why a prefilter is inserted in the system. The functions of the prefilter will be investigated in more detail later. The function of the postfilter is obvious; to reconstruct the desired signal from the discretized and possibly contaminated data received at the other end. The functions of the quantizer are:

1. To combat the additive noises in the channel (55).
2. To reduce the amount of information to be transmitted (31).

In this work, the design of the quantizer is not undertaken; only its effects on the design of the optimal pre- and postfilters are considered. In Fig. 1.1_b, the input-output relationship of the quantizer is specified by an additive noise $n_q(t_k)$ called the quantization noise. If the sampling process is periodic, $t_k = kT$, otherwise t_k denotes the time the k th sampling takes place. In

Fig. 1.1, $n_c(t_k)$ and the channel additive noise $n_c(t_k)$ are combined to become $n_2(t_k)$. The characterization of the quantization noise will be discussed in Section 1.5.

Depending on whether n_1 , n_c or/and n_q are present or not, the system shown in Fig. 1.1 can either be regarded as a PCM (Pulse Code Modulation) system, a PAM (Pulse Amplitude Modulation) system, or a data processing system.

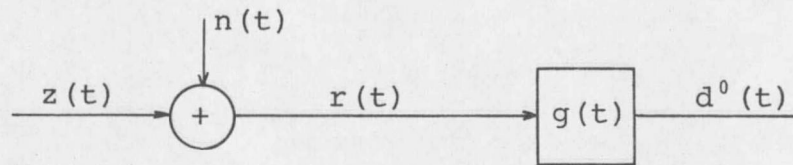
The distortion between $d(t)$ and $d^0(t)$ is due to n_1 , n_2 , and the sampling process. The object is to apply the innovations approach to find the optimal realizable pre- and postfilters for reconstructing $d(t)$ according to a mean-square-error criterion, and to discuss their physical interpretations.

1.2 Historical Summary

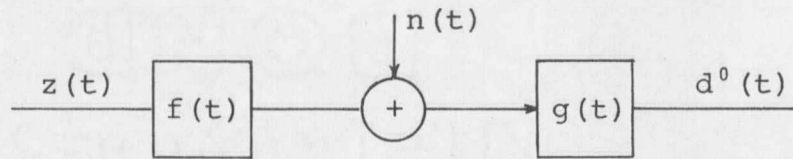
Previous work in this area can be grouped into three categories: 1. The frequency domain analysis, where the optimal filters are specified by their transfer functions which are the frequency domain representation of the impulse response of the optimal filters. 2. The state-space approach, where the optimal filters are specified by differential or difference equations which relate the

inputs to the outputs of the optimal filters. 3. Numerical methods; while both of the above two methods can be used in conjunction with numerical methods, the one that will be mentioned in this thesis is that of Kellog's (31). The three methods are discussed in more detail as follows.

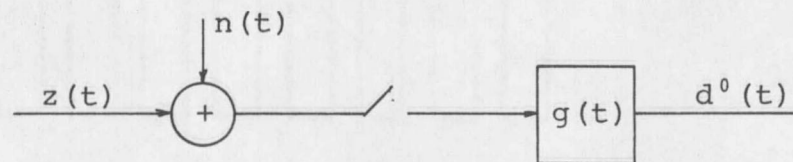
1. Frequency domain analysis: Ever since Wiener's work in linear least-squares estimation of continuous random processes (the system is shown in Fig. 1.2_a), the result has been extended by various people. Costas (10) studied the system shown in Fig. 1.2_b, where an unrealizable pre-transmission filter with a power constraint is added. The error reduction resulting from the usage of this prefilter were shown to be moderate for a simple signal such as a first-order G.M. (Gauss-Markov) process with a white additive noise. Franklin (47), and Tretter and Steiglitz (57) considered the system shown in Fig. 1.2_c, where a sampler is included. They also considered fixed-lag postfilters but did not provide a detailed quantitative error analysis. Katzenelson (30), Steiglitz (54), and Ruchkin et al. (50) investigate the system shown in Fig. 1.2_d, where quantization is included. Since the effect of the quantization process can always be



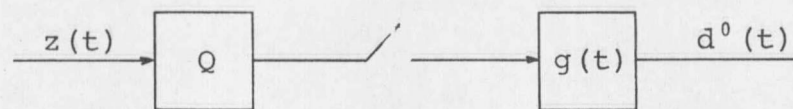
(a)



(b)

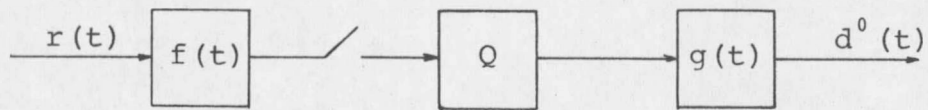


(c)

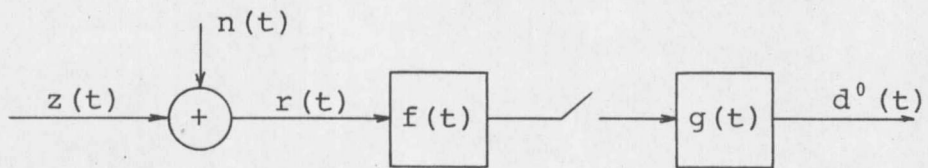


(d)

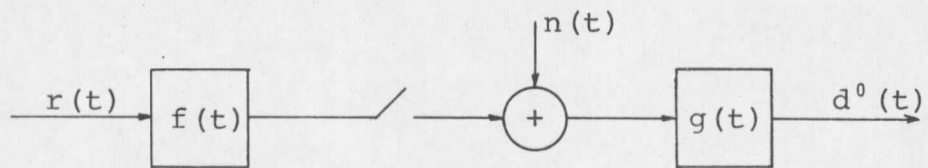
Figure 1.2 Models of communication systems.



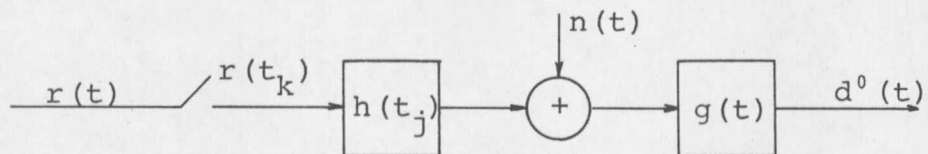
(e)



(f)



(g)



(h)

Figure 1.2 (cont.) Models of communication systems.

characterized as adding noise to the signal, this system reduces to that shown in Fig. 1.2_c. They studied the tradeoff between the number of quantization levels and the sampling rate.

Goodman and Drouilhet (19) studied the system shown in Fig. 1.2_e which includes a pre-sampling filter. They assume that the power spectrum of the input signal is bandlimited, and modeled the quantization noise as an uncorrelated white noise. Thus, this system reduces to the one considered by Costas. Robbins (49), Derusso (12), Chang (9), and others investigated the system shown in Fig. 1.2_f, where the filters are constrained to be zero-lag realizable filters. Quantization was not considered in their work. No error analysis has been found in the above three papers; hence, it is not quantitatively clear what role the optimal realizable prefilter plays. Brown (5) and Spilker (53) studied the same system with unrealizable filters. The optimal unrealizable filters found in this case are ideal bandpass filters. A similar result was obtained by Chan and Donaldson (7) in their work on the system shown in Fig. 1.2_g, where $n(t)$ can be quantization noise or/and channel additive noise. One of the

major contributions of their work was that they found that the jointly optimized unrealizable pre- and postfilters eliminate the aliasing effect which will be explained in Section 1.4. Hayes (20) studied the system shown in Fig. 1.2_h in which the prefilter was put behind the sampler. Thus, it is not surprising that this prefilter has no effect on the aliasing error.

The frequency-domain analysis is restricted to stationary processes, time-invariant systems, and steady-state conditions. Furthermore, it only applies to those stationary processes having power spectral densities that can be factored. The solutions for jointly optimized pre- and postfilters found by frequency-domain methods are in general rather messy when realizability is required. As the expressions for the transfer function of the optimal filters get more complicated, it becomes increasingly difficult to realize and analyze these filters.

2. State-space approach: Kalman (28) and Kalman and Bucy (29) originated the state-space approach to linear estimation problems as extensions of Wiener's work. Meditch (37), Fraser (13), and Mehra (38) extended this work to the smoothing problem. Kailath and others (23-27)

successfully applied the innovations approach to this problem. The innovations approach will be further discussed in Section 1.4. The systems considered by these researchers were either purely continuous or purely discrete, whereas in this thesis, a partly continuous and partly discrete system is considered.

Moore and Hetrakul (40) applied the state-space approach to a PAM system. They considered the problem of reconstructing the signal at discrete points in time based on continuous observations of the contaminated input. The problem considered in this thesis is how to reconstruct the analog signal continuously from discrete observations. Petersen (42) considered the system shown in Fig. 1.1_c with no word about quantization. He started the solution by proposing a particular model for the pre- and post-filters structure. The prefilter in his model consists of a continuous KSVF (Kalman State Vector Filter) followed by a coder which is simply a gain matrix. The optimal postfilter in his model can also be represented by a KSVF. This structure is intuitively appealing but needs to be generalized. Further, no error analysis is contained in Petersen's work. Also, no clue can be found about

suboptimum systems. For example, what is the situation when the optimal prefilter is absent?

The basic ideas behind the state-space approach are:

a. Instead of characterizing the signals by their covariance functions, or their power spectral densities, characterize them by differential or difference equations in the state variables with a white noise as the driving function. The signals can be represented as linear combinations of the state variables.

b. Instead of estimating the desired signal directly, estimate the state variables first. The best estimate of the signal is simply a linear combination of the best estimates of the state variables.

c. Instead of finding the impulse responses of the optimum filters, which are specified by integral equations (or the analogous equations with a summation term on one side of the equations in the discrete case), or by transfer functions in frequency domain, specify the optimal filters by differential or difference equations which relate the inputs to the outputs of the optimal filters.

The major advantages of the state-space approach are:

a. Both the differential or difference equations which specify the optimal estimates exhibit the characteristic of recursiveness. That means that it is not necessary to store all the past information. Thus, the amount of storage required is reduced. If one cannot solve the differential or difference equations analytically, he can use an analog or a digital computer.

b. This approach can be applied to nonstationary processes and time-variant systems. The only requirement is that the state models of the signals and the noises be given. Furthermore, transient behavior and nonperiodic sampling can be analyzed using the state-space approach.

The major problem encountered in the state-space approach is the modeling of the signal and noise by state variables. For a discussion of this modeling problem, see Kailath (24), Kailath and Geesey (27), and Moore and Anderson (39).

3. Numerical methods: both of the above two methods can be used in conjunction with the numerical methods. The one that is of interest here is that of

Kellog's (31). He investigated the systems of Figs. 1.2_f and 1.2_g numerically. The key points in his approach are:

a. The impulse response of the optimal filters are assumed to vanish outside a certain time interval.

b. The approximate impulse responses of the optimum filters are found by locating their sampled values at discrete points of time. For example, if the impulse response of the optimal filter is $f(t)$, the numerical method finds the values of $f(t)$ at $t=b \cdot k \cdot T$, where $k=0, 1, 2, \dots$, and $0 < b \leq 1$. One then plots $f(t)$ from this discrete set of sampled values to yield an approximate solution of $f(t)$.

While the numerical method is certainly one possible way to deal with the filtering problem, its difficulty lies in its computational complexity. Also, one has to be careful of the roundoff error in the numerical analysis itself.

1.3 Outline of Contents and Statements of Results

In Section 1.4, some of the basic concepts and tools, such as the orthogonality principle of linear estimation, used in this work are discussed. This discussion will provide some insight into the problem under consideration. In Section 1.5, the characterization of

quantization noise is discussed. The approximations made by other researchers about the quantization noise are given a heuristic justification in that section.

The method used in Chapters 2, 3, and 4 is the state-space and the innovations approach. The optimal filters found are realizable ones.

In Chapter 2, the reconstruction problem is studied. Since the system under consideration is partly discrete and partly continuous, it is necessary to have both continuous and discrete state models of the signal. The continuous model is assumed to be given, and the problem is to discretize it; this problem is solved in Section 2.2.

Both the zero-lag and the fixed-lag optimum filters are specified by the coefficients of differential equations. The amount of delay is constrained to be an integer multiple of the sampling period. Error expressions have also been found for both the filtering and the smoothing problem. For the smoothing problem, we defined an error-reduction factor, and a delay-efficiency factor; and investigate this problem using numerical examples. It is found that the stronger the noise, the greater the delay required for the performance of an optimum realizable

filter to approximate that of the optimum unrealizable filter. It is also found that error reduction due to the first sample of delay usually contributes a major portion of the overall error reduction, i.e., the error reduction using infinite-lag postfilter (the unrealizable postfilter). For the case of a first-order G.M. process with a white additive noise, it is found that the maximum error reduction that can be achieved by introducing delay is essentially accomplished within five samples of delay; i.e., the performance of the optimal realizable fixed-lag postfilter approaches that of the optimal unrealizable postfilter within five samples of delay in this case.

The design of the optimal prefilter with a given postfilter is investigated in Chapter 3. We first notice that if the number of sampled-data links can be made equal to the number of states, and n_2 in Fig. 1.1_c is absent; then the prefilter of the jointly optimized pre- and postfilters is simply a continuous KSVF. In the case when there is only one sampled-data link, the optimal prefilter can be represented as the combination of a continuous KSVF followed by a coder where the impulse response of the coder is in general an impulse train.

Joint optimization of the pre- and postfilters with a single sampled-data link is discussed in Chapter 4. In the case when n_2 in Fig. 1.1_c is absent, the jointly optimized prefilter can be represented as a cascade of a continuous KSVF followed by a constant coder, which is the model proposed by Petersen (42). In this case, since the optimal postfilter can be found easily by the method derived in Chapter 2 of this thesis once the constant coder is known, and since that KSVF is a well known filter, the problem reduces to finding the constant coder. Nonlinear programming has been successfully applied to find the constant coder in the steady state.

A summary and some extensions of the frequency-domain approach to this problem are presented in Chapter 5. The results found in previous chapters are checked here by examples. Numerical examples are then worked out and performances of various systems are compared.

For examples worked out in Chapter 5, it is found that in the absence of n_2 , the main function of the realizable prefilter is to prevent the aliasing of the primary observation noise (n_1 in Fig. 1.1).

Finally, in Chapter 6, the results of this work are summarized, and suggestions for further work in this area are provided.

1.4 Basic Concepts and Theorems

Some of the basic concepts and Theorems which are fundamental to the problem being investigated are discussed in this section. The discussion will provide some insight into the problem under consideration.

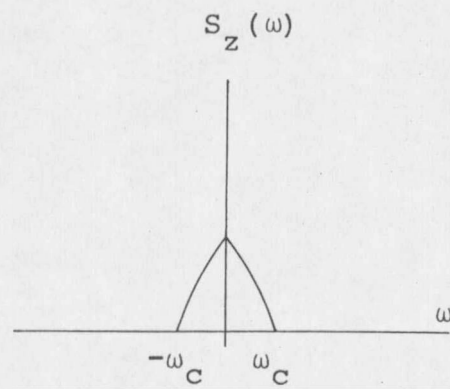
1.4.1 Sampling theorem and aliasing effect

Assume that we are given the sampled value of a certain continuous function $f(t)$ at $t = \dots, -T, 0, T, \dots$. It is natural to ask whether one will be able to find the value of $f(t)$ at an arbitrary t from this set of sampled values. The answer to this question is the content of the Sampling Theorem (56). The theorem states that if a signal does not contain any frequency components higher than ω_c radians per second, it is completely characterized by the values of the signal measured at instants of time separated by $*T = (\pi/\omega_c)^+$ seconds. If the signal is a

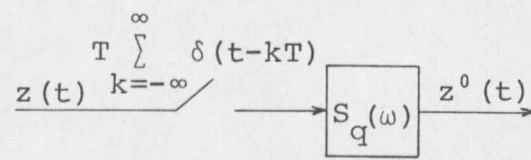
$$*(\pi/\omega_c)^+ = \lim_{\Delta \rightarrow 0} [(\pi/\omega_c) + \Delta]$$

random process, then the phrase "completely characterized" is in the sense of zero mean-square-error. To give a simple interpretation of this theorem, assume that a random process with a bandlimited spectral density as shown in Fig. 1.3_a is being sampled by the sampler shown in Fig. 1.3_b. The sampling period T is assumed to be equal to π/ω_c . The output of the sampler has a power spectrum shown in Fig. 1.3_c. An ideal low-pass filter with cut-off frequency $\omega = \omega_c$ is used to reconstruct the original signal. The output of this low pass filter has a power spectrum exactly equal to that of the original one, and the mean-square-error between z and z^0 is zero.

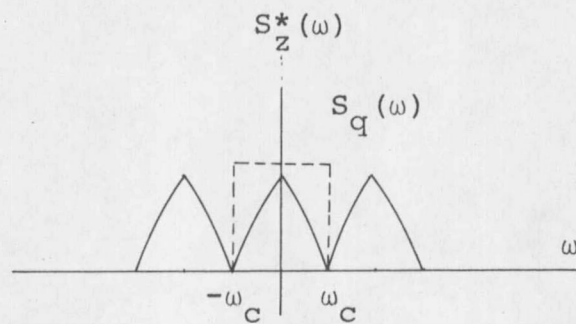
Signals with perfect bandlimited spectra do not occur in nature. Frequently, a random process might have a power spectrum as shown in Fig. 1.4_a. If one passes this signal through the sampler shown in Fig. 1.3_b, the output of the sampler will have a power spectrum which is the sum of the waveforms shown in Fig. 1.4_b. If we use an ideal low-pass filter to reconstruct the original signal, the distortion due to the sampling process will be observed as a result of two factors: the information lost outside of the frequency band of the filter, and the



(a)

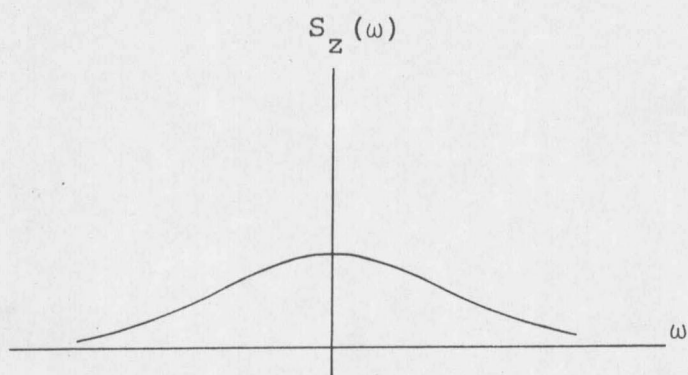


(b)

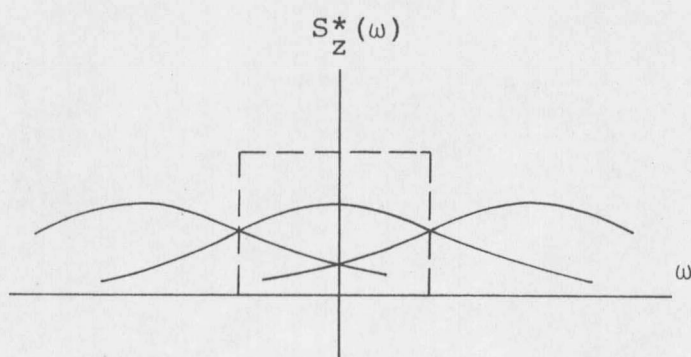


(c)

Figure 1.3 Reconstruction of a bandlimited signal after sampling.



(a)



(b)

Figure 1.4 Power spectrum of (a) a nonbandlimited signal;
(b) its sampled output.

tail effect of the induced frequency bands. The second effect is called aliasing. If an ideal low-pass filter with the same cut-off frequency as the postfilter is added to the system as a pre-sampling filter, the aliasing effect can be eliminated and the distortion can be reduced. As shown in Fig. 1.5, by using this combination of the pre- and postfilters, the aliasing effect is eliminated, whereas the information lost outside of the frequency band of the filters remains the same as in the case with a postfilter only. It can be proved easily (32) that the use of the prefilter in this case halves the distortion. It turns out that the combination of the ideal low pass pre- and postfilters in this case is optimal.

1.4.2 Orthogonality principle in linear estimation

This is the basic tool used in solving the linear-least-squares estimation problems. Assume that $r(t)$ is the observation process, $d(t)$ is the desired signal, and $\hat{d}(t)$ is the linear-least-squares estimate of $d(t)$ based on $r(\tau)$ for $T_1 \leq \tau \leq T_f$, the orthogonality principle states that the difference between $d(t)$ and $\hat{d}(t)$ is orthogonal to $r(\tau)$ for all $T_1 \leq \tau \leq T_f$, i.e.,

$$E[(d(t) - \hat{d}(t)) r(\tau)] = 0 \quad \text{for } T_1 \leq \tau \leq T_f.$$

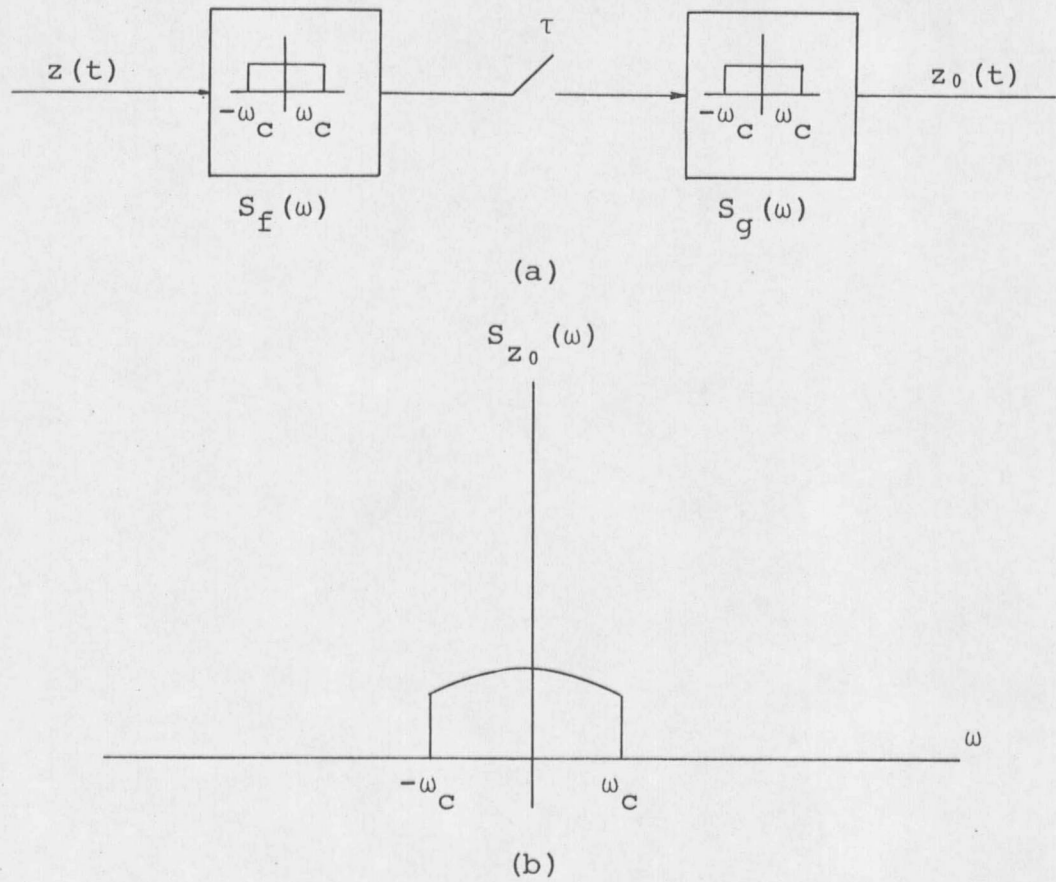


Figure 1.5 (a) Optimal pre- and postfilters for the signal with power spectrum shown in Fig. 1.4_a; (b) the power spectrum of the resulting output.

Depending on the values of T_i and T_f , applications of the orthogonality principle yields the solution to the linear filtering, predication, or smoothing problems.

To interpret this principle geometrically (58), see Fig. 1.6, where the shaded area represents those points that can be obtained by a linear operation on $r(\tau)$. One tries to select the vector lying within the shaded area so as to minimize the mean square error between it and $d(t)$. It is intuitively clear that the projection of $d(t)$ on the shaded area yields the optimum estimate of $d(t)$. Thus, the error is orthogonal to $r(\tau)$ and to any linear combination of $r(\tau)$ for $T_i \leq \tau \leq T_f$.

1.4.3 The innovations approach

The innovations approach was first defined by Kailath (23), but the approach itself has been used by various researchers before. This approach is equivalent to the whitening approach in the frequency-domain analysis. The word 'innovations' refers to the innovation processes, or equivalently, the whitening processes which represent the new information contains in some signal processes as time goes on.

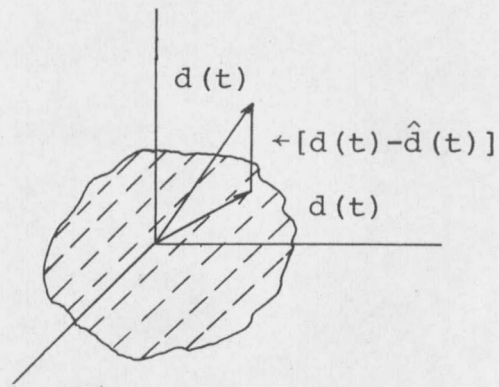


Figure 1.6 Geometric interpretation of the orthogonality principle in linear estimation.

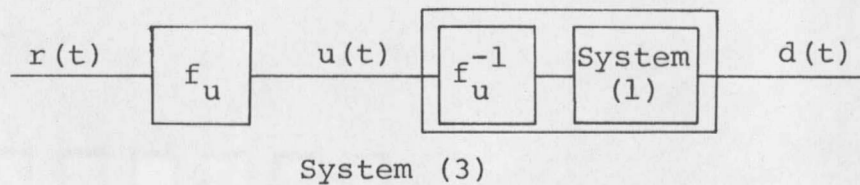
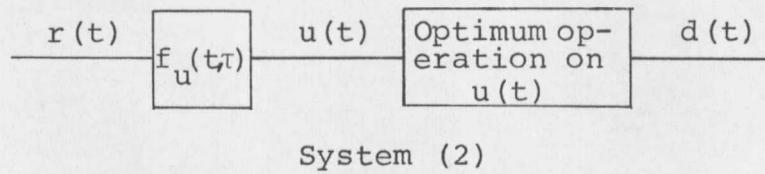
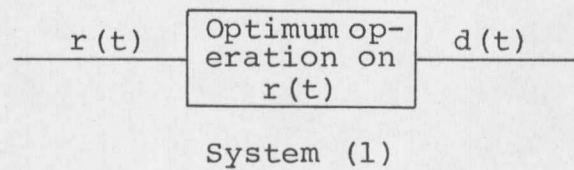


Figure 1.7 The innovations approach.

Kolmogorov (34) seems to be the first one to use this approach to discrete linear estimation problems. Bode and Shannon (4) used this approach to derive Wiener's result. Kalman (28) applied this method in deriving his result of the state-space approach to linear discrete estimation. Recently, Kailath (23,24), Kailath and Frost (25), and Kailath and Geesey (26) applied this method to derive Kalman and Bucy's (29) results, and to extend these results to smoothing, and nonlinear estimation problems.

There are two basic concepts imbedded in the innovations approach:

1. Since a signal with a flat spectral density is easy to deal with both in the frequency and in the time domains, one first converts the original observation process into a white noise by an invertible filter. The resulting white noise is to be called the innovation process, and the invertible whitening filter is to be called the innovation filter. For a discussion of the existence of such an innovation process, refer to Kailath (24).

2. If there does exist such an invertible innovation filter, then the original process and the innovation process are equivalent to each other in the sense

that they contain the same statistical information. Thus, the optimum estimation of the desired signal based on the innovation process will be the same as the one based on the original observation process.

A simple proof of the above statement, due to Van Trees (58), is given below. The proof hinges on the invertibility of the innovation filter.

System (1) shown in Fig. 1.7_a operates on $r(t)$ to yield an output which is optimum according to some fidelity criterion. In system (2), one first operates on $r(t)$ with an invertible innovation filter $f_u(t, \tau)$ to obtain the innovation process $u(t)$. One then operates on $u(t)$ to yield an output which is optimum according to the same fidelity criterion used in system (1). We claim that the two systems have the same performance. Apparently, system (2) cannot perform better than system (1); otherwise, it will contradict the statement that system (1) is the optimum system given $r(t)$ as its input. We claim that system (2) cannot perform worse than system (1) either. If it does, we can design the third system shown in Fig. 1.7_c, recalling that f_u is invertible. System (3) is equivalent to system (1). That contradicts the statement

that system (2) is the optimum system given $u(t)$ as its input. Thus, the proof is completed.

The major advantages of the innovations approach are:

1. The derivation of the optimum filter is simpler in general.
2. The smoothing problem can not only be handled more easily using the innovations approach, but will yield to a better physical interpretation.
3. The innovations approach can be extended to deal with nonlinear estimation problems (14).

1.4.4 The smoothing problem

In general, the greater the delay allowed, the closer the performance of an optimum realizable filter will approximate that of the corresponding optimum unrealizable filter (47,57). It is of interest to investigate quantitatively the efficiency of this delay factor in a sampled-data system. For example, assume that the signal under consideration is the first-order Gauss-Markov process with autocorrelation function $R_z(t) = \exp(-a|t|)$, and the system under consideration is shown in Fig. 1.3_b. If one uses the optimum fixed-lag realizable reconstruction

filter, the mean square error between $z(t)$ and $z^0(t)$ is determined by the reconstruction density function $\psi(t)$ where

$$\psi(t) = \begin{cases} 0 & -\infty < t \leq -(T-D) \\ \frac{e^{-Da}}{(1-e^{-2aT})^{1/2}} [e^{at} - e^{-2a(T-D)} e^{-at}] & -(T-D) < t \leq D \\ (1 - e^{-2aT})^{1/2} e^{Da} e^{-at} & D < t < \infty. \end{cases} \quad (1.1)$$

The general shape of $\psi(t)$ specified by Eq. (1.1) is shown in Fig. 1.8. Depending on whether $D=0, >0$, or <0 , we are dealing with the zero-lag, fixed-lag, or prediction problems. It is obvious from Fig. 1.8 that a delay of $D=T$ is all that is needed for the performance of this optimal realizable filter to equal that of the optimum unrealizable one. Shown in Fig. 1.9 are the following: a plot of the error versus aT for the optimum zero-lag realizable postfilter (reconstruction filter), for the optimum unrealizable postfilter in the absence of the prefilter; and for the jointly optimized unrealizable pre- and postfilters (see Section 1.4.1 of this thesis). The jointly optimized zero-lag realizable prefilter is simply a straight line in this case. Fig. 1.9 gives some

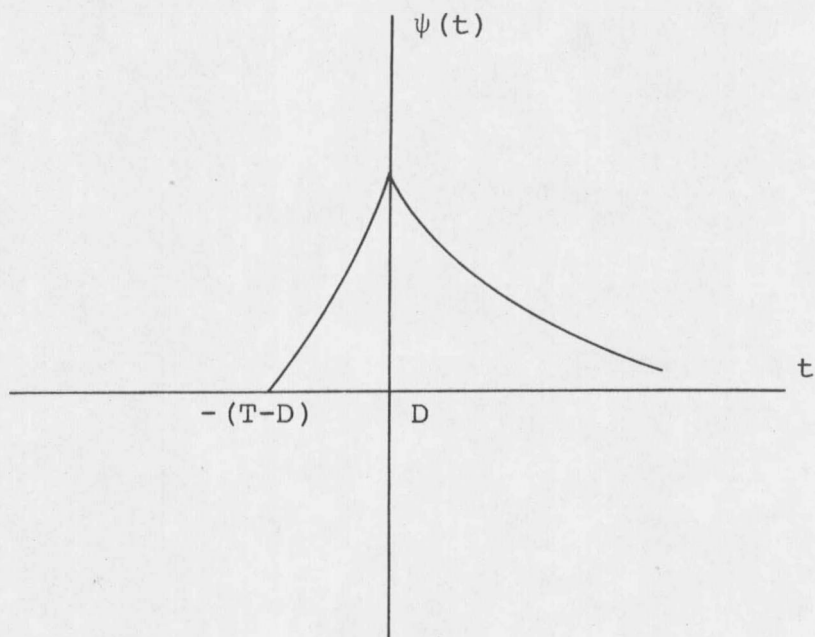


Figure 1.8 The general shape of the 'Reconstruction density function' of a first-order G.M. (Gauss-Markov) process.

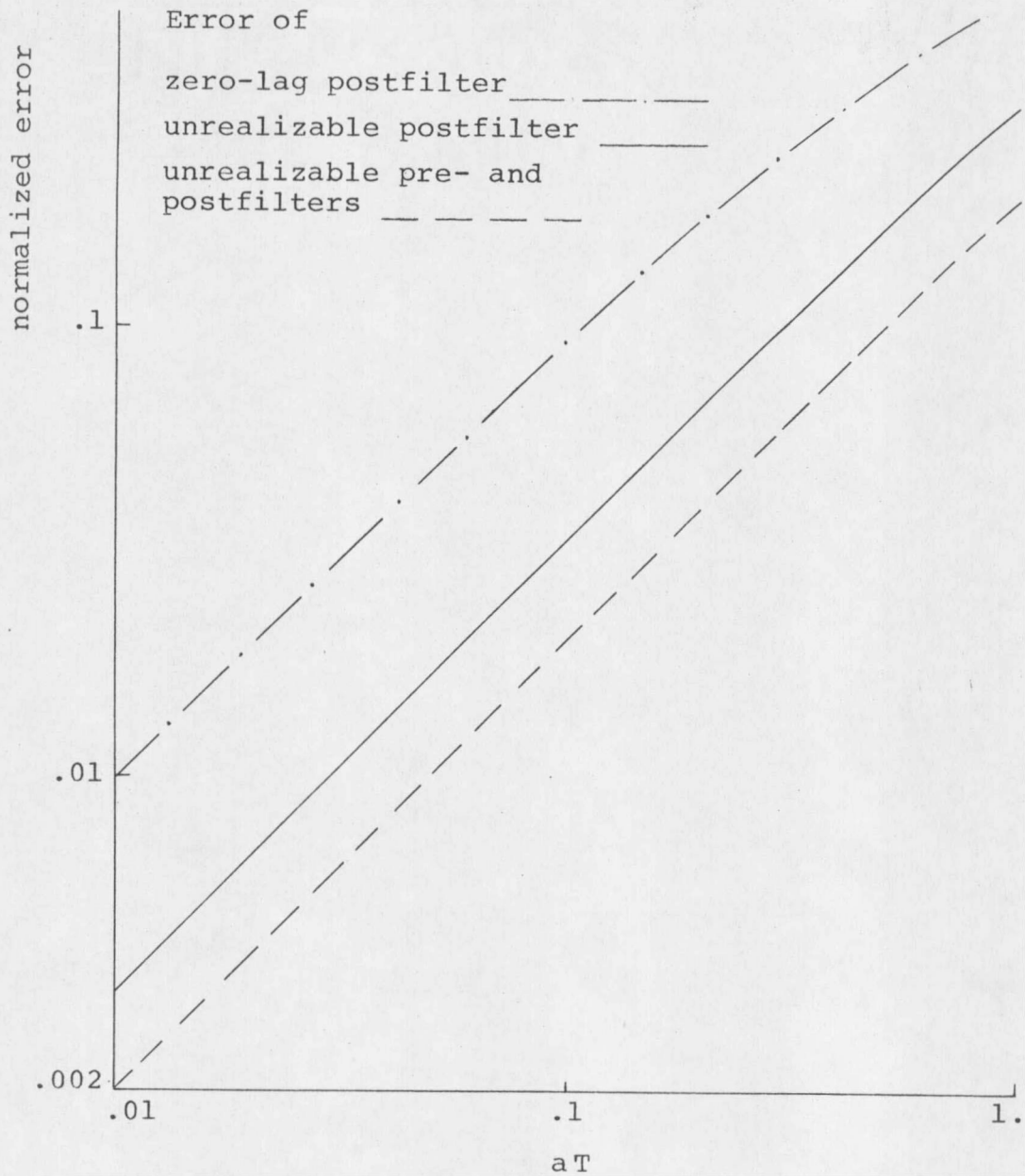


Figure 1.9 Normalized reconstruction error of a first-order G.M. process after sampling.

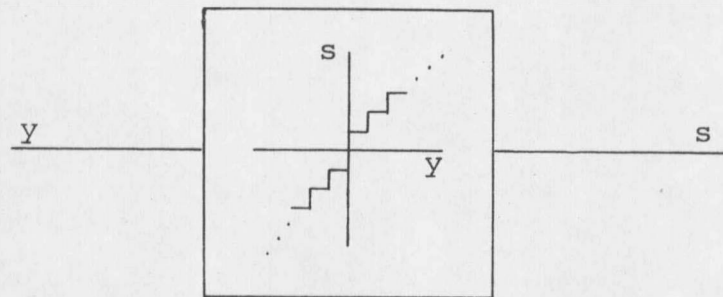
indication about how much error reduction can be achieved by using unrealizable filters, and by using a prefilter.

1.5 Characterization of the Quantization Noise

The effect of including a quantizer in the system can always be characterized as adding noise to the signal being quantized. Assume that y is the input of the quantizer, s is the output, and n is the quantization noise as shown in Fig. 1.10. The model usually used to characterize the quantizer is

$$s = y + n \quad (1.2)$$

Various researchers have studied the characterization of this quantization noise. Widrow (60) and Katzenelson (30), for example, have shown that for many cases, $n(t)$ can be approximated by an uncorrelated noise with an autocorrelation function $R_n(\tau)$ which decreases with τ much faster than does the autocorrelation function of the input signal. The approximate variance of the quantization noise depends on the number of the quantization levels only. Ruchkin (50) and others have shown that except for very coarse quantizing, and very small sampling period (with periodic sampling), the power spectrum of the



y^i : end points of the M input ranges

s^i : output level for $y^{i-1} \leq y < y^i$

Figure 1.10 An equal-interval quantizer.

sampled version of the quantization noise for a large class of y can be approximated by

$$S_n^*(f) \approx T R_n(0) \quad (1.3)$$

where

$$R_n(0) = \sum_{i=1}^M \int_{y^{i-1}}^{y^i} (s^i - y) P_r(y) dy \quad (1.4)$$

y^i is the i th end point of the M input ranges; s^i is the output level for $y^{i-1} \leq y < y^i$; and $P_r(y)$ is the probability density function of y . But since $S_n^*(f) = T \cdot \sum_{k=-\infty}^{\infty} e^{-j2\pi f kT} R_n(kT)$

(1.3) essentially says that $R_n(kT)$ for $k \neq 0$ is negligible in comparison with $R_n(0)$; i.e., $n(kT)$ is approximately a discrete white noise with

$$R_n(k-j) \approx R_n(0) \delta_{kj} \quad (1.5)$$

The validity of (1.5) will be examined later using a time-domain analysis.

Bussgang (6) has found that for Gaussian inputs, the cross correlation between the input and the output of the quantizer is given by

$$E[y(t) \cdot s(\tau)] = a_1 E[y(t) \cdot y(\tau)] = a_1 R_y(t-\tau), \quad (1.6)$$

where

$$a_1 = \sum_{i=1}^M s^i \int_{y^{i-1}}^{y^i} \frac{y}{\sqrt{2\pi}} \exp(-x^2/2) dx. \quad (1.7)$$

Kellog (31) has found that the autocorrelation function of the output of the quantizer can be written as

$$R_s(\tau) = \sum_{i=1}^M (s^i / \sqrt{2\pi}) \int_{y^{i-1}}^{y^i} h(\rho, \eta) \exp(-\eta^2/2) d\eta, \quad (1.8)$$

where

$$h(\rho, \eta) = \sum_{i=1}^M (s^i / 2) \left[\operatorname{erf} \left(\frac{y^i - \rho\eta}{\sqrt{2(1-\rho^2)}} \right) - \operatorname{erf} \left(\frac{y^{i-1} - \rho\eta}{\sqrt{2(1-\rho^2)}} \right) \right], \quad (1.9)$$

$$\rho(\tau) = R_Y(\tau) \quad (1.10)$$

and

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \exp(-t^2) dt. \quad (1.11)$$

Equations (1.8) and (1.9) can be solved by numerical integration.

Curry (11) has found that if we characterize the quantizer by

$$s = c \cdot y + n_c, \quad (1.12)$$

instead of the commonly used $s=y+n$ (i.e., $c=1$); then by letting $c=a_1$, the signal to noise ratio ($a_1^2 E[y^2]$ to $E[n_c^2]$) is maximized, and the quantization noise $n_c(t)$ is as white as possible. Furthermore, $n_c(t)$ is uncorrelated with the input y , i.e., $E[y \cdot n_c] = 0$. For our purposes either (1.2) or (1.12) is adequate to model the quantizer. With $\rho(\tau)$ given, one can find $R_s(\tau)$ from (1.8) and (1.9); then $R_n(\tau)$ and $R_{n_c}(\tau)$ can be found from (1.2) and (1.12), respectively.

Assume that the input to the quantizer has an autocorrelation function given by $R_y(\tau) = e^{-a|\tau|}$. Using the procedure outlined above with $c=a_1$, plots of $R_y(\tau)$, $R_n(\tau)$, and $R_{n_c}(\tau)$ versus $a\tau$ for a Max equal-interval quantizer (35) with 2, 4, and 8 quantization levels are shown in Fig. 1.11 and Table 1.1. Let us investigate the ratio $R_n(kT)/R_n(0)$ and $R_{n_c}(kT)/R_{n_c}(0)$ for $k \neq 0$ since those values indicate when the sampled version of the quantization noise can be approximated as a discrete white noise. It is obvious from Fig. 1.11 that the larger the sampling period T is, the smaller the ratios $R_n(kT)/R_n(0)$ and $R_{n_c}(kT)/R_{n_c}(0)$ for $k \neq 0$ become. To see the effect of the number of quantization levels on those ratios, a plot of

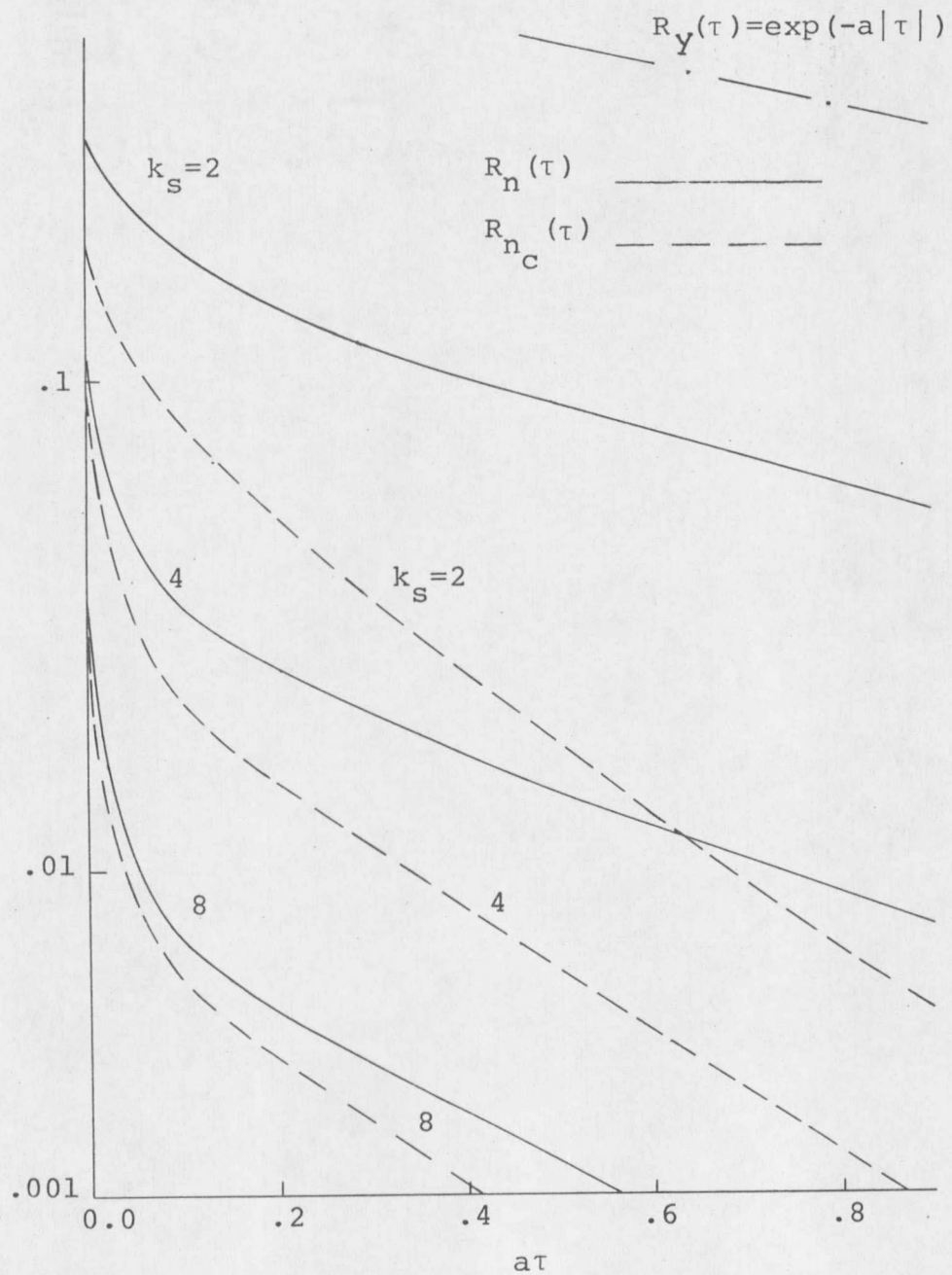


Figure 1.11 Autocorrelation functions of quantization noises.

Table 1.1 Autocorrelation functions of the input, output, and the quantization noise of a Max. equal-interval quantizer

k_s	k	$R_Y(kT)$	$R_S(kT)$	$R_{n_c}(kT)$	$R_n(kT)$
2	0	1.0	0.6368	0.2314	0.3634
	1	0.9048	.4585	.0917	.2111
	3	.7408	.3382	.0379	.1357
	5	.6065	.2641	.0183	.0983
	7	.4966	.2107	.0094	.0749
	9	.4066	.1697	.0048	.0586
4	0		.8814	.1050	.1191
	1		.7256	.0230	.0358
	3		.5864	.0112	.0217
	5		.4770	.0060	.0146
	7		.3889	.0033	.0103
	9		.3175	.0018	.0075
8	0		.9632	.0367	.0381
	1		.8440	.0057	.0070
	3		.6892	.0029	.0039
	5		.5635	.0015	.0024
	7		.4609	.0008	.0015
	9		.3771	.0004	.0010

* The number of the quantization levels.

those ratios versus the number of quantization levels (up to 32) for a Max equal-interval quantizer is shown in Fig. 1.12. It is shown in Fig. 1.12 that the larger the number of quantization levels, the whiter the quantization noise becomes. It is also shown in Fig. 1.12 that the difference between the modeling of the quantizer with $c=1$ and $c=a_1$ in Eq. (1.12) (Eq. (1.12) reduces to Eq. (1.2) with $c=1$) becomes less significant as the number of quantization levels increases. This is also evident from the fact that a_1 approaches 1 as the number of quantization levels increases. Whether the quantization noise specified by Eq. (1.2) can be approximated by an uncorrelated noise depends on the value of a_1 . The closer a_1 approaches 1, the less correlation there will be between the input signal and the quantization noise. These figures also provide a heuristic look at the result of Katzenelson (30) and others about the fact that the autocorrelation function of the quantization noise specified in (1.2), $R_n(\tau)$, decreases with τ much faster than that of the input signal.

Thus, in most of the practical cases, it is reasonable to assume that the sampled version of the quantization noise specified by (1.12) with $c=a_1$ can be approximated by

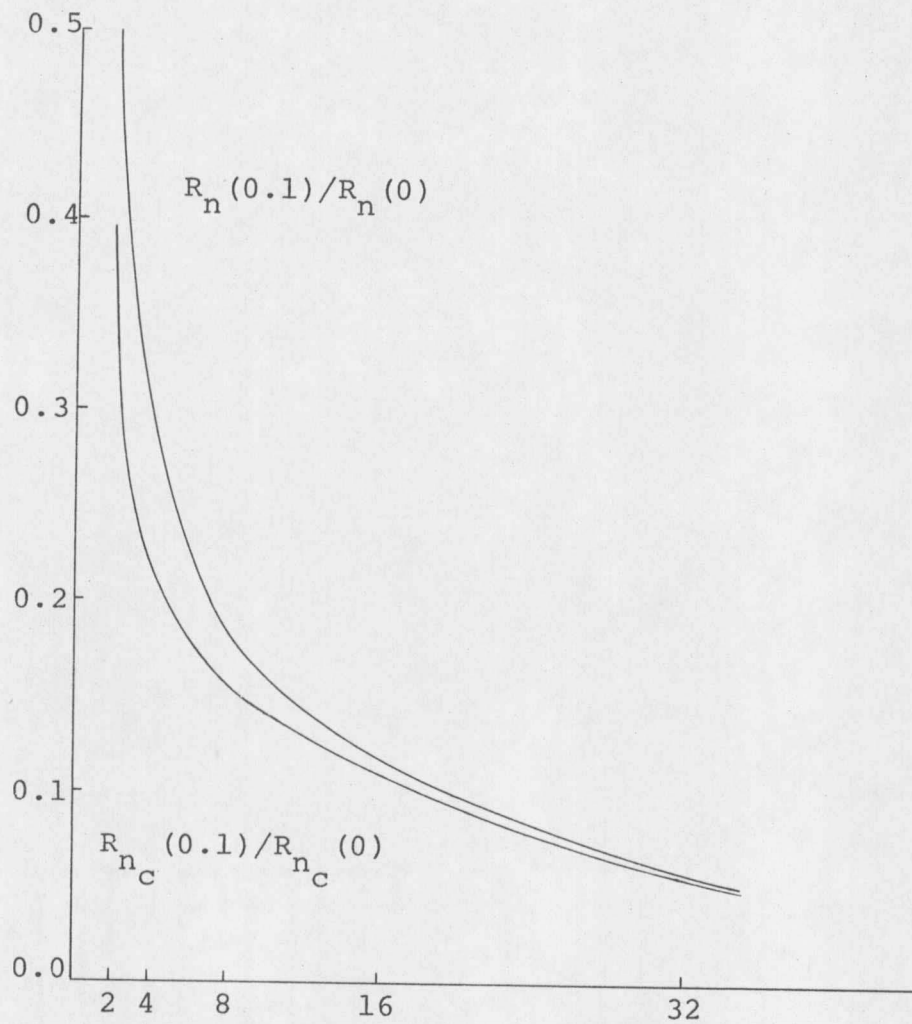


Figure 1.12 Ratios of $R_n(0.1)$ (or $R_{n_c}(0.1)$) to $R_n(0)$ ($R_{n_c}(0)$) versus the number of quantization levels.

a discrete white noise uncorrelated with the input signal and with its autocorrelation function given by

$$R_{n_c}(kT-jT) = R_{n_c}(0) \delta_{kj}, \quad (1.13)$$

where

$$R_{n_c}(0) = \sum_{i=1}^M \int_{y^{i-1}}^{y^i} (s^i - a_1 y)^2 P_r(y) dy. \quad (1.14)$$

Notice that even though the quantization noise can be approximated by an uncorrelated additive noise, it is still dependent on the input signal since $R_{n_c}(0)$ depends on the power of the input signal.

Chapter 2

THE OPTIMAL POSTFILTER

2.1 Problem Statement and Preliminary Remarks

The system under consideration in this chapter is depicted in Fig. 2.1. Shown in Fig. 2.1_a is the system that generates the signal $z(t)$ to be sampled, possibly quantized, and then transmitted to the receiving end. Fig. 2.1_b specifies the relation between the desired signal and the state vector. Shown in Fig. 2.1_c is the overall system being considered; where the sampling is periodic; $n(k)$ can be quantization noise or/and channel additive noise; and $g(t, kT)$ is the postfilter to be determined. The systems equations are

$$\dot{x}(t) = A(t) x(t) + \Gamma(t) w(t), \quad (2.1)$$

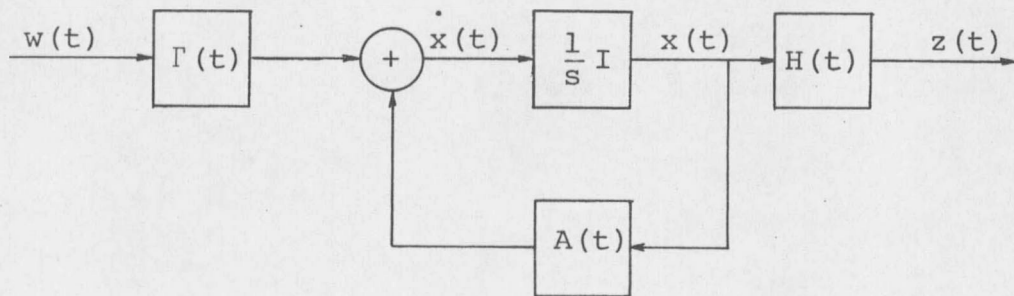
$$z(t) = H(t) x(t), \quad (2.2)$$

$$r(kT) = z(kT) + n(kT), \quad (2.3)$$

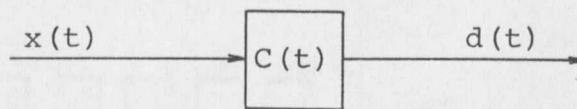
and the desired signal is given by

$$d(t) = C(t) x(t). \quad (2.4)$$

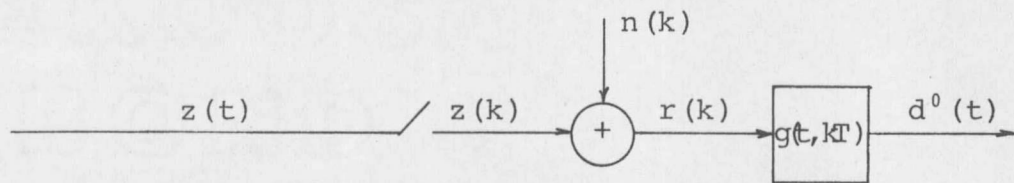
The T in Eq. (2.3) is the sampling period. In order to simplify notation, we will occasionally denote kT by k from here on if it does not cause confusion. The state



(a)



(b)



(c)

Figure 2.1 The design of the postfilter: (a) the message model; (b) the generation of the desired signal; (c) the system.

vector $x(t)$ is an n -vector; $w(t)$ is an ℓ -vector; $d(t)$ is an p -vector, and $r(t)$, $z(t)$, and $\eta(t)$ are i -vectors. The dimensions of $A(t)$, $\Gamma(t)$, $H(t)$, and $C(t)$ are $n \times n$, $n \times \ell$, $i \times n$, and $p \times n$, respectively. Let

$$R_{\xi\eta}(t, \tau) = E[\xi(t) \eta'(\tau)],$$

$$R_{\eta}(t, \tau) = E[\eta(t) \eta'(\tau)]$$

and

$$V_{\eta}(t) = R_{\eta}(t, t) = E[\eta^2(t)],$$

where ξ, η are any related signal.

Prior statistics of the system are

$$E[w(t)] = E[n(k)] = E[x(iT)] = 0 \quad (2.5)$$

$$R_w(t, \tau) = V_w(t) \delta(t - \tau) \quad (2.6)$$

$$R_{xw}(\ell_i T, \tau) = 0 \text{ for } \tau > \ell_i T, \quad (2.7)$$

$$R_{wn}(t, kT) = 0 \quad (2.8)$$

and the initial covariance $V_x(\ell_i T)$ is known, where $\ell_i T$ denotes the initial time the estimation begins. Let

$\hat{\xi}(t/\ell_i, \ell_f, \eta(k)) \triangleq$ linear least-squares estimate of $\xi(t)$ based on the observation of the $\eta(k)$'s for $\ell_i \leq k \leq \ell_f$. Depending on whether $m > \ell_f$, $m = \ell_f$, or $m < \ell_f$, where $m = \lfloor t/T \rfloor$, we are dealing with the prediction, filtering, and

smoothing problems. A steady-state estimation corresponds to $\ell_i \rightarrow -\infty$. If an unrealizable filter with an infinite memory capacity is used as the postfilter, $\ell_f = -\ell_i = \infty$. Either one of the above two situations is usually assumed a prior condition in the frequency-domain approach to this problem. In some cases, one or more of the quantities ℓ_i, ℓ_f , or/and $\eta(k)$ can be omitted from the notation defined above when there is no need to specify them. For example, when there is no need to specify ℓ_i, ℓ_f and $\eta(k)$, $\hat{\xi}(t)$ will be used to denote $\hat{\xi}(t/\ell_i, \ell_f, \eta(k))$. When there is no need to specify ℓ_i and $\eta(k)$, $\hat{\xi}(t/\ell_f)$ will be used to denote $\hat{\xi}(t/\ell_i, \ell_f, \eta(k))$. In this work, the initial time $\ell_i T$ does not have to be $-\infty$, but $\ell_i T < t$ and $\ell_i < \ell_f$. Thus, finite time observations and transient behaviors can be analyzed using this state-space approach.

At a first glance, it seems that the problem can be solved in two steps:

1. Find $\hat{d}(k)$ from the $r(j)$'s using a discrete filter.
2. Find $\hat{d}(t)$ from the $\hat{d}(k)$'s using a filter which is the optimal reconstruction filter if the $d(k)$'s are the inputs.

This is indeed the case if the optimum filter under consideration is an unrealizable filter with an infinite memory capacity. To prove this point, assume $d(t)=z(t)$ for easy presentation, and let

$$y(t) = \sum_{k=-\infty}^{\infty} g_0(t, kT) \hat{z}(k), \quad (2.9)$$

where the quantities $g_0(t, kT)$'s satisfy the equation

$$E\left\{ \left[z(t) - \sum_{k=-\infty}^{\infty} g_0(t, kT) z(k) \right] z'(j) \right\} = 0 \quad (2.10)$$

for all t, j ; i.e., $g_0(t, kT)$ represents the impulse response of the optimum filter to reconstruct $z(t)$ from the $z(k)$'s. $\hat{z}(k)$ in Eq. (2.9) is the optimum discrete estimation of $z(k)$ from the $r(j)$'s for all j , i.e.,

$$E\{[z(k) - \hat{z}(k)] r'(j)\} = 0 \quad \text{for all } j, k. \quad (2.11)$$

Let

$$\tilde{z}(\cdot) = \xi(\cdot) - \hat{\xi}(\cdot). \quad (2.12)$$

Then (2.11) becomes

$$E[\tilde{z}(k) r'(j)] = 0 \quad \text{for all } j, k. \quad (2.13)$$

We want to prove that $y(t)=\hat{z}(t)$; i.e., to prove that (from the orthogonality principle)

$$E\{[z(t) - y(t)] r'(j)\} = 0 \quad \text{for all } j.$$

From (2.12) it follows that

$$\hat{z}(k) = z(k) - \tilde{z}(k). \quad (2.14)$$

Substituting (2.14) into (2.9) yields

$$\begin{aligned} y(t) &= \sum_{k=-\infty}^{\infty} g_0(t, kT) [z(k) - \tilde{z}(k)] \\ &= \sum_{k=-\infty}^{\infty} g_0(t, kT) z(k) - \sum_{k=-\infty}^{\infty} g_0(t, kT) \tilde{z}(k). \end{aligned} \quad (2.15)$$

Thus,

$$\begin{aligned} E\{[z(t) - y(t)] r'(j)\} &= E\left\{[z(t) - \sum_{k=-\infty}^{\infty} g_0(t, kT) z(k) \right. \\ &\quad \left. + \sum_{k=-\infty}^{\infty} g_0(t, kT) \right. \\ &\quad \left. \times z(k)] r'(j)\right\} \\ &= E\left\{[z(t) - \sum_{k=-\infty}^{\infty} g_0(t, kT) z(k)] r'(j)\right\} \\ &\quad + \sum_{k=-\infty}^{\infty} g_0(t, kT) E[\tilde{z}(k) r'(j)] \end{aligned} \quad (2.16)$$

The second term on the right-hand side of (2.16) is zero from (2.13). For the first term,

$$\begin{aligned}
E\{[z(t) - \sum_{k=-\infty}^{\infty} g_0(t, kT) z(k)] r'(j)\} &= E\{[z(t) \\
&- \sum_{k=-\infty}^{\infty} g_0(t, kT) z(k)] \\
&\times [z'(j) + n'(j)]\} \\
&= E\{[z(t) - \sum_{k=-\infty}^{\infty} g_0(t, kT) z(k)] z'(j)\} + \\
&E\{[z(t) - \sum_{k=-\infty}^{\infty} g_0(t, kT) z(k)] n'(j)\}.
\end{aligned}
\tag{2.17}$$

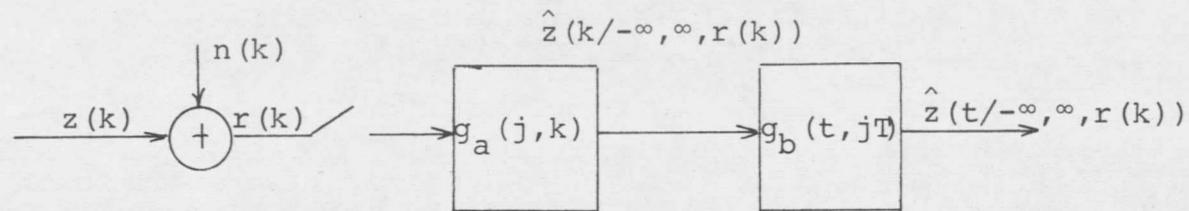
The first term on the right-hand of (2.17) is zero from (2.10). The second term is also zero since $n(k)$ is assumed to be uncorrelated with the signal. Thus, from the above arguments,

$$E\{[z(t) - y(t)] r'(j)\} = 0 \quad \text{for all } j; \tag{2.18}$$

i.e., $y(t) = \hat{z}(t)$ from the orthogonality principle. The postfilter that implements the two steps mentioned before is shown in Fig. 2.2.

If realizability is required, let $\hat{z}(t)$ denote the zero-lag optimum estimate of z based on the $r(k)$'s for $k \leq m$; i.e., $z(t) = \hat{z}(t/l_1, m, r(k))$, where $m = \langle t/T \rangle$.

Eq. (2.10) becomes



$g_a(j, k) \triangleq$ the optimal unrealizable discrete filter to reconstruct $z(j)$ from the $r(k)$'s.

$g_b(t, jT) \triangleq$ the optimal unrealizable continuous filter to reconstruct $z(t)$ from the $z(j)$'s.

Figure 2.2 A two-filter structure for the optimal unrealizable postfilter.

$$\dot{E}\left\{\left[z(t) - \sum_{k=-\infty}^m g_0(t, kT) z(k)\right] z'(j)\right\} = 0 \quad \text{for } j \leq m, \quad (2.19)$$

and (2.13) becomes

$$E[\hat{d}(k) r'(j)] = 0 \quad \text{for } j \leq k. \quad (2.20)$$

Following the same procedure used for the case of unrealizable filter, (2.16) becomes

$$\begin{aligned} \dot{E}\{[z(t) - y(t)] r'(j)\} &= \dot{E}\{[z(t) \\ &\quad - \sum_{k=-\infty}^m g_0(t, kT) z(k)] r'(j)\} \\ &\quad + \sum_{k=-\infty}^m g_0(t, kT) E[\hat{d}(k) r'(j)] \end{aligned} \quad (2.21)$$

If $\dot{E}\{[z(t) - y(t)] r'(j)\}$ is zero for all $j \leq m$, the two steps mentioned before will yield the optimal zero-lag estimate of $z(t)$. The first term on the right-hand side of (2.21) is zero for all $j \leq m$ from (2.19) and the fact that $n(j)$ is uncorrelated with $z(t)$. But the second term might not be zero since $E[\hat{z}(k) r'(j)]$ is guaranteed zero only if $j \leq k$, whereas the summation is from $k = -\infty$ to $k = m$. If $z(t)$ is a first-order Gauss-Markov process, and $A(t) = -a$ and $H(t) = 1$ in Equations (2.1) and (2.2), respectively, then

$$g_0(t, kT) = \begin{cases} \exp(-a(t-kT)) & \text{if } k=m \\ 0 & \text{otherwise} \end{cases} \quad (2.22)$$

Substituting (2.22) into (2.21) yields

$$\begin{aligned} E\{[z(t) - y(t)] r'(j)\} &= E\{[z(t) - g_0(t, mT) \cdot z(m)] r'(j)\} \\ &\quad + g_0(t, mT) E[\tilde{d}(m) r'(j)]. \end{aligned} \quad (2.23)$$

It is easy to prove that $E\{[z(t) - y(t)] r'(j)\} = 0$ for all $j \leq m$ in this case. Thus, the two steps mentioned before will yield the optimum zero-lag estimate of $z(t)$ in the case when $z(t)$ is a first-order Gauss-Markov process. The fact that $z(t) = x(t)$ is the state variable in the case when $z(t)$ is a first-order Gauss-Markov process whereas for a higher order signal, $z(t)$ is usually a linear combination of the state variables suggests that instead of estimating $z(k)$ alone, one should estimate the entire state vector $x(k)$ at first in the case with a realizable filter. This point will become clear later.

Since results pertaining to both the continuous and discrete KSVF (Kalman State Vector Filter) are needed as backgrounds in the following work, summaries and discussions of those results are presented in Appendices A and B, respectively (51).

The system under consideration is partly continuous and partly discrete. In order to use the state-space approach, one must discretize the continuous state equation given by (2.1) first. The discretization of the continuous state equation, and some discussions in this respect are given in Appendix C (22).

Prediction and filtering are considered in the following section. Smoothing is considered in Section 2.3.

2.2 Optimum Filtering and Prediction

In this section, let

$$\hat{\xi}(t/\langle \rho/T \rangle) = \hat{\xi}(t/\ell_i, \langle \rho/T \rangle, r(k)) = \text{linear least-squares prediction (or filtering) of } \xi(t) \text{ based on the } r(k) \text{'s for } \ell_i \leq k \leq \langle \rho/T \rangle, \text{ where } \rho \leq t.$$

and let

$$\hat{\xi}(t) = \hat{\xi}(t/\langle t/T \rangle) = \text{linear least-squares zero-lag estimate of } \xi(t) \text{ based on the } r(k) \text{'s.}$$

From (C.1), it follows that

$$\hat{x}(t/\langle \rho/T \rangle) = \Phi(t, \rho) \hat{x}(\rho) + \int_{\rho}^t \Phi(t, v) \Gamma(v) \hat{w}(v/\langle \rho/T \rangle) dv \quad (2.24)$$

Since $w(v)$ is independent of $x(j)$ and is uncorrelated with $n(j)$ for $jT < v$,

$$\hat{w}(v/\langle \rho/T \rangle) = 0 \quad (2.25)$$

Hence (2.24) becomes

$$\hat{x}(t/\langle \rho/T \rangle) = \Phi(t, \rho) \hat{x}(\rho). \quad (2.26)$$

Thus, optimum prediction of the desired signal can be found easily from Eq. (2.26) once the optimum zero-lag estimate of the desired signal is known. Hence, our efforts will be concentrated on solving the filtering problem in the remainder of this section. Two alternative derivations which lead to two different forms of the optimal zero-lag postfilter are presented in the following two subsections.

2.2.1 A two-filters structure

A straightforward derivation of the optimal postfilter, which results in a two-filter structure, is given as follows. Substituting $\rho = mT$ into (2.26) where $m = \langle t/T \rangle$, yields

$$\hat{x}(t) = \Phi(t, mT) \hat{x}(mT), \quad \text{for } mT \leq t < (m+1)T \quad (2.27)$$

or in another form,

$$\hat{x}(t) = \sum_{k=\ell_i}^{\infty} \Phi(t, kT) \hat{x}(k) [u_s(t-kT) - u_s(t-(k+1)T)], \quad (2.28)$$

where $u_s(t)$ is the unit step function. Thus, the optimal postfilter can be represented by the cascade of a discrete KSVF which generates $\hat{x}(m)$'s followed by a filter with impulse response given by

$$g_2(t, mT) = \begin{cases} \Phi(t, mT) & \text{for } mT \leq t < (m+1)T \\ 0 & \text{otherwise.} \end{cases} \quad (2.29)$$

The $g_2(t, mT)$ given by (2.29) is the impulse response of the optimal zero-lag postfilter for reconstructing $x(t)$ from the $x(k)$'s. Finally, $\hat{d}(t)$ is obtained by the equation

$$\hat{d}(t) = C(t) \hat{x}(t) \quad (2.30)$$

The block diagram of this system is shown in Fig. 2.3.

It is interesting to notice the differences between the system shown in Fig. 2.3 and the system shown in Fig. 2.2, where both systems have a two-filter structure. The first filter in the system shown in Fig. 2.2 is to obtain the infinite-lag optimal estimate of the desired signal $d(k)$'s at discrete points in time. Whereas when realizability is required, the first filter in the system shown in Fig. 2.3 is to obtain the zero-lag optimal estimate of the entire state vector $x(k)$'s instead of the

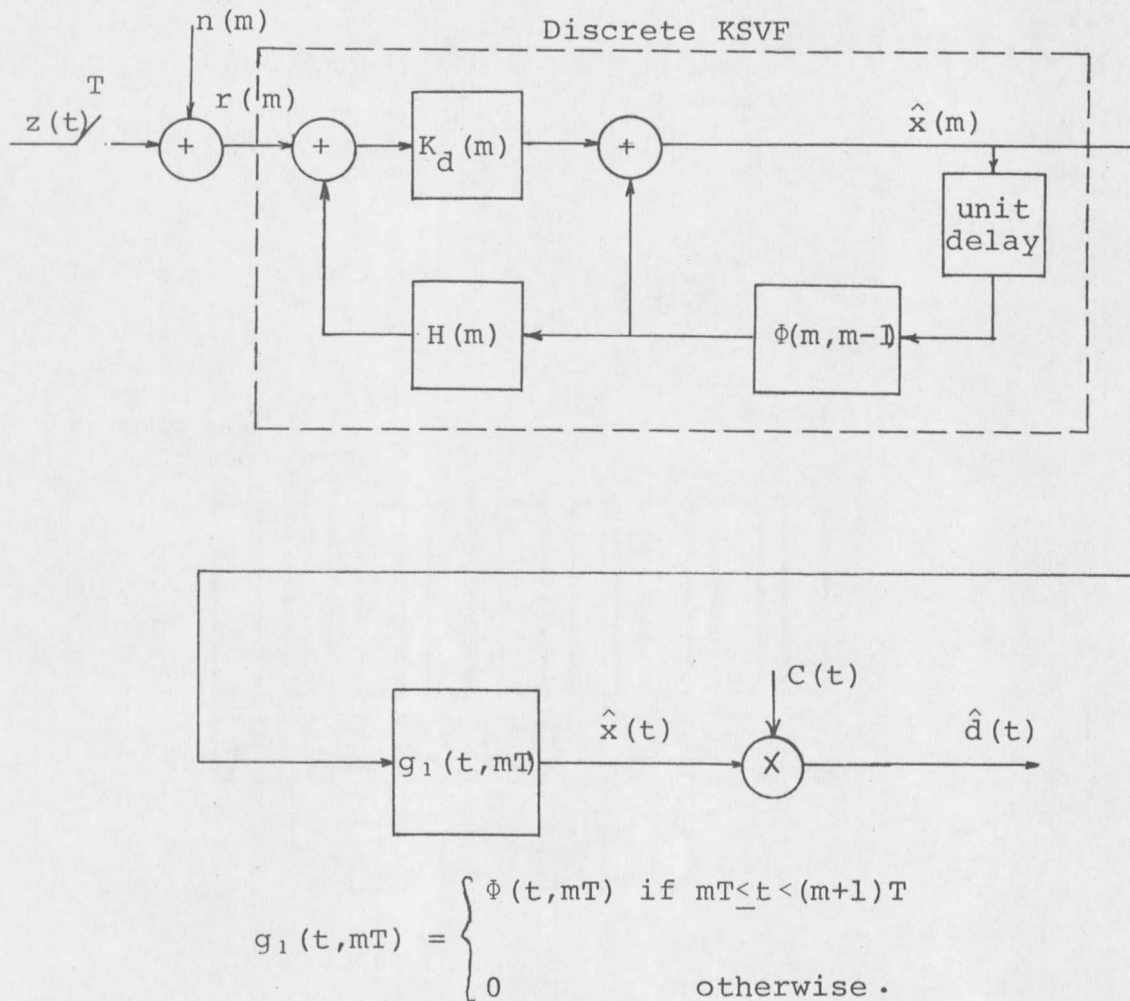


Figure 2.3 A two-filter structure for the optimum zero-lag postfilter.

desired signal $d(k)$ which is usually a linear combination of the state variables. The second filters in the two systems differ from each other accordingly.

An alternative derivation based on the innovations approach, which will also lead to the solution of the optimal filtering, is presented next. This alternative derivation will not only yield a different filter structure, but will also enable one to solve the smoothing problem.

2.2.2 The innovations approach

There are three basic concepts behind this development:

1. The innovations concept which has already been discussed in Section 1.4.3.
2. It can be seen from (2.28) that the zero-lag optimal estimate of the state vector, $\hat{x}(t)$ (and thus, $\hat{d}(t)$), is a piece-wise continuous function in time.
3. Recall that if a function $f(t)$ has a jump at certain points, then the derivative of $f(t)$ will contain impulse functions at those points. An example is shown in Fig. 2.4.

The discrete innovation process $u_d(k)$ of $r(k)$ is given by Eq. (B.11). $u_d(k)$ can either be generated by a

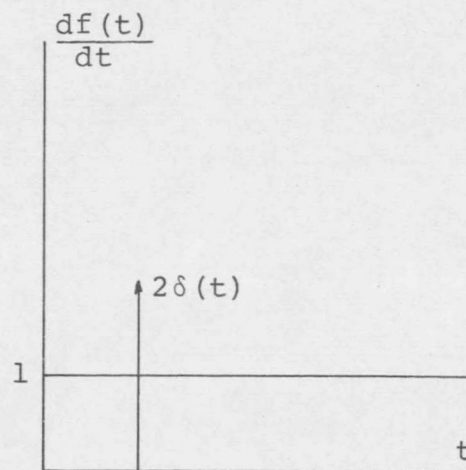
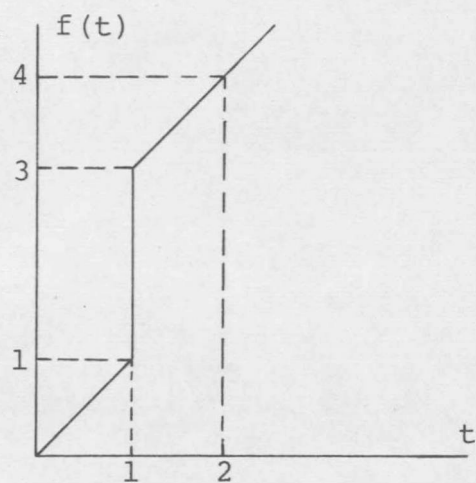


Figure 2.4 A piece-wise continuous function, and its derivative.

discrete filter, or by a feedback loop in the optimum system for filtering the desired signal, as will be explained later. The proof that $u_d(k)$ is a discrete white noise is given in Appendix B.

At this point, let us look at the general situation including the smoothing, the prediction, and the filtering problems. Let $\hat{x}(t/\ell_f) = \hat{x}(t/\ell_i, \ell_f, r(k))$, since $\hat{x}(t/\ell_i, \ell_f, r(k)) = \hat{x}(t/\ell_i, \ell_f, u_d(k))$ from the discussion made in Section 1.4.3.

$$\hat{x}(t/\ell_f) = \sum_{k=\ell_i}^{\ell_f} g(t, kT) u_d(k). \quad (2.31)$$

From the orthogonality principle, for $\ell_i \leq k \leq \ell_f$,

$$E\{[x(t) - \hat{x}(t/\ell_f)] u_d^T(k)\} = 0. \quad (2.32)$$

Substituting (2.31) into (2.32), and using (B.29) to obtain, for $\ell_i \leq k \leq \ell_f$

$$g(t, kT) = E[x(t) u_d^T(kT)] V_{u_d}^{-1}(k), \quad (2.33)$$

where V_{u_d} is given by (B.28). Substituting (2.33) into (2.31) yields

$$\hat{x}(t/\ell_f) = \sum_{k=\ell_i}^{\ell_f} E[x(t) u_d^T(kT)] V_{u_d}^{-1}(k) u_d(k). \quad (2.34)$$

For a zero-lag estimation, $\ell_f = m = \langle t/T \rangle$,
 $\hat{x}(t/\ell_f) = \hat{x}(t/m) = \hat{x}(t)$, and (2.34) becomes

$$\hat{x}(t) = \sum_{k=\ell_i}^m E[x(t) u_d'(kT)] V_{u_d}^{-1}(k) u_d(k). \quad (2.35)$$

Substituting $\tau = kT$ into (C.1) yields, for $kT < t$

$$x(t) = \Phi(t, kT) x(kT) + \int_{kT}^t \Phi(t, v) \Gamma(v) w(v) dv. \quad (2.36)$$

Since $u_d(k)$ depends on $x(j)$ and $n(j)$ for $j \leq k$ only,

$$E[u_d(k) w(v)] = 0 \quad \text{for } v > kT. \quad (2.37)$$

Thus, for $kT \leq t$,

$$\begin{aligned} E[x(t) u_d'(kT)] &= \Phi(t, kT) E[x(k) u_d'(k)] + \int_{kT}^t \Phi(t, v) \Gamma(v) \\ &\quad E[w(v) u_d'(kT)] dv \\ &= \Phi(t, kT) E[x(k) u_d'(k)]. \end{aligned} \quad (2.38)$$

From (B.18), it follows that,

$$\begin{aligned} E[x(k) u_d'(k)] &= E\{x(k) [H(k) \hat{x}(k/k-1) + n(k)]'\} \\ &= E[x(k) \hat{x}(k/k-1)] H'(k) \\ &\quad + E[x(k) n'(k)]. \end{aligned} \quad (2.39)$$

Since n is assumed to be uncorrelated with x ,

$$\begin{aligned} E[x(k) u_d'(k)] &= E[x(k) \tilde{x}(k/k-1)] H'(k) \\ &= E\{[\hat{x}(k/k-1) + \tilde{x}(k/k-1)] \tilde{x}(k/k-1)\} \cdot H'(k) \\ &= V_x^{\sim}(k/k-1) \cdot H'(k), \end{aligned} \quad (2.40)$$

where the second equality is due to the relation $x(k) = \hat{x}(k/k-1) + \tilde{x}(k/k-1)$, the third equality is due to the orthogonality principle; and $V_x^{\sim}(k/k-1) = E[\tilde{x}^2(k/k-1)]$. Substituting (2.40) into (2.38), then substituting the result into (2.33) to obtain, for $kT \leq t$,

$$g(t, kT) = \Phi(t, kT) V_x^{\sim}(k/k-1) H'(k) V_{u_d}^{-1}(k). \quad (2.41)$$

From Equations (2.35), (2.41), and (B.13), it follows that

$$\hat{x}(t) = \sum_{k=\ell_i}^m \Phi(t, kT) K_d(k) u_d(k). \quad (2.42)$$

where $K_d(k)$ is given by (B.13), $u_d(k)$ is given by (B.11), and $m = \langle t/T \rangle$. Differentiating both sides of (2.42) with respect to t for $mT < t < (m+1)T$ to obtain

$$\begin{aligned}
\frac{d\hat{x}(t)}{dt} &= \sum_{k=\ell_i}^m \frac{\partial \Phi(t, kT)}{2t} K_d(k) u_d(k) \\
&= \sum_{k=\ell_i}^m A(t) \Phi(t, kT) K_d(k) u_d(k) \\
&= A(t) \hat{x}(t),
\end{aligned} \tag{2.43}$$

where the second equality comes from Eq. (C.2), and the third equality comes from (2.42). At the sampling instants, let

$$mT^- = \lim_{\Delta \rightarrow 0} (mT - \Delta),$$

and

$$mT^+ = \lim_{\Delta \rightarrow 0} (mT + \Delta).$$

Then from (2.42),

$$\hat{x}(mT^-) = \sum_{k=\ell_i}^{m-1} \Phi(mT, kT) K_d(k) u_d(k), \tag{2.44}$$

and

$$\hat{x}(mT^+) = \sum_{k=\ell_i}^m \Phi(mT, kT) K_d(k) u_d(k). \tag{2.45}$$

Thus,

$$\hat{x}(mT^+) - \hat{x}(mT^-) = K_d(m) u_d(m) \quad (2.46)$$

From (2.43) and (2.46), it follows that for all t ,

$$\dot{\hat{x}}(t) = A(t) \hat{x}(t) + \sum_{\ell=\ell_i}^{\infty} K_d(\ell) u_d(\ell) \delta(t-\ell T), \quad (2.47)$$

where $K_d(\ell)$ is the discrete KSVF gain given by (B.13), $u_d(\ell)$ is the discrete innovation process of $r(\ell)$ given by (B.11), and $\delta(t)$ is the unit impulse function. Two different block diagrams which realize (2.47) are shown in Figures 2.5_a and 2.5_b. In Fig. 2.5_a, $u_d(\ell)$ is generated by a separated discrete filter. In Fig. 2.5_b, $u_d(\ell)$ is generated by a feedback loop. Since $\hat{x}(t)$ contains discontinuities at the sampling instants, one has to keep in mind that $y(\ell T)$ in Fig. 2.5_b is $\hat{x}(\ell T/(\ell-1)T)$ instead of $\hat{x}(\ell T)$. In other words, S_1 and S_2 in Fig. 2.5_b must be in synchronism.

The structure of the optimal postfilter shown in Fig. 2.5_b is similar to the postfilter structure obtained by Petersen (42). The difference is that: in Petersen's work, the prefilter is constrained to be the combination of a continuous KSVF followed by a constant coder, thus,

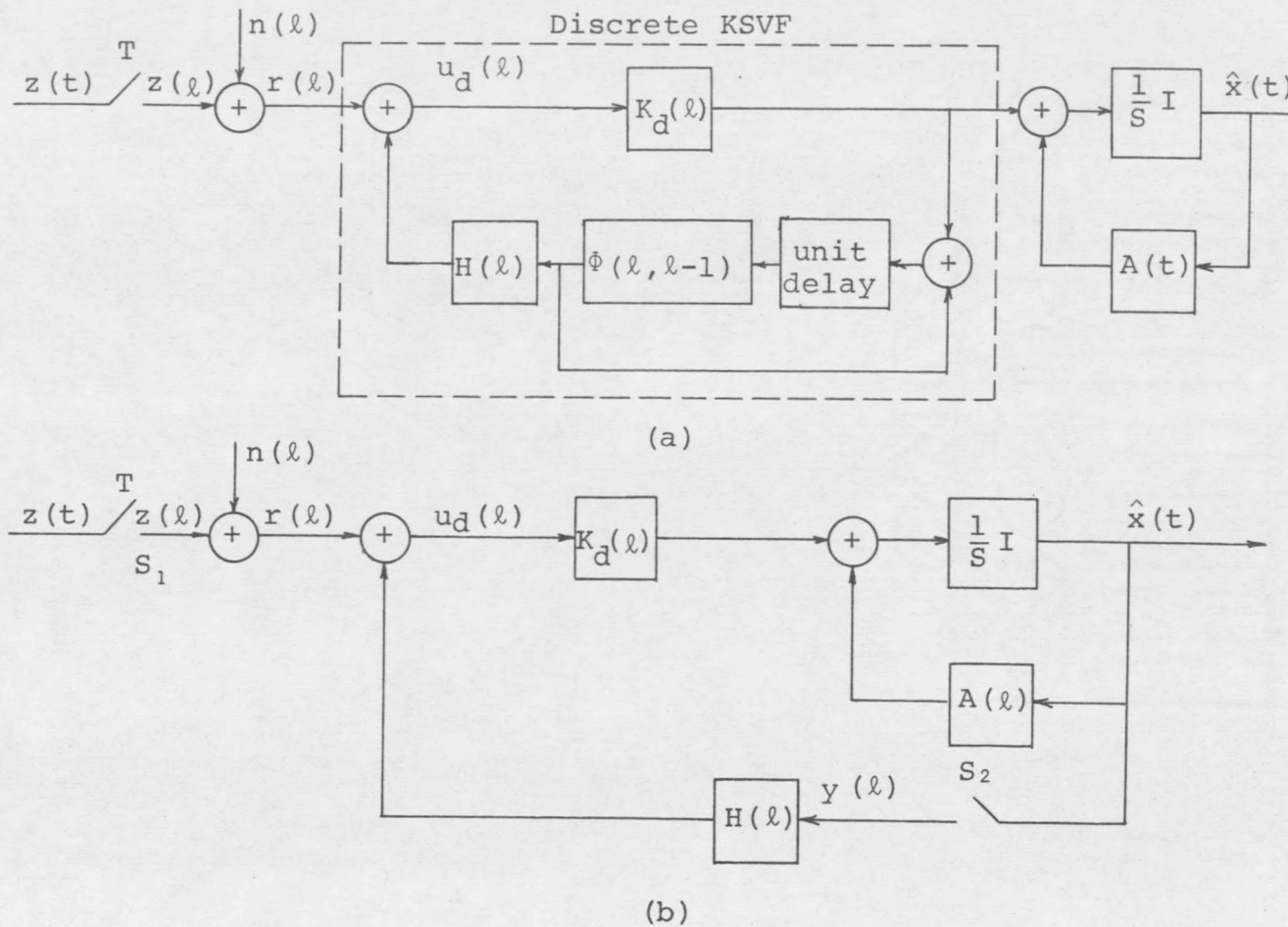


Figure 2.5 The optimum zero-lag postfilter: (a) The innovation process is generated by a separate discrete KSVF. (b) The innovation process is generated by a feedback loop.

the observation process at the receiving end is restricted to be the sampled version of the class of random processes which can be represented as linear combinations of the output of the continuous KSVF. Whereas in this work, the only restriction about the observation process at the receiving end is that they must be characterizable by state equations such as Equations (2.1) through (2.3) and (C.7). Thus, the prefilter can either be absent, or can be a filter whose outputs can be characterized by state equations.

The structure of the optimal postfilter shown in Fig. 2.5_b bears the form of a KSVF (between a continuous and a discrete KSVF). The major difference is that the boundary conditions of the integrator shown in Fig. 2.5_b are reset at the sampling instants. A similar result has been found by Moore and Hetrakul (40).

Once the optimal filtering is solved, the optimal prediction can be obtained by Eq. (2.26).

The error covariance of the state vector, $E[\hat{\mathbf{x}}^2(t)]$; and of the desired signal, $E[\hat{\mathbf{d}}^2(t)]$, are obtained in Appendix D.

Smoothing problems are considered next.

2.3 Optimal Smoothing

Only delays which are integer multiples of T are considered here. We will solve this problem in three steps. First, we will look at the problem in which the delay is one sample. Then, we will treat delays of more than one sample. Finally, we will put the solution to the above two cases in one form with the former one considered as an initial condition. Let

$$\begin{aligned}\hat{\xi}(t, n_d) &= \hat{\xi}(t/\ell_i, m+n_d, r(k)) \\ &= \hat{\xi}(t/\ell_i, m+n_d, u_d(k))\end{aligned}$$

$\triangleq$ linear-least-squares estimate of $\xi(t)$
given the $r(k)$'s for $\ell_i \leq k \leq (m+n_d)$, where
 $m = \langle t/T \rangle$.

Then from (2.31),

$$\hat{x}(t, n_d) = \hat{x}(t) + \sum_{\ell=m+1}^{m+n_d} g(t, \ell T) u_d(\ell), \quad (2.48)$$

where $g(t, \ell T)$ is given by Eq. (2.33), and

$$u_d(m+k) = r(m+k) - \hat{z}[(m+k)/(m+k-1)] \quad (2.49)$$

from (B.18), where $\hat{z}[(m+k)/(m+k-1)]$ is the best estimate of $z(m+k)$ given the $r(i)$'s for $i \leq m+k-1$. If $k=1$ (delay of

one sample), $(m+k-1)T = mT \leq t$. If $k \geq 2$, $(m+k-1)T \geq (m+1)T > t$. Thus, it is necessary to consider separately the problem of a delay of one sample.

2.3.1 Delay of one sample

Substituting $n_d=1$ into (2.48) yields

$$\hat{x}(t,1) = \hat{x}(t) + g(t, (m+1)T) u_d(m+1), \quad (2.50)$$

where from (2.33),

$$g(t, (m+1)T) = E[x(t) u_d'(m+1)] V_{u_d}^{-1}(m+1), \quad (2.51)$$

and from (B.11),

$$\begin{aligned} u_d(m+1) &= r(m+1) - \hat{z}(m+1/m) \\ &= H(m+1) x(m+1) + n(m+1) - H(m+1) \hat{x}(m+1/m). \end{aligned} \quad (2.52)$$

From (2.26),

$$\hat{x}(m+1/m) = \Phi(m+1, m) \hat{x}(m) \quad (2.53)$$

and

$$\hat{x}(t) = \Phi(t, mT) \hat{x}(m) \quad \text{for } mT \leq t < (m+1)T \quad (2.54)$$

From Equations (2.53), (2.54), and the properties of a transition matrix,

$$\hat{x}(m+1/m) = \Phi(m+1, t) \hat{x}(t) \quad \text{for } mT \leq t < (m+1)T. \quad (2.55)$$

Substituting $t=(m+1)T$, and $\tau=t$ into (C.1) to obtain

$$x(m+1) = \Phi((m+1)T, t) x(t) + \int_t^{(m+1)T} \Phi((m+1)T, v) \Gamma(v) w(v) dv. \quad (2.56)$$

Substituting (2.55) and (2.56) into (2.52) to obtain

$$u_d(m+1) = H(m+1)\Phi((m+1), t)[x(t) - \hat{x}(t)] + H(m+1) \times \\ \int_t^{(m+1)T} \Phi((m+1)T, v) \Gamma(v) w(v) dv + n(m+1) \quad (2.57)$$

Thus,

$$E[x(t) u_d'(m+1)] = E[x(t) \hat{x}'(t)] \Phi'((m+1)T, t) H'(m+1) + \\ \int_t^{(m+1)T} E[x(t) w'(v)] \Gamma'(v) \Phi'((m+1)T, v) \\ H'(m+1) dv + E[x(t) n'(m+1)]. \quad (2.58)$$

From (2.6) and (C.1),

$$E[x(t) w'(v)] = 0 \quad \text{for } v < t, \quad (2.59)$$

and

$$E[x(t) w'(t)] = (1/2)\Gamma(t) V_w(t) \quad (2.60)$$

Thus, the second term on the right-hand of (2.58) is zero from (2.59) and (2.60). The third term is also zero since the noise is assumed to be uncorrelated with the signal.

Hence, (2.58) becomes

$$\begin{aligned} E[x(t) u_d'(m+1)] &= E[x(t) \hat{x}'(t)] \Phi'((m+1)T, t) H'(m+1) \\ &= V_x^v(t) \Phi'((m+1)T, t) H'(m+1) \end{aligned} \quad (2.61)$$

from the orthogonality principle. Substituting (2.61) into (2.51) yields

$$g(t, (m+1)T) = V_x^v(t) \Phi'((m+1)T, t) H'(m+1) V_{u_d}^{-1}(m+1) \quad (2.62)$$

Substituting (2.62) into (2.50) to obtain

$$\hat{x}(t, 1) = \hat{x}(t) + V_x^v(t) \Phi'((m+1)T, t) H'(m+1) V_{u_d}^{-1}(m+1) u_d(m+1), \quad (2.63)$$

For the moment, we will leave (2.63) as it is, and divert our attention to the case of a delay of more than one sample.

2.3.2 Delays of more than one sample

For $j \geq m+2$, we write (B.11) in another form ($m = \lfloor t/T \rfloor$)

$$\begin{aligned} u_d(j) &= r(j) - \hat{z}(j/j-1) \\ &= H(j) x(j) + n(j) - H(j) \hat{x}(j/j-1). \end{aligned} \quad (2.64)$$

Substituting (C.7) and (2.26) into (2.64) yields

$$\begin{aligned}
u_d(j) &= H(j) [\Phi(j, j-1) x(j-1) + w_d(j-1)] - H(j) \Phi(j, j-1) \hat{x}(j-1) \\
&\quad + n(j) \\
&= H(j) \Phi(j, j-1) \tilde{x}(j-1) + H(j) w_d(j-1) + n(j). \quad (2.65)
\end{aligned}$$

Thus,

$$\begin{aligned}
E[x(t) u_d'(jT)] &= E[x(t) \tilde{x}'(j-1)] \Phi'(j, j-1) H'(j) \\
&\quad + E[x(t) w_d'(j-1)] H'(j) + E[x(t) n'(jT)]. \quad (2.66)
\end{aligned}$$

Since $t < (j-1)T$, $E[x(t) w_d'(j-1)] = 0$ from (C.13).

$E[x(t) n'(jT)]$ is also zero since the noise is assumed to be uncorrelated with the signal. Thus, (2.66) becomes, for $j \geq m+2$

$$E[x(t) u_d'(jT)] = E[x(t) \tilde{x}'((j-1)T)] \Phi'(j, j-1) H'(j) \quad (2.67)$$

From (2.47), it follows that

$$\begin{aligned}
\hat{x}(j-1) &= \Phi((j-1)T, t) \hat{x}(t) \\
&\quad + \int_t^{[(j-1)T]} \Phi((j-1)T, v) \sum_{\ell=\ell_i}^{\infty} K_d(\ell) u_d(\ell) \delta(v - \ell T) dv \\
&= \Phi((j-1)T, t) \hat{x}(t) + \sum_{\ell=m+1}^{j-1} \Phi((j-1)T, T) K_d(\ell) u_d(\ell). \quad (2.68)
\end{aligned}$$

Substituting $t=(j-1)T$ and $\tau=t$ into (C.1) to obtain

$$x(j-1) = \Phi((j-1)T, t) x(t) + \int_t^{(j-1)T} \Phi((j-1)T, v) \Gamma(v) w(v) dv. \quad (2.69)$$

Subtracting (2.68) from (2.69) yields, for $j \geq m+2$

$$\begin{aligned} \tilde{x}(j-1) &= \Phi((j-1)T, t) \tilde{x}(t) + \int_t^{(j-1)T} \Phi((j-1)T, v) \Gamma(v) w(v) dv \\ &\quad - \sum_{\ell=m+1}^{j-1} \Phi((j-1)T, T) K_d(\ell) u_d(\ell). \end{aligned} \quad (2.70)$$

Through the use of Equations (2.59) and (2.60), and the orthogonality principle,

$$\begin{aligned} E[x(t) \tilde{x}'(j-1)] &= V_x(t) \Phi'((j-1)T, t) \\ &\quad - \sum_{\ell=m+1}^{j-1} E[x(t) u_d'(\ell)] K_d(\ell) \\ &\quad \times \Phi'((j-1)T, \ell T). \end{aligned} \quad (2.71)$$

for $j \geq m+2$. Assume first that $j=m+2$. Then $j-1=m+1$, and

$$E[x(t) \tilde{x}'(m+1)] = V_x(t) \Phi'((m+1)T, t) - E[x(t) u_d'(m+1)] K_d'(m+1). \quad (2.72)$$

Substituting (2.61) into (2.72) to obtain

$$E[x(t) \tilde{x}'(m+1)] = V_x(t) \Phi'((m+1)T, t) [I - H'(m+1) K_d'(m+1)]. \quad (2.73)$$

Substituting (2.73) into (2.67) yields

$$E[x(t) u_d'(m+2)] = V_x^v(t) \Phi'((m+1)T, t) [I - H'(m+1) K_d'(m+1)] \\ \times \Phi'(m+2, m+1) H'(m+2). \quad (2.74)$$

Substituting (2.74) into (2.33), we obtain

$$g(t, (m+2)T) = V_x^v(t) \Phi'((m+1)T, t) [I - H'(m+1) K_d'(m+1)] \Phi'(m+1, m+2) \\ \times H'(m+2) V_{u_d}^{-1}(m+2). \quad (2.75)$$

For $j > m+2$, we have

$$\Phi'((j-1)T, t) = \Phi'((j-2)T, t) \Phi'(j-1, j-2) \quad (2.76)$$

and

$$\Phi'((j-1)T, \ell T) = \Phi'(j-2, \ell) \Phi'(j-1, j-2), \quad (2.77)$$

Substituting (2.76) and (2.77) into (2.71) yields, for $j > m+2$,

$$E[x(t) \tilde{x}'((j-1)T)] = V_x^v(t) \Phi'((j-2)T, t) \Phi'(j-1, j-2) \\ - \left[\sum_{\ell=m+1}^{j-2} E[x(t) u_d'(\ell)] K_d'(\ell) \Phi'(j-2, \ell) \Phi'(j-1, j-2) \right. \\ \left. + E[x(t) u_d'(j-1)] K_d'(j-1) \Phi'(j-1, j-1) \right]$$

$$\begin{aligned}
&= [V_{\tilde{x}}(t)\Phi'((j-2)T, t) - \sum_{\ell=m+1}^{j-2} E[x(t) u_d'(\ell)] K_d'(\ell)\Phi'(j-2, \ell)] \\
&\quad \times \Phi'(j-1, j-2) - E[x(t) u_d'(j-1)] K_d'(j-1). \quad (2.78)
\end{aligned}$$

Through the use of Equations (2.71) and (2.67), we have

$$\begin{aligned}
E[x(t) \tilde{x}'((j-1)T)] &= E[x(t) \tilde{x}'((j-2)T)] \Phi'(j-1, j-2) \\
&\quad - E[x(t) \tilde{x}'((j-2)T)] \Phi'(j-1, j-2) \\
&\quad \times H'(j-1) K_d'(j-1); \quad (2.79)
\end{aligned}$$

i.e.,

$$\begin{aligned}
E[x(t) \tilde{x}'((j-1)T)] &= \\
&E[x(t) \tilde{x}'((j-2)T)] \Phi'(j-1, j-2) [I - H'(j-1) K_d'(j-1)]. \quad (2.80)
\end{aligned}$$

for $j \geq m+2$. Substituting (2.80) into (2.67) and then substituting the result into (2.33), we obtain

$$\begin{aligned}
g(t, jT) &= E[x(t) \tilde{x}'((j-2)T)] \Phi'(j-1, j-2) [I - H'(j-1) K_d'(j-1)] \\
&\quad \Phi'(j, j-1) H'(j) v_{u_d}^{-1}(j) \quad (2.81)
\end{aligned}$$

for $j > m+2$.

2.3.3 General solution

The general solution can be expressed in two different forms: the first one represents the solution of the

fixed-point smoothing (51). The second one represents the solution of the fixed-lag smoothing.

1. We first obtain the solution of the fixed-point smoothing. Let

$$B(t, j) = E[x(t) \hat{x}'(jT)] \quad \text{for } j \geq m+1. \quad (2.82)$$

Then

$$B(t, j) = B(t, j-1) \cdot \Phi'(j, j-1) [I - H'(j) K_d'(j)] \quad \text{for } j \geq m+1 \quad (2.83)$$

with

$$B(t, m) = V_x(t) \Phi'((m+1)T, t) \Phi'(m, m+1). \quad (2.84)$$

Substituting (2.82) into (2.81) to obtain

$$g(t, jT) = B(t, j-1) \Phi'(j, j-1) H'(j) V_{u_d}^{-1}(j) \quad \text{for } j \geq m+1. \quad (2.85)$$

Finally, for $n_d \geq 1$,

$$\begin{aligned} \hat{x}(t, n_d) &= \hat{x}(t, n_d - 1) + g(t, m+n_d) u_d(m+n_d) \\ &= \hat{x}(t, n_d - 1) + B(t, m+n_d - 1) \Phi'(m+n_d, m+n_d - 1) H'(m+n_d) \\ &\quad \times V_{u_d}^{-1}(m+n_d) u_d(m+n_d), \end{aligned} \quad (2.86)$$

where $\hat{x}(t, 0) = \hat{x}(t)$, $B(t, j)$ is specified by (2.83) and (2.84), and $m = \lfloor t/T \rfloor$. At the sampling instants, the result obtained here checks with the known result for the optimal

discrete fixed-point smoothing (51; pp.374, Eqs. (8.3-106) and (8.3-107)).

2. To put the general solution in another form which corresponds to fixed-lag smoothing, define

$$D_m(1) = 1, \quad (2.87_a)$$

$$D_m(i) = [I - H'(m+i-1) K'_d(m+i-1)] \Phi'(m+i, m+i-1) \text{ for } i > 1, \quad (2.87_b)$$

$$\Lambda_m(1) = 1, \quad (2.88_a)$$

and

$$\Lambda_m(i) = \prod_{j=1}^i D_m(j) \text{ with } c^m(j) \text{ given by (2.87)} \quad (2.88_b)$$

Using the above definitions, Equations (2.62), (2.75), and (2.81) become

$$g(t, (m+1)T) = V_x^\nu(t) \Phi'((m+1)T, t) \Lambda_m(1) H'(m+1) V_{u_d}^{-1}(m+1), \quad (2.89_a)$$

$$\begin{aligned} g(t, (m+2)T) &= V_x^\nu(t) \Phi'((m+1)T, t) D_m(2) H'(m+2) V_{u_d}^{-1}(m+2) \\ &= V_x^\nu(t) \Phi'((m+1)T, t) \Lambda_m(2) H'(m+2) V_{u_d}^{-1}(m+2), \end{aligned} \quad (2.89_b)$$

and, for $k > 2$,

$$g(t, m+k) = E[x(t) \hat{x}'((j-2)T)] \Phi'(m+k-1, m+k-2) D_m(k) \\ H'(m+k) V_{u_d}^{-1}(m+k), \quad (2.89_c)$$

respectively. By repeatedly using (2.80), we obtain,
for $k \geq 1$

$$g(t, m+k) = V_x^{-1}(t) \Phi'((m+1)T, t) \Lambda_m(k) H'(m+k) V_{u_d}^{-1}(m+k). \quad (2.90)$$

Substituting (2.90) into (2.48) to obtain

$$\hat{x}(t, n_d) = \hat{x}(t) + V_x^{-1}(t) \Phi'((m+1)T, t) \sum_{k=1}^{n_d} \eta_m(k), \quad (2.91)$$

where

$$\eta_m(k) = \Lambda_m(k) H'(m+k) V_{u_d}^{-1}(m+k) u_d(m+k) \quad (2.92)$$

The block diagram which realize Eq. (2.91) is shown in Fig. 2.6, where $u_d(l)$ and $\hat{x}(t)$ can be generated by the system shown in Fig. 2.5. Eq. (2.91) corresponds to the solution of the fixed-lag smoothing. To obtain a solution in the form of a differential equation in the state variables, differentiating both sides of (2.91) with respect to t for $mT < t < (m+1)T$ to obtain

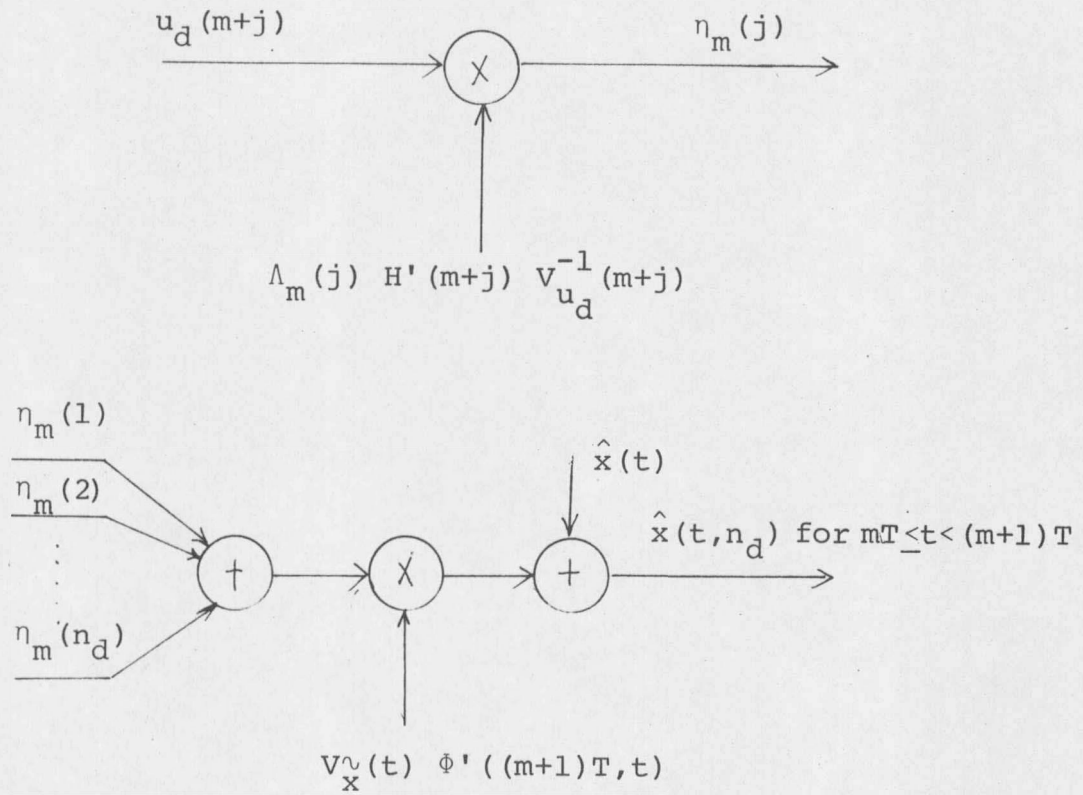


Figure 2.6 The optimal fixed-lag postfilter specified by Eq. (2.91).

$$\begin{aligned} \dot{\hat{x}}(t, n_d) = & \dot{\hat{x}}(t) + [\dot{V}_x(t) \Phi'((m+1)T, t) + V_x(t) \dot{\Phi}'((m+1)T, t)] \\ & \times \sum_{k=1}^{n_d} \eta_m(k). \end{aligned} \quad (2.93)$$

Since

$$\frac{\partial \Phi(t, (m+1)T)}{\partial t} = A(t) \Phi(t, (m+1)T) \quad (2.94)$$

from (C.2), and

$$\Phi'((m+1)T, t) \Phi'(t, (m+1)T) = I \quad (2.95)$$

from (C.3) and (C.4), we have

$$\dot{\Phi}'((m+1)T, t) = -A'(t) \Phi'((m+1)T, t). \quad (2.96)$$

Substituting (D.26) and (2.96) into (2.93), and after some manipulation, we obtain for $mT < t < (m+1)T$,

$$\begin{aligned} \dot{\hat{x}}(t, n_d) = & A(t) \hat{x}(t, n_d) + \Gamma(t) V_w(t) \Gamma'(t) \Phi'((m+1)T, t) \sum_{k=1}^{n_d} \eta_m(k). \end{aligned} \quad (2.97)$$

At the sampling instants,

$$\hat{x}(mT^-, n_d) = \hat{x}(mT^-) + \sum_{k=m}^{m+n_d-1} g(m, k) u_d(k), \quad (2.98)$$

and

$$\hat{x}(mT^+, n_d) = \hat{x}(mT^+) + \sum_{k=m+1}^{m+n_d} g(m, k) u_d(k). \quad (2.99)$$

In this problem, since $g(t, jT)$ has different forms for $t < jT$ and $t > jT$, it is necessary to check the continuity of $g(t, jT)$ at $t = jT$ before proceeding. From (2.41), it follows that for $t \rightarrow jT^+$

$$\lim_{t \rightarrow jT^+} g(t, jT) = V_x^u(j/j-1) H'(j) V_{u_d}^{-1}(j). \quad (2.100)$$

For $t \rightarrow jT^-$, $m = \lfloor t/T \rfloor = j-1$. Thus, $\hat{x}(jT^-) = \hat{x}(jT^-/mT) = \hat{x}(j/j-1)$, and $V_x^u(jT^-) = V_x^u(j/j-1)$. Hence, from (2.62),

$$\begin{aligned} \lim_{t \rightarrow jT^-} g(t, jT) &= \lim_{t \rightarrow jT^-} g(t, (m+1)T) \\ &= \lim_{t \rightarrow jT^-} [V_x^u(t) \Phi'((m+1)T, t) H'(m+1) V_{u_d}^{-1}(m+1)] \\ &= V_x^u(j/j-1) H'(j) V_{u_d}^{-1}(j). \end{aligned} \quad (2.101)$$

From (2.100) and (2.101), it follows that

$$\lim_{t \rightarrow jT^+} g(t, jT) = \lim_{t \rightarrow jT^-} g(t, jT). \quad (2.102)$$

From (2.46), (2.90), (2.98), (2.99), and (2.102), it follows that

$$\begin{aligned}
\hat{x}(mT^+, n_d) - \hat{x}(mT^-, n_d) &= V_x(m) \Phi'(m+1, m) \Lambda_m(n_d) H'(m+n_d) \\
&\quad V_{u_d}^{-1}(m+n_d) u_d(m+n_d) \\
&= V_x(m) \Phi'(m+1, m) \eta_m(n_d), \quad (2.103)
\end{aligned}$$

where the second equality follows from the substitution of (2.92). Combining (2.97) and (2.103) to obtain for all $t > \ell_i T$ (where $\ell_i T$ is the initial time the estimation begins).

$$\begin{aligned}
\dot{\hat{x}}(t, n_d) &= A(t) \hat{x}(t, n_d) + \Gamma(t) V_w(t) \Gamma'(t) \Phi'((m+1)T, t) \sum_{k=1}^{n_d} \eta_m(k) \\
&\quad + \sum_{\ell=\ell_i}^{\infty} V_x(\ell) \Phi'(\ell+1, \ell) \eta_\ell(n_d) \delta(t-\ell T). \quad (2.104)
\end{aligned}$$

where $m = \langle t/T \rangle$. In order to get rid of m , let

$$\lambda(t, n_d) = \Phi'((m+1)T, t) \sum_{k=1}^{n_d} \eta_m(k) \quad (2.105)$$

Then for $mT < t < (m+1)T$,

$$\dot{\lambda}(t, n_d) = -A'(t) \lambda(t, n_d) \quad (2.106)$$

from (2.96). At the sampling instants,

$$\lambda(mT^+, n_d) = \Phi'(m+1, m) \sum_{k=1}^{n_d} \eta_m(k), \quad (2.107)$$

and

$$\lambda(mT^-, n_d) = \Phi'(m, m) \sum_{k=1}^{n_d} \eta_{m-1}(k) \quad (2.108)$$

Thus,

$$\lambda(mT^+, n_d) - \lambda(mT^-, n_d) = \Phi'(m+1, m) \sum_{k=1}^{n_d} \eta_m(k) - \sum_{k=1}^{n_d} \eta_{m-1}(k). \quad (2.109)$$

Finally, through the use of Eqs. (2.104), (2.106), and (2.109), we have

$$\begin{aligned} \dot{\hat{x}}(t, n_d) = & A(t) \hat{x}(t, n_d) + \Gamma(t) V_w(t) \Gamma'(t) \lambda(t, n_d) \\ & + \sum_{\ell=\ell_i}^{\infty} V_X(\ell) \Phi'(\ell+1, \ell) \eta_{\ell}(n_d) \delta(t-\ell T), \end{aligned} \quad (2.110)$$

where

$$\begin{aligned} \dot{\lambda}(t, n_d) = & -A'(t) \lambda(t, n_d) \\ & + \sum_{\ell=\ell_i}^{\infty} \Phi(\ell+1, \ell) \left[\sum_{k=1}^{n_d} \eta_{\ell}(k) - \sum_{k=1}^{n_d} \eta_{\ell-1}(k) \right] \delta(t-\ell T), \end{aligned} \quad (2.111)$$

and $\eta_{\ell}(k)$ is given by (2.92). The block diagram of the system which implements (2.110) and (2.111) is shown in Fig. 2.7, where the innovation process is generated by a discrete filter. The advantages of the system shown in Fig. 2.7 over the system shown in Fig. 2.6 are the following: 1. It is no longer necessary to calculate $\hat{x}(t)$ and $V_x^v(t)$ in Fig. 2.7. 2. There is no need to find out during which sampling instants t lies; i.e., there is no need to find out $m = \langle t/T \rangle$. 3. In the case in which the inputs are stationary, there will be no need to calculate such parameters as $\Phi((m+1)T, t)$ which varies continuously with time (notice that in this case, Γ , V_w , A , etc., become constants).

The derivations of the error covariance of the smoothing problems are presented in Appendix D.

2.3.4 Delay efficiency

From (D.11), it follows that

$$\begin{aligned}
 V_d^v(t, n_d) &= V_d^v(t) - \sum_{k=1}^{n_d} g_d^2(t, (m+k)T) \\
 &= V_d^v(t) - g_d^2(t, (m+1)T) - \dots - g_d^2(t, (m+i)T) \\
 &\quad - g_d^2(t, (m+n_d)T)
 \end{aligned} \tag{2.112}$$

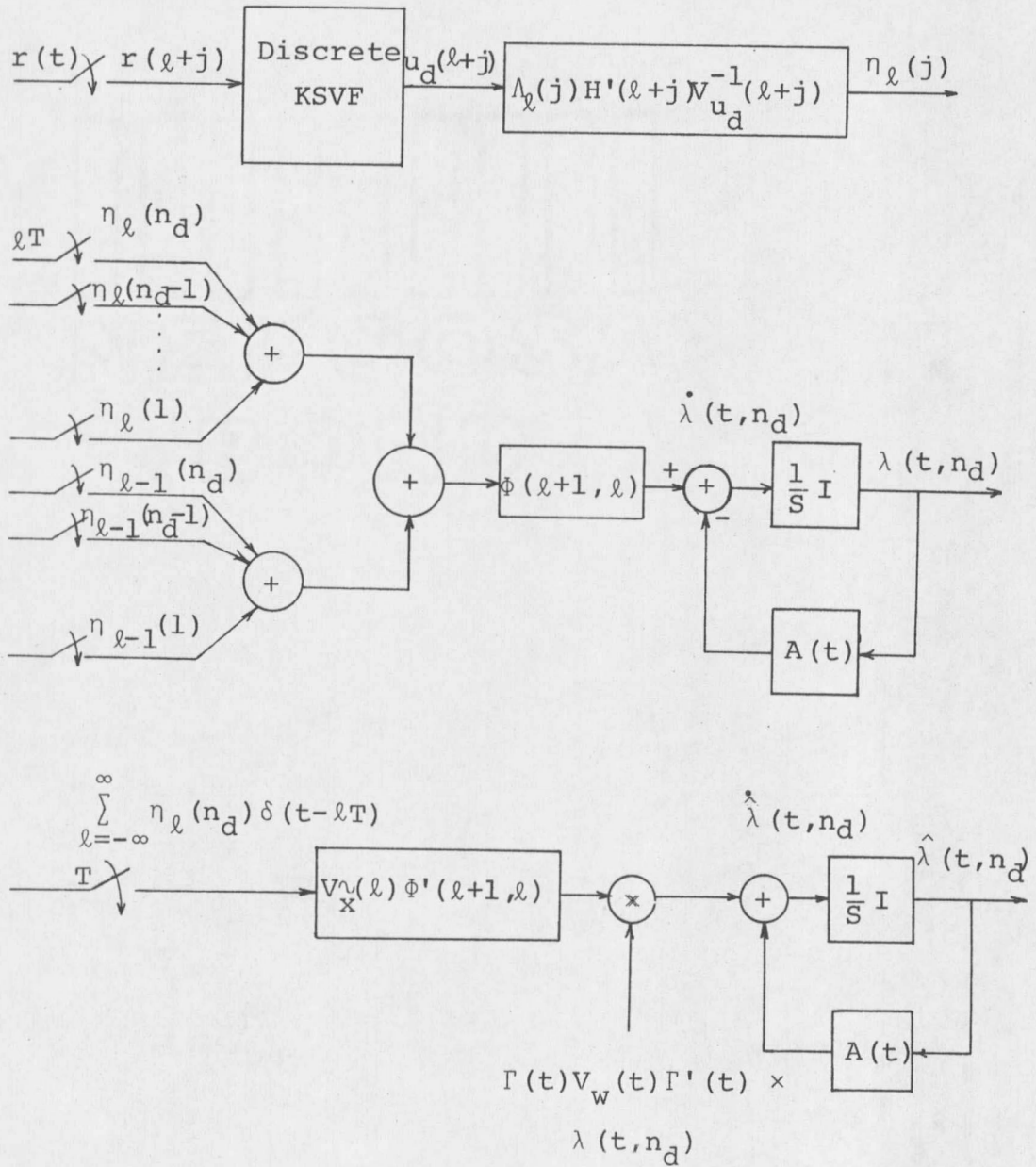


Figure 2.7 The optimal fixed-lag postfilter specified by (2.110) and (2.111).

where $n_d T$ denotes the total amount of delay, and $g_d^2(t, (m+k)T)$ for $k > 0$ is given by (D.12). The $V_d^v(t)$ on the right hand side of (2.112) represents the zero-lag minimum-mean-square error of estimation. The $g_d^2(t, (m+1)T)$ represents the error reduction due to the first sample of delay, and $g_d^2(t, (m+i)T)$ represents the error reduction due to the i th sample of delay, etc. Notice that the error reduction due to the i th sample of delay for $i \geq 1$ is independent of n_d , the total amount of delay. Also notice that the error reduction due to the i th sample of delay, $g_d^2(t, (m+i)T)$, is uncorrelated with the error reduction due to other samples of delay, $g_d^2(t, (m+j)T)$ for $j \neq i$. Define

$$ED(t, i) = g_d^2(t, (m+i)T) \quad (2.113)$$

and

$$EN(t, i) = \frac{g_d^2(t, (m+i)T)}{V_d^v(t)} \quad (2.114)$$

for $i \geq 1$. For steady state conditions and stationary inputs, define

$$EDS(i) = (1/T) \int_0^t ED(t, i) dt \quad (2.115)$$

$$ENS(i) = (1/T) \int_0^t EN(t, i) dt \quad (2.116)$$

for $i \geq 1$. $ED(t,i)$ and $EDS(i)$ indicate the amount of the error reduction, and $EN(t,i)$ and $ENS(i)$ indicate the efficiency of the error reduction by allowing the delay extended to the i th sample. Numerical examples are worked out in the next section.

2.4 Examples

Example 2.1: A first-order Gauss-Markov process with a white additive noise. Related equations are given by (2.1) through (2.8) with $A(t) = -\sigma$, $\Gamma(t) = 1$, $V_w(t) = 2$, $V_x(0) = \sigma$, and $d(t) = z(t) = x(t)$. The discretized state equations are given by (C.6) and (C.7). The transition matrix of A is

$$\Phi(t) = \exp(-\sigma T) \quad (2.117)$$

Let

$$\alpha = \Phi(T) = \exp(-\sigma T) \quad (2.118)$$

and

$$q = \frac{V_n(1+\alpha^2) + \sigma(1-\alpha^2) - \sqrt{[V_n(1+\alpha^2) + \sigma(1-\alpha^2)]^2 - 4}}{2\alpha V_n} \quad (2.119)$$

where V_n is the variance of n , i.e., $V_n = E[n^2]$ which is assumed given. From (B.13) through (B.16), the steady-state discrete Kalman-Bucy filter gain is given by

$$K_d = \frac{\alpha - q}{\alpha}, \quad (2.120)$$

and the variance of the discrete innovation process at the steady-state is

$$V_{u_d} = V_n \alpha / q \quad (2.121)$$

Through the use of (2.87), (2.88), and (2.120), it follows that for $i \geq 2$,

$$\Lambda_m(i) = q^{i-1}, \quad (2.122)$$

where q is given by (2.119). From (D.37), it follows that

$$V_x^v(t) = \exp(-2(t-mT)) \cdot V_x^v(m) + \sigma [1 - \exp(-2(t-mT))], \quad (2.123)$$

where the steady-state $V_x^v(m)$ can again be found by using Equations (B.13) through (B.16),

$$V_x^v(m) = V_{xd}^v = V_n \cdot K_d = V_n \cdot (\alpha - q) / \alpha \quad (2.124)$$

Substituting (2.122) and (2.121) into (D.12) to obtain

$$g_d(t, (m+1)T) = V_x^v(t) \exp(-((m+1)T-t)) q^{i-1/2} / (V_n \alpha)^{1/2}. \quad (2.125)$$

where $V_x^v(t)$ is given by (2.123). To obtain the steady-state solution, define

$$ERS(n_d) = (1/T) \int_0^T V_z^v(t, n_d) dt, \quad (2.126)$$

where $V_z^v(t, n_d)$ is given by (D.11) with $d(t)=z(t)$ (i.e., $V_d^v(t, n_d)=V_z^v(t, n_d)$), $g_d(t, (m+k)T)$ is given by (2.125), and $V_d^v(t)=V_x^v(t)$ is given by (2.123). After some manipulation, we obtain,

$$ERS(n_d) = \sigma - ((1-\alpha^2)/2T) (\sigma - V_n \frac{\alpha-q}{\alpha}) + \sum_{i=1}^{n_d} EDS(i), \quad (2.127)$$

where $EDS(i)$ is given by

$$EDS(i) = (\alpha/2T V_n) q^{2i-1} [V_{xd}^v - \sigma)(1-\alpha^2) + \sigma^2(1/\alpha^2-1) + 2(V_{xd}^v - \sigma)] \quad (2.128)$$

with $V_{xd}^v = V_x^v(m)$, as given by (2.124). From (2.116), it follows that

$$ENS(1) = \frac{EDS(1)}{\sigma - ((1-\alpha^2)/2T) (\sigma - V_n \cdot (\alpha-q)/\alpha)} \quad (2.129)$$

and

$$ENS(i) = ENS(1) q^{2(i-1)} \quad \text{for } i \geq 2. \quad (2.130)$$

Plots of $ERS(n_d)$ versus n_d with V_n/σ and aT as parameters are shown in Figs. 2.8 and 2.10. aT is a constant in Fig. 2.8, and V_n/σ is a constant in Fig. 2.10. $\log(ENS(i))$ versus i are plotted in Figs. 2.9 and 2.11.

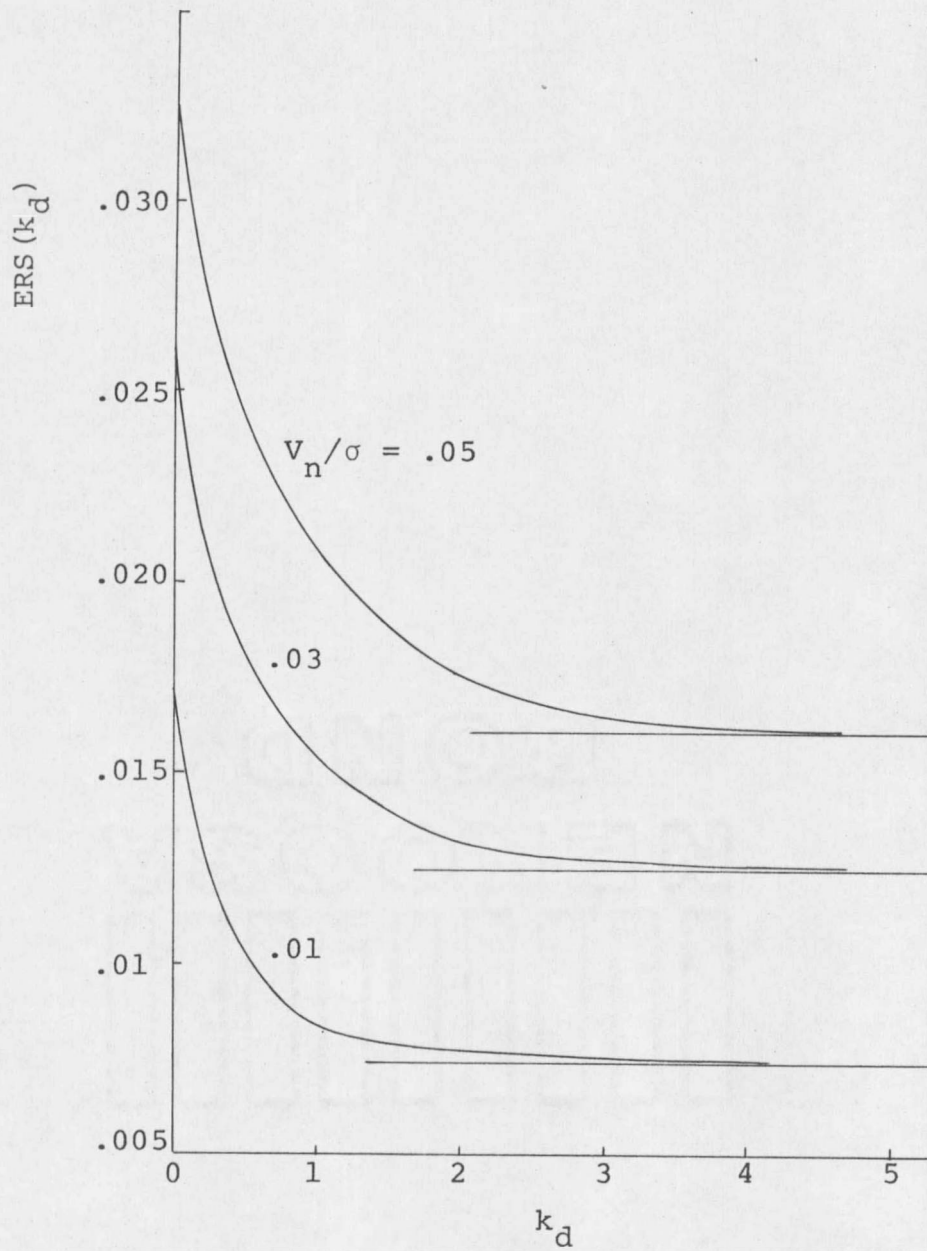


Figure 2.8 Normalized m.s.e. (mean square error) versus the amount of delay allowed for Example 2.1, $aT=0.01$.

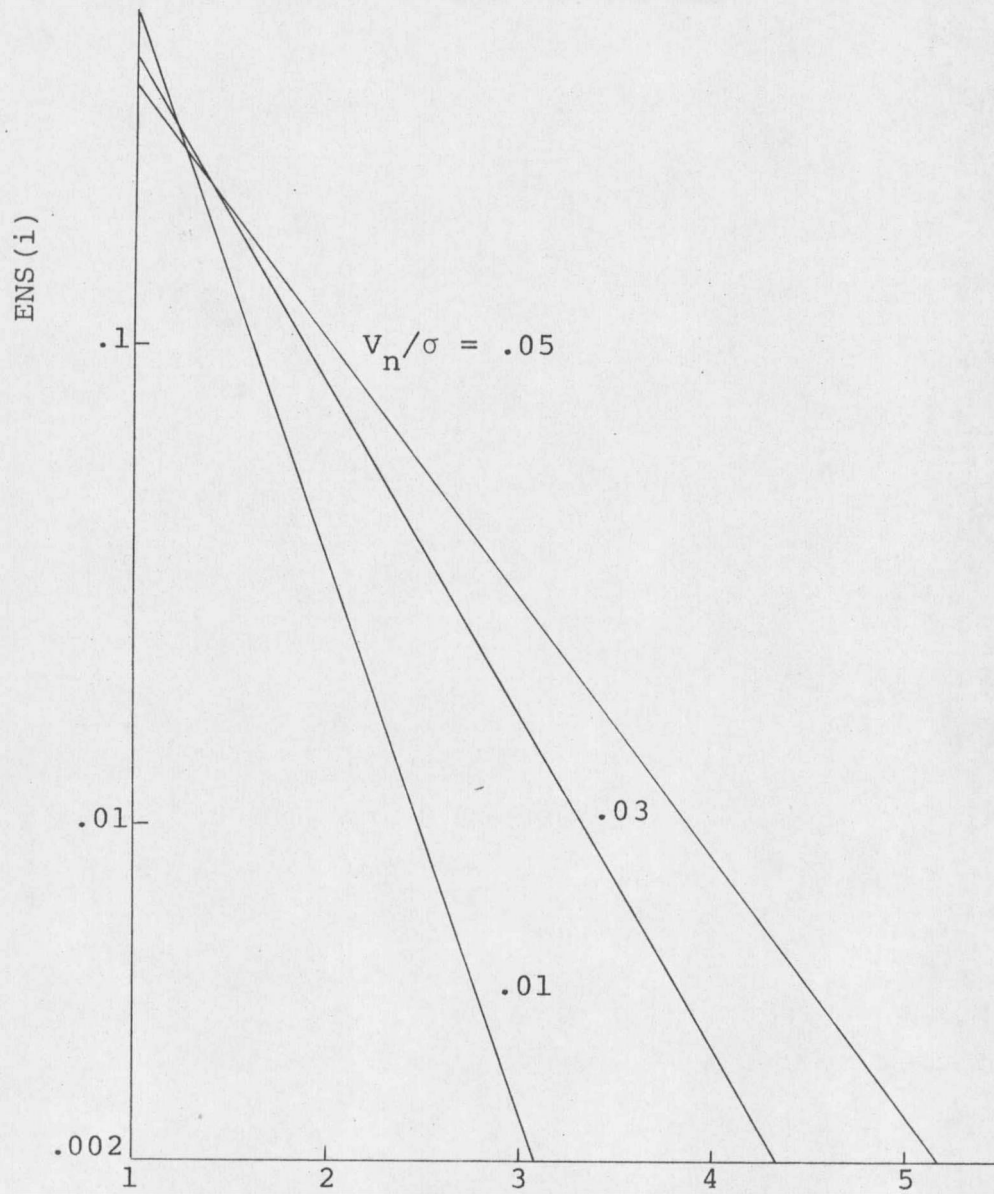


Figure 2.9 Delay efficiency of Example 2.1, $aT=0.01$.

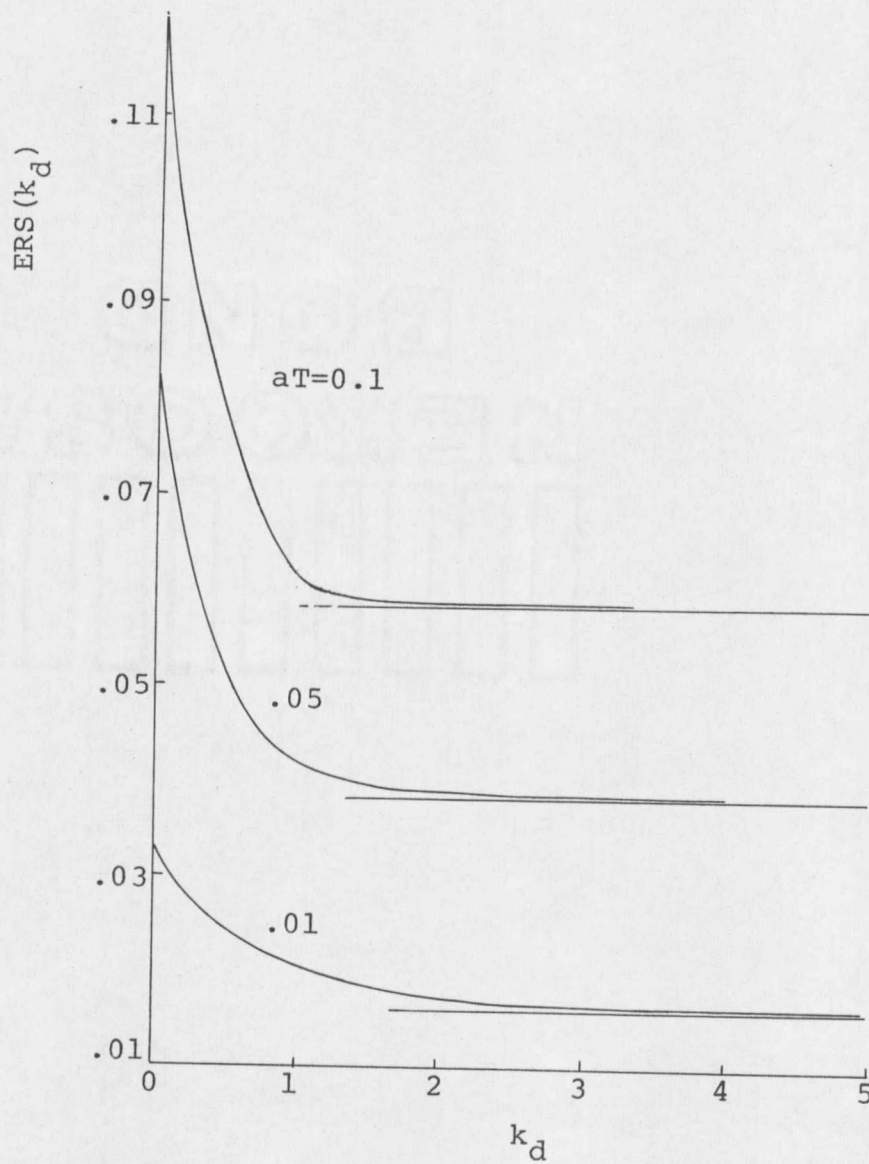


Figure 2.10 Normalized m.s.e. versus the amount of delay allowed for Example 2.1, $V_n/\sigma=0.05$.

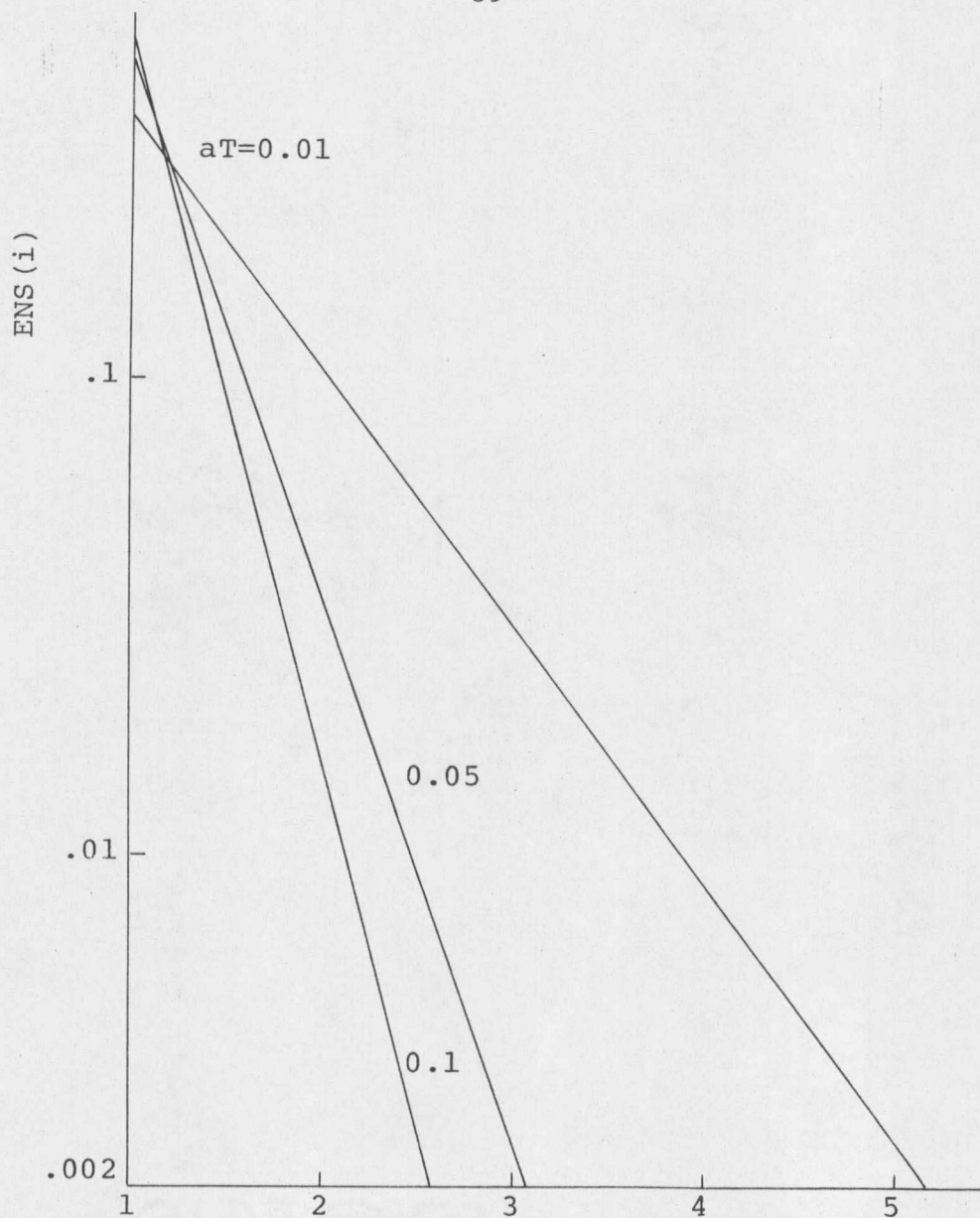


Figure 2.11 Delay efficiency of Example 2.1, $V_n/\sigma=0.05$.

Numerical data are given in Tables 2.1 and 2.2. The asymptotes shown in Figs. 2.8 and 2.10 indicate the performance of the optimal infinite-lag postfilter. Notice that different scales are used for different figures.

First of all, it can be seen from Tables 2.1 and 2.2 that the mean-square-error with an infinite-lag optimal postfilter is about half of the error of an optimal zero-lag postfilter in this example.

From Figs. 2.8 and 2.9, and Table 2.1, it is clear that as V_n increases, more delay is required for the performance of the optimal realizable filter to approximate that of the optimal unrealizable filter. Similar situation occurs as T decreases, as shown by Figs. 2.10 and 2.11, and Table 2.2. That is, delay efficiency converges more slowly to zero as V_n increases or as T decreases.

From Eq. (2.130), it follows that as q increases, $\text{ENS}(i)$ decreases more slowly with i ; which in turn means that more delay is required for the performance of a realizable filter to approach that of the corresponding optimum unrealizable one. Thus, we plot q versus V_n/σ , and q versus aT in Figs. 2.12 and 2.13, respectively. It is shown that q increases with V_n but decreases with T .

Table 2.1 Effect of delay on error reduction for Example 2.1, $aT=0.01$

V_n/σ	$k_{d,i}$	ERS(k_d)	EDS(i)	ENS(i) %
0.01	0	0.0172		
	1	.0084	0.0088	51.77
	2	.0077	.0006	3.67
	3	.0077	.0000	0.26
	4	.0077	.0000	0.02
	5	.0077	.0000	0.00
0.03	0	.0261		
	1	.0153	.0107	41.18
	2	.0132	.0022	8.39
	3	.0127	.0004	1.71
	4	.0126	.0000	0.35
	5	.0126	.0000	0.07
0.05	0	.0326		
	1	.0208	.0117	36.09
	2	.0174	.0034	10.39
	3	.0165	.0010	2.99
	4	.0162	.0003	0.86
	5	.0161	.0001	0.25

Table 2.2 Effect of delay on error reduction for Example 2.1, $V_n/\sigma=0.05$

aT	$k_{d,i}$	ERS(k_d)	EDS(i)	ENS(i) %
0.01	0	0.0326		
	1	.0208	0.0118	36.09
	2	.0174	.0034	10.39
	3	.0164	.0010	2.99
	4	.0162	.0003	0.86
	5	.0161	.0001	0.25
.05	0	.0826		
	1	.0416	.0405	49.59
	2	.0387	.0029	3.55
	3	.0385	.0002	0.25
	4	.0385	.0000	0.02
	5	.0385	.0000	0.00
.10	0	.1304		
	1	.0609	.0695	53.28
	2	.0589	.0020	5.16
	3	.0589	.0000	0.05
	4	.0589	.0000	0.00
	5	.0589	.0000	0.00

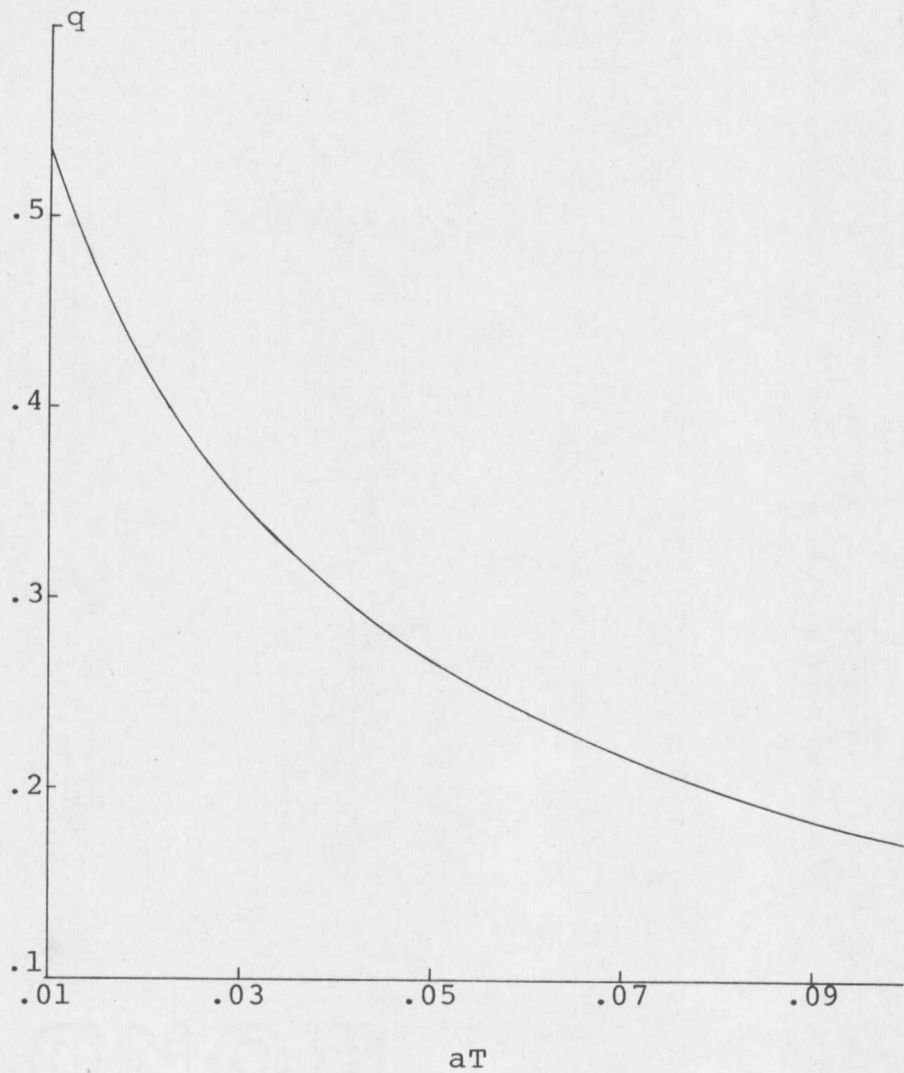


Figure 2.12 q versus aT , an indication of the rate of convergence of the delay efficiency for Example 2.1.

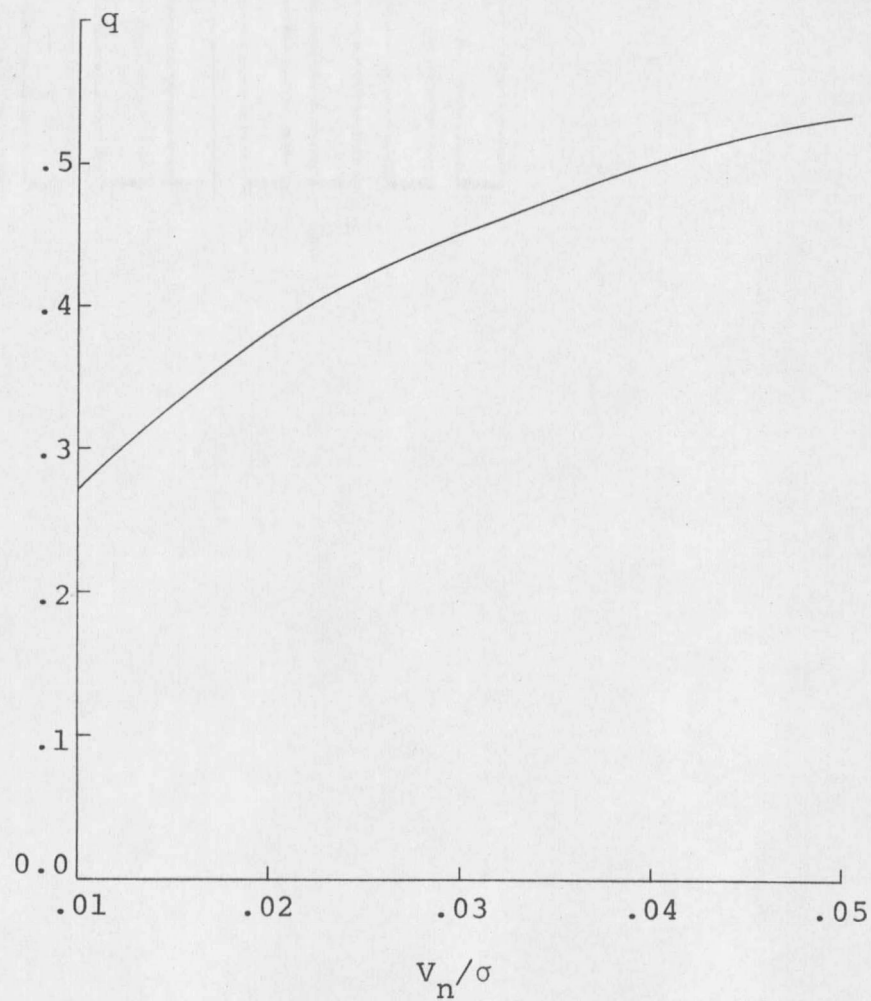


Figure 2.13 q versus V_n/σ , an indication of the rate of convergence of the delay efficiency for Example 2.1.

It is shown in Figs. 8 and 10 that the error reduction due to the first sample of delay dominates the overall error reduction--the error reduction with an infinite-lag in this example. As T increases, or as V_n decreases, this domination becomes more prominent. It is also shown that for this example, the performance of the fixed-lag optimal realizable postfilter approach that of the optimal unrealizable filter when the delay is of five samples or less.

Example 2.2: A first-order Gauss-Markov process with a colored noise. The desired signal and the observation models are the same as in Example 2.1, except now $n(t)$, the continuous version of $n(k)$ is characterized by the state equation (see the discussion in Appendix C)

$$\dot{n}(t) = -b n(t) + w_n(t)$$

where

$$E[w(t) w_n(\tau)] = 0$$

and

$$E[w_n(t) w_n(\tau)] = V_n \delta(t-\tau).$$

Augmenting the state equation by letting

$$x_a(t) = [x(t), n(t)]'.$$

Then

$$\dot{x}_a(t) = \begin{bmatrix} -a & 0 \\ 0 & -b \end{bmatrix} x_a(t) + \begin{bmatrix} w(t) \\ w_n(t) \end{bmatrix}$$

and

$$r(k) = [1 \ 1] x_a(k).$$

The corresponding discretized state equation is

$$x_a(k) = \begin{bmatrix} \alpha & 0 \\ 0 & \beta \end{bmatrix} x_a(k-1) + w_d(k-1),$$

where

$$w_d(k-1) = \int_{(k-1)T}^{kT} \begin{bmatrix} e^{-(kT-t)a} & 0 \\ 0 & e^{-(kT-t)b} \end{bmatrix} \cdot \begin{bmatrix} w(t) \\ w_n(t) \end{bmatrix} dt$$

and

$$\beta = \exp(-bT).$$

Following the same procedure used in Example 2.1, the results we obtain are

$$ERS(n_d) = \sigma - ((1-\alpha^2)/2T)(\sigma-\theta) - \sum_{i=1}^{n_d} EDS(i),$$

where

$$\theta = \frac{(k_1 - \beta)(1 - \alpha^2)}{(\alpha - \beta)(1 - \alpha k_1)} \cdot \sigma,$$

$$k_1 = (k_2 - k_3)/2k_4,$$

$$k_2 = \sigma(1 - \alpha^2)(1 + \beta^2) + V_n(1 - \beta^2)(1 + \alpha^2),$$

$$k_3 = k_2^2 - 4k_4,$$

$$k_4 = \sigma \cdot \beta(1 - \alpha^2) + V_n \cdot \alpha(1 - \beta^2),$$

and EDS(i) is obtained from (2.115), i.e.,

$$\begin{aligned} \text{EDS}(1) = (k_1/k_4) \cdot \{ [\theta^2 k_1^2 (\alpha - \beta)^2 - 2\theta k_1 (\alpha - \beta) \sigma - \alpha(k_1 - \beta)] \times \\ ((1 - \alpha^2)/2T) + \sigma^2 \alpha^2 (k_1 - \beta)^2 (1/\alpha^2 - \alpha^2) + 2\sigma \alpha (k_1 - \beta) \theta k_1 \} \end{aligned}$$

$$\begin{aligned} \text{EDS}(2) = (k_1/k_4) \cdot \{ [\theta^2 (\alpha - \beta)^2 - 2\theta (\alpha - \beta) \cdot \sigma \cdot \alpha] ((1 - \alpha^2)/2T) + \\ \sigma^2 \alpha^2 (1/\alpha^2 - \alpha^2) + 2\sigma \theta \alpha (\alpha - \beta) \}. \end{aligned}$$

and

$$\text{EDS}(i) = \text{EDS}(2) k_2^{2(i-2)} \quad \text{for } i \geq 2.$$

From (2.116), it follows that

$$\text{ENS}(1) = \frac{\text{EDS}(1)}{[\sigma - ((1 - \alpha^2)/2T) (\sigma - \theta)]}, \quad (2.131_a)$$

$$\text{ENS}(2) = \text{ENS}(1) (\text{EDS}(2)/\text{EDS}(1)), \quad (2.131_b)$$

and

$$\text{ENS}(i) = \text{ENS}(2) k_1^{2(i-2)} \quad \text{for } i \geq 2. \quad (2.131_c)$$

With $V_n/\sigma=0.1$, $aT=0.02$, and b varying from $7a$ to $13a$, plots of $\text{ERS}(n_d)$ versus n_d , and $\text{Log}(\text{ENS}(i))$ versus i are shown in Figs. 2.14 and 2.15, respectively, and numerical data are given in Table 2.3. With $V_n/\sigma=0.1$, $b=7a$, and aT varying from 0.02 to 0.1 , the same quantities are shown in Figs. 2.16 and 2.17, and numerical data given in Table 2.4. Finally, with $b=7a$, $aT=0.02$, and V_n/σ varying from 0.1 to 1 , we have Figs. 2.18 and 2.19, and Table 2.5.

The delay efficiency curves consist of two straight lines, as is also evident from (2.131).

As can be seen from Tables 2.3 through 2.4, the maximum error reduction that can be achieved (i.e., with infinite delay) varies from approximately 20% to 40% of the zero-lag error in this example.

It is shown in Fig. 2.15 that as b increases, delay becomes more efficient. This means that as the noise becomes whiter, delay becomes more desirable.

From Figs. 2.16 through 2.19, it is shown that the larger T or/and the smaller V_n are, the smaller the delay required for the performance of the optimal realizable

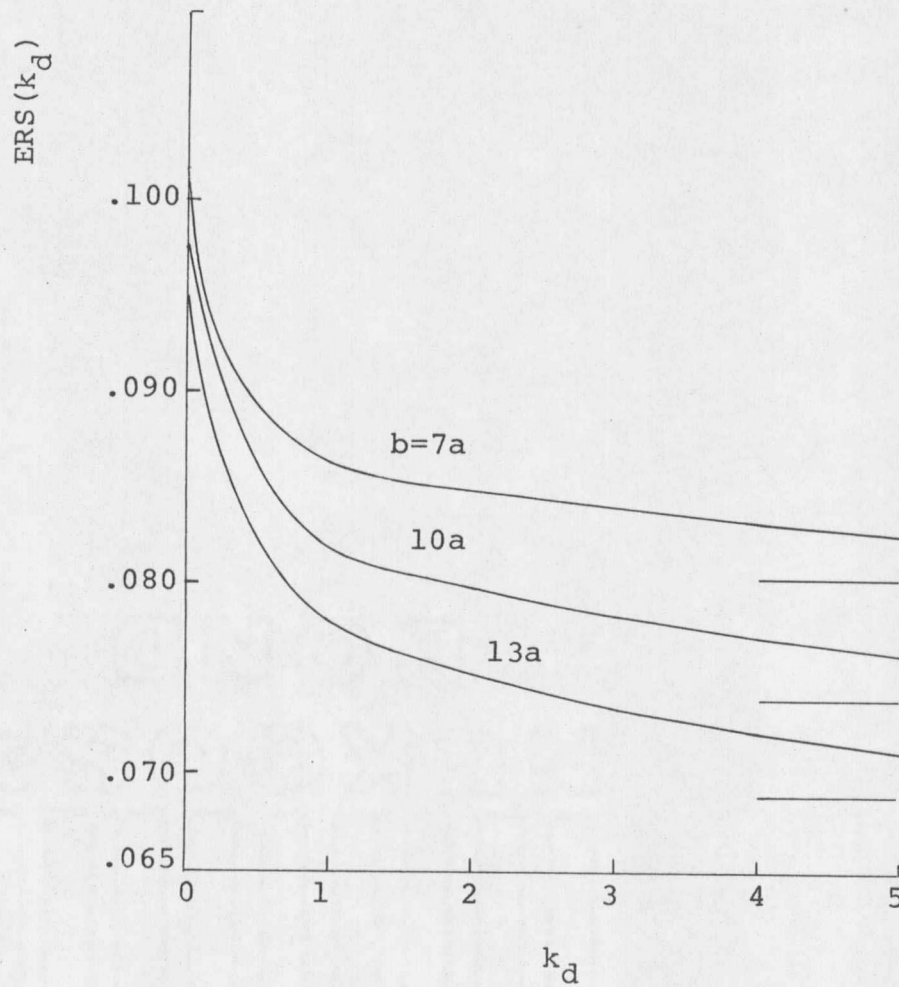


Figure 2.14 Normalized m.s.e. versus the amount of delay allowed for Example 2.2, $V_n/\sigma = 0.1$ and $aT=0.02$.

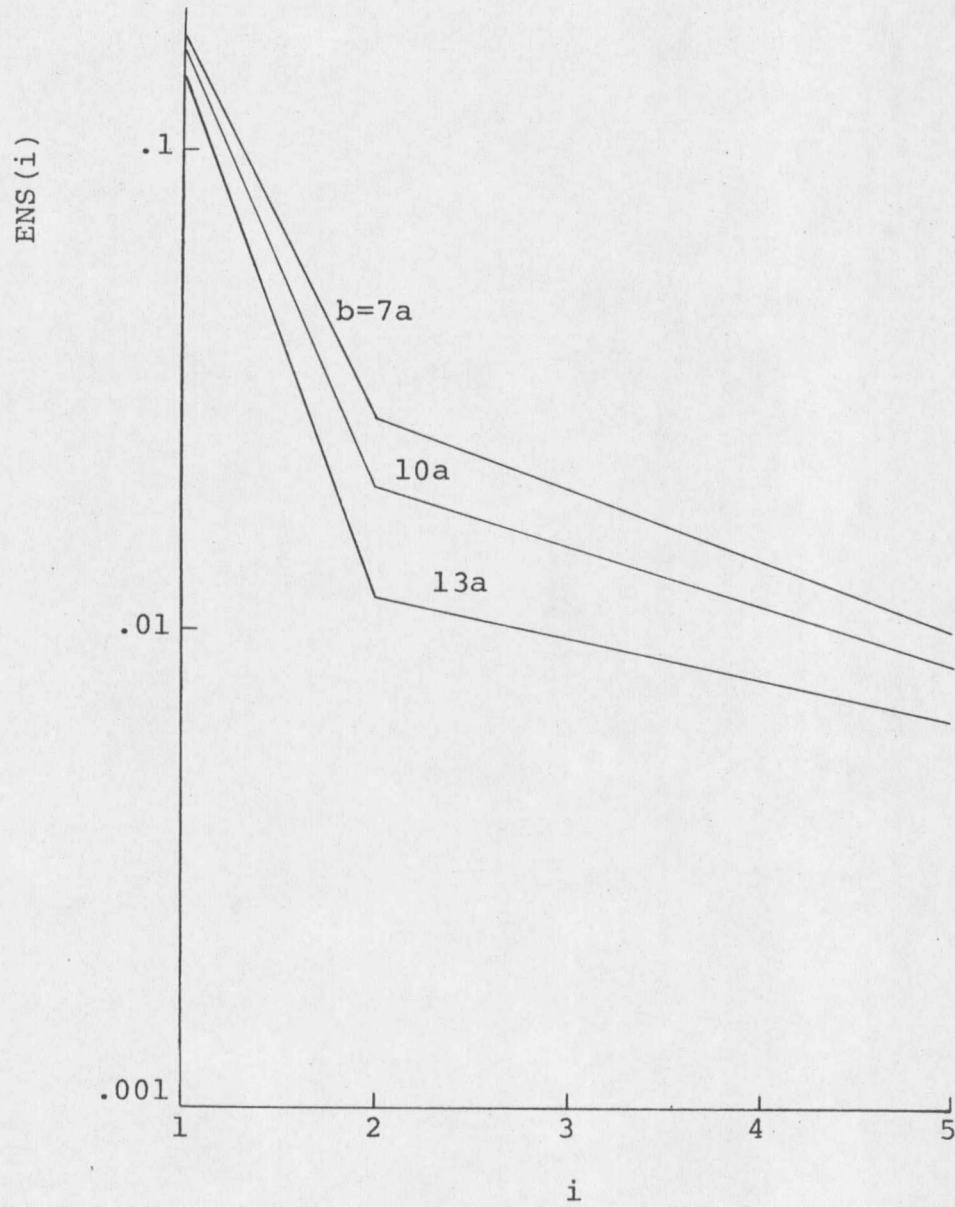


Figure 2.15 Delay efficiency of Example 2.2, $V_n/\sigma=0.1$ and $aT=0.02$.

Table 2.3 Effect of delay on error reduction for Example 2.2,
 $V_n/\sigma=0.1$, $aT=0.02$

b	$k_{d,i}$	ERS(k_d)	EDS(i)	ENS(i) %
7a	0	0.1009		
	1	.0863	0.0147	14.56
	2	.0851	.0012	1.17
	3	.0842	.0009	0.94
	4	.0834	.0008	.76
	5	.0828	.0006	.60
		.0802		
10a	0	.0975		
	1	.0818	.0157	16.15
	2	.0799	.0019	1.98
	3	.0784	.0015	1.49
	4	.0773	.0011	1.12
	5	.0765	.0008	0.84
		.0740		
13a	0	.0946		
	1	.0780	.0166	17.56
	2	.0753	.0026	2.76
	3	.0735	.0019	1.96
	4	.0722	.0013	1.39
	5	.0713	.0009	0.98
		.0689		

Table 2.4 Effect of delay on error reduction for Example 2.2,
 $b=7a$ and $V_n/\sigma=0.1$

aT	$k_{d,i}$	ERS(k_d)	EDS(i)	ENS(i) %
0.02	0	0.1009		
	1	.0863	0.0147	14.56
	2	.0851	.0012	1.17
	3	.0842	.0009	0.94
	4	.0834	.0008	0.76
	5	.0828	.0006	0.60
		.0802		
.06	0	.1359		
	1	.0943	.0157	30.64
	2	.0919	.0036	1.74
	3	.0907	.0012	0.91
	4	.0900	.0006	0.48
	5	.0897	.0003	0.25
		.0893		
.10	0	.1690		
	1	.1035	.0656	38.80
	2	.1007	.0027	1.62
	3	.0998	.0009	0.55
	4	.0995	.0003	0.19
	5	.0994	.0002	0.06
		.0993		

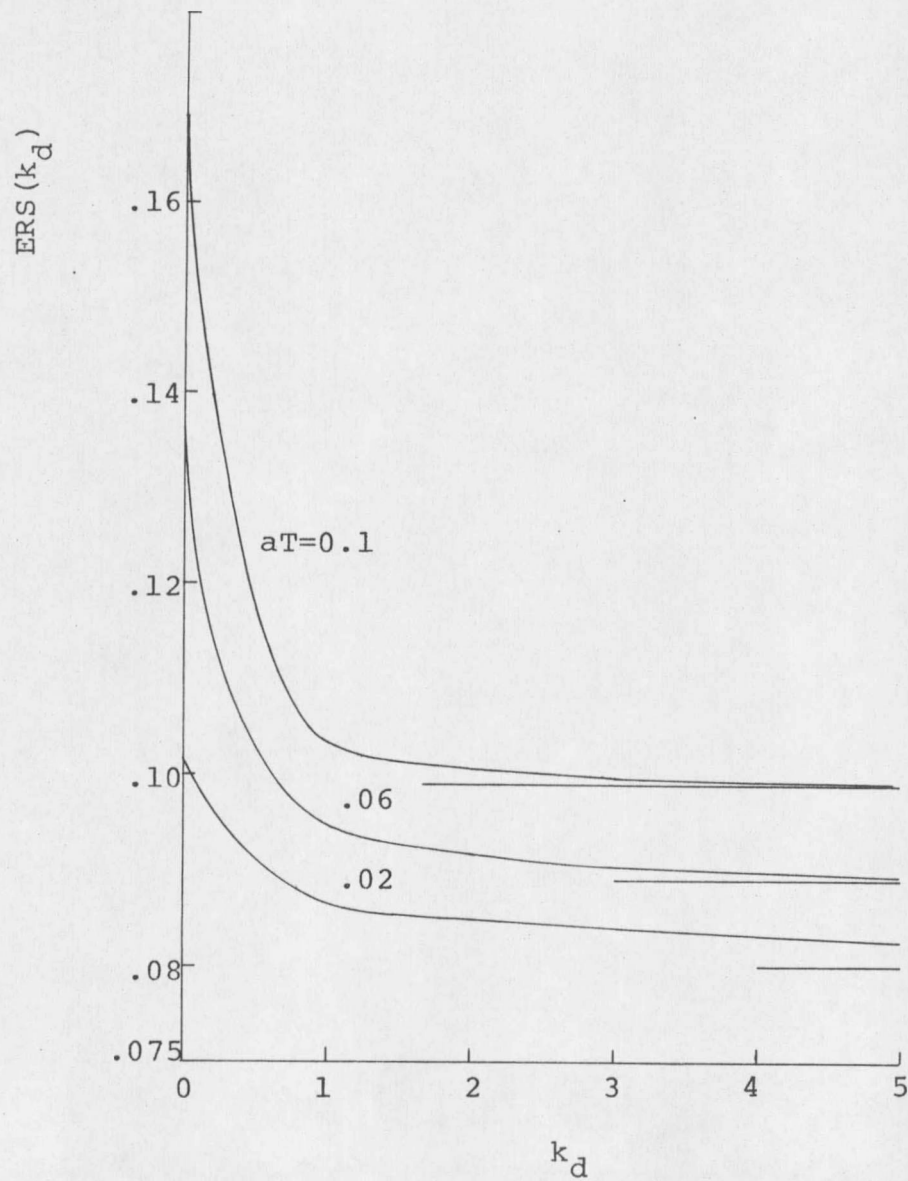


Figure 2.16 Normalized m.s.e. versus the amount of delay allowed for Example 2.2, $V_n/\sigma=0.1$ and $b=7a$.

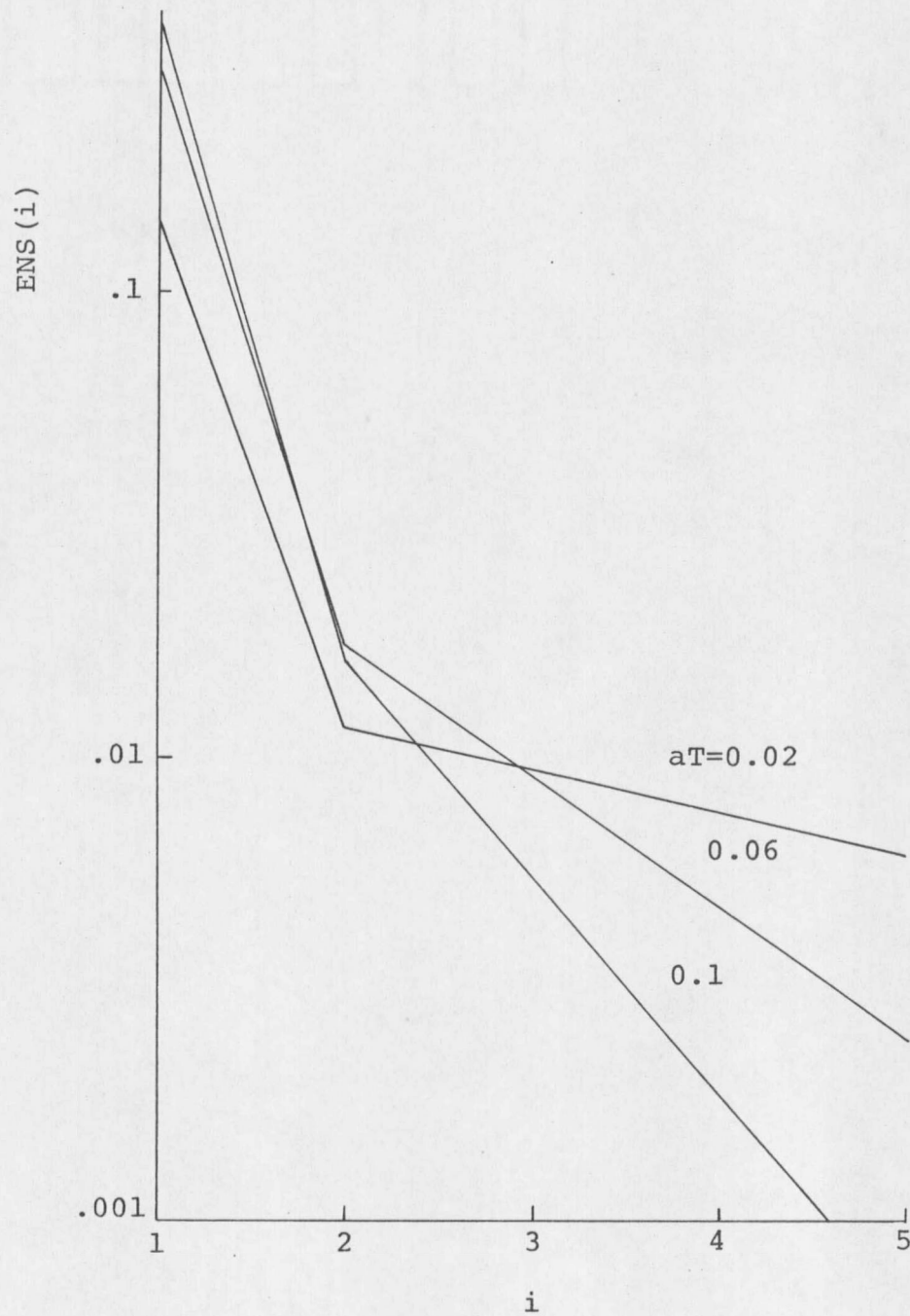


Figure 2.17 Delay efficiency of Example 2.2, $V_n/\sigma=0.1$ and $b=7a$.

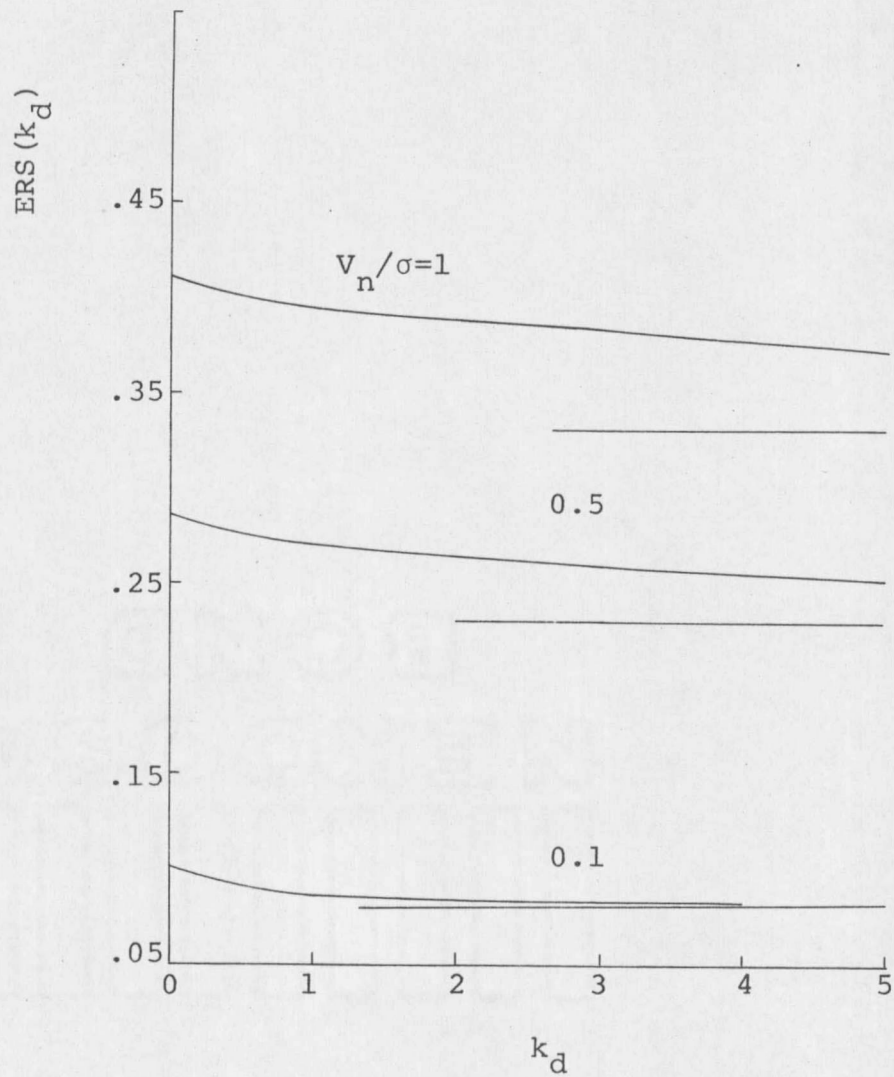


Figure 2.18 Normalized m.s.e. versus the amount of delay allowed for Example 2.2, $aT=0.02$ and $b=7a$.

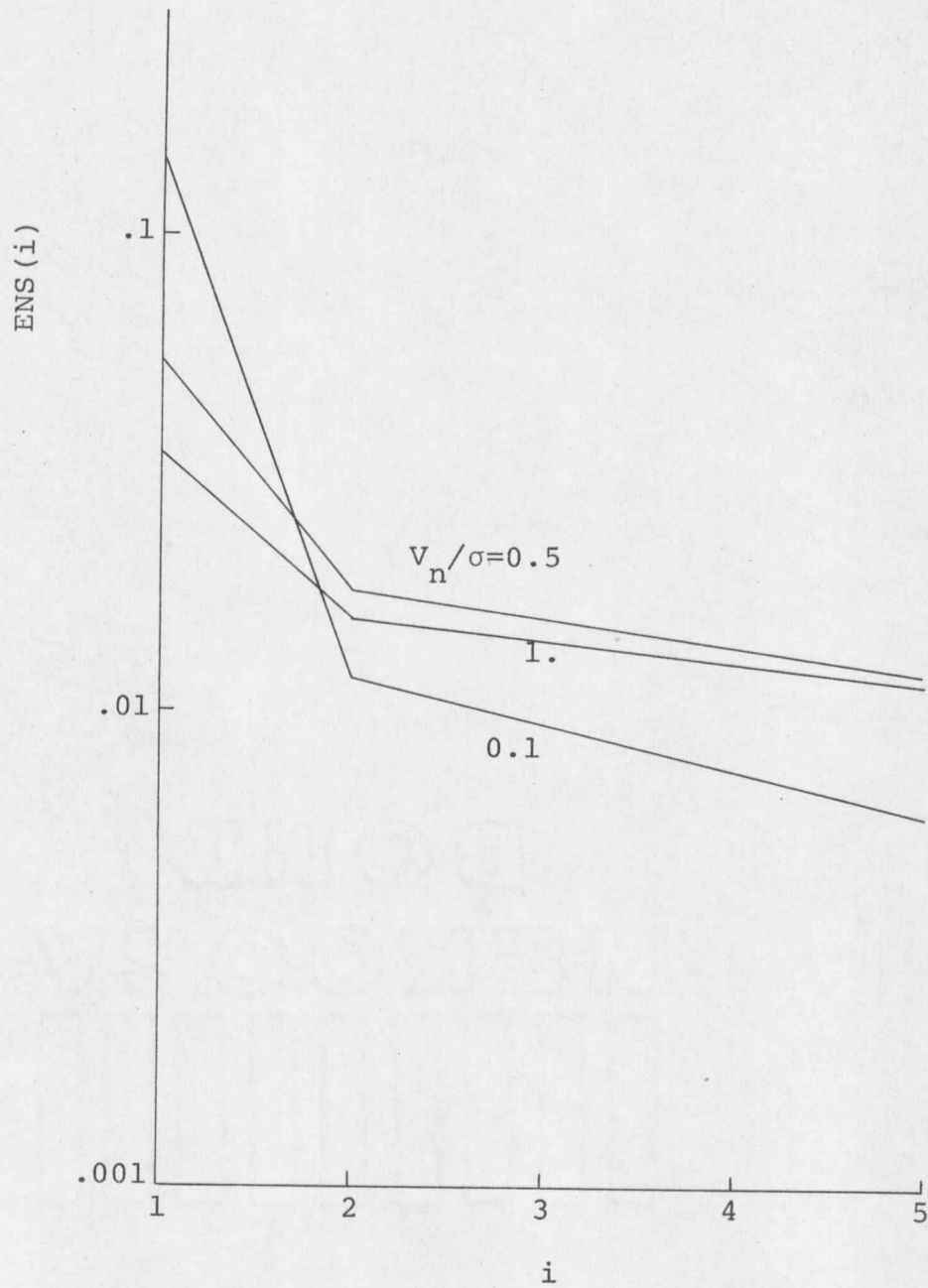


Figure 2.19 Delay efficiency of Example 2.2, $aT=0.02$ and $b=7a$.

Table 2.5 Effect of delay on error reduction for
Example 2.2, $aT=0.02$ and $b=7a$

V_n/σ	$k_{d,i}$	ERS (k_d)	EDS (i)	ENS (i) %
0.1	0	0.1009		
	1	.0863	0.0147	14.56
	2	.0851	.0012	1.17
	3	.0842	.0009	0.94
	4	.0834	.0008	0.76
	5	.0828	.0006	0.60
		.0803		
0.5	0	.2851		
	1	.2695	.0155	5.45
	2	.2644	.0051	1.79
	3	.2600	.0045	1.56
	4	.2560	.0039	1.37
	5	.2527	.0034	1.19
		.2295		
1.0	0	.4101		
	1	.3958	.0143	3.49
	2	.3984	.0064	1.56
	3	.3836	.0057	1.40
	4	.3784	.0052	1.26
	5	.3738	.0047	1.14
		.3319		

Table 2.6 Effect of delay on error reduction for
Example 2.3

$k_{d,i}$	ERS (k_d)	EDS (i)	ENS (i) %
0	0.5327		
1	.3373	.1954	36.67
2	.3353	.0020	0.39
3	.3349	.0004	0.06
4	.3349	.0000	0.00
5	.3349	.0000	0.00
	.3349		

filter to approximate that of the optimal unrealizable filter.

In this Example, the amount of error reduction due to the first sample of delay still contributes most of the overall error reduction. Also, the error reduction due to subsequent samples of delay converges more slowly to zero than in the white-noise case.

Example 2.3: A second-order Gauss-Markov process without noise. Related equations are given by (2.1) through (2.8) with $n(k)=0$, $d(t)=z(t)$, $\Gamma(t)=$ the identity matrix, $H(t)=[1, 1]$, and

$$A(t) = \begin{bmatrix} -a & 0 \\ 0 & -b \end{bmatrix},$$

$$V_w(t) = \begin{bmatrix} 2a\sigma_a & 0 \\ 0 & 2b\sigma_b \end{bmatrix},$$

$$R_x(0,0) = \begin{bmatrix} \sigma_a & 0 \\ 0 & \sigma_b \end{bmatrix},$$

and $V_z(0)=\sigma=\sigma_a+\sigma_b$. This problem is similar to Example 2.2. It is found that, in this Example, most of the error reduction is achieved by the first sample of delay for

various combinations of parameters. With $b=10a$, $aT=0.3$, $\sigma_b=0.6\sigma_a$, plots of $ERS(n_d)$ versus n_d , and $\text{Log}(ENS(i))$ versus i are shown in Figs. 2.20 and 2.21, respectively.

For more complicated problems, the simple relationships exhibited by (2.130) and (2.131) no longer exist. Hand calculation is still possible, but with additional effort. In those cases, computer solutions are more feasible.

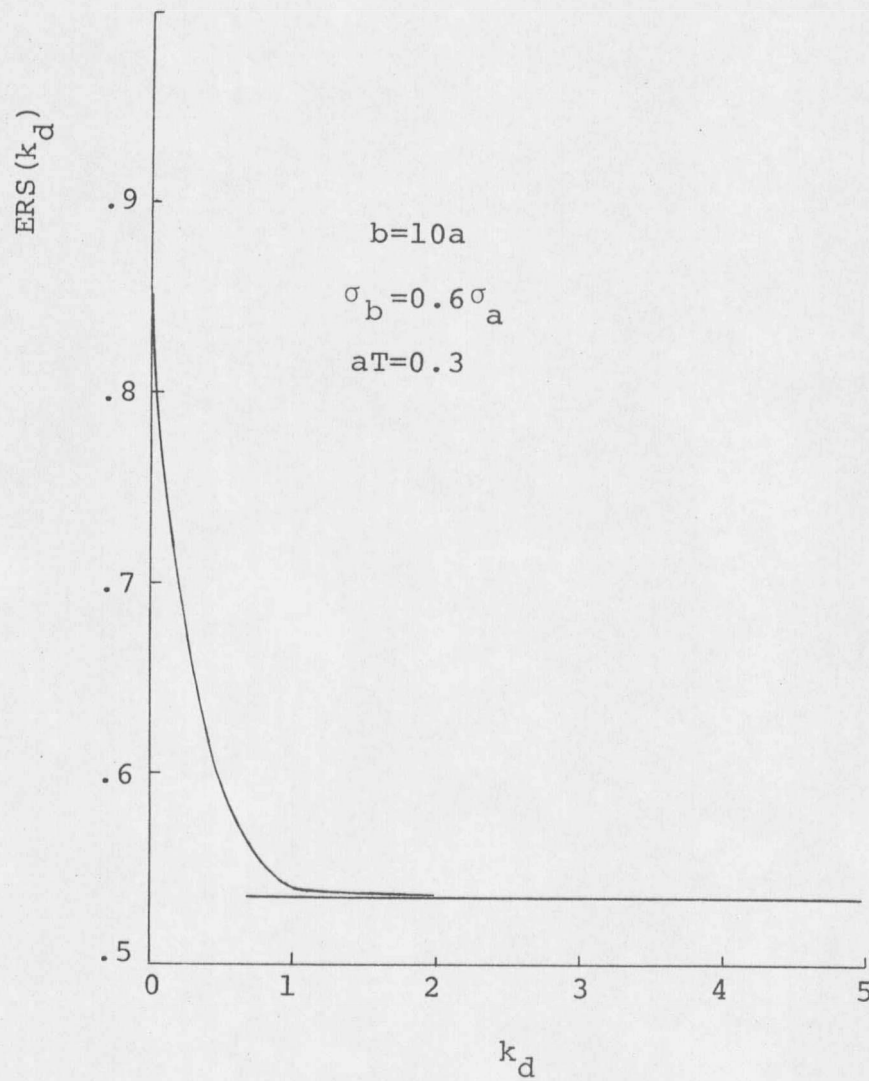


Figure 2.20 Normalized m.s.e. versus the amount of delay allowed for Example 2.3.

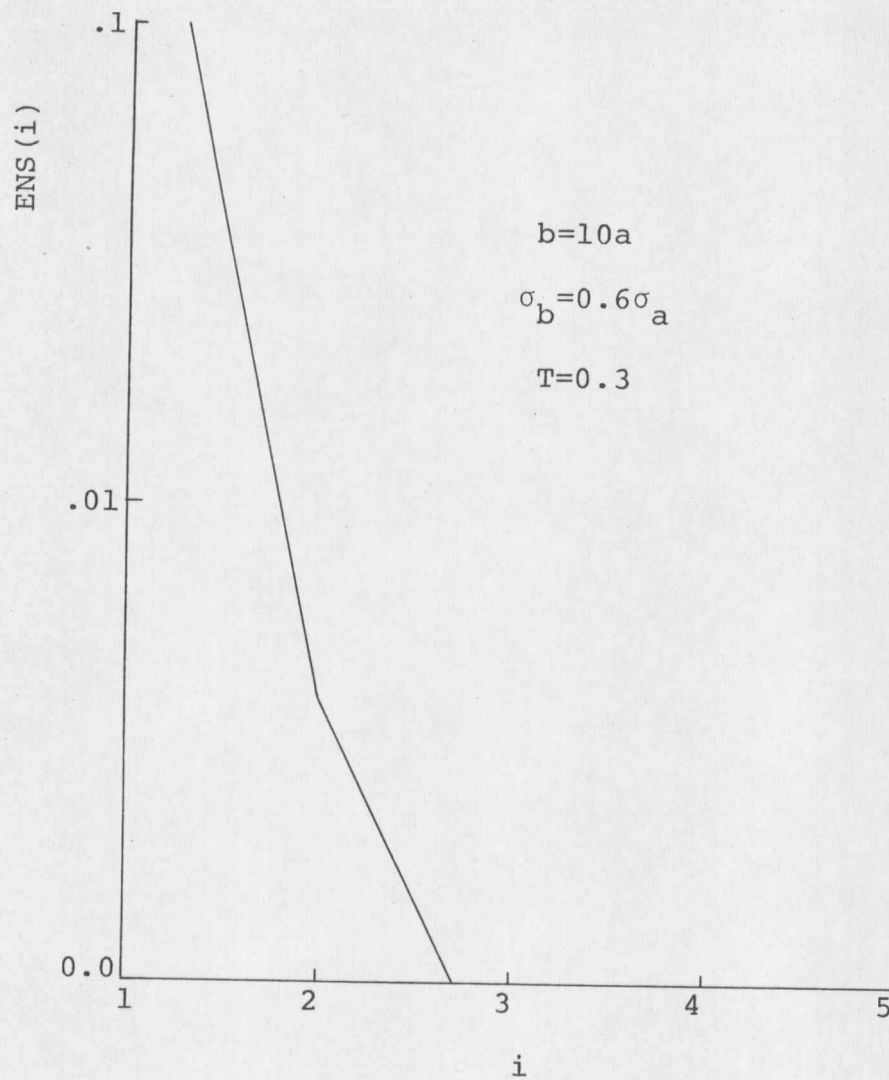


Figure 2.21 Delay efficiency of Example 2.3.

Chapter 3

THE OPTIMAL PREFILTER

3.1 Problem Statement and Preliminary Remarks

The system under consideration is shown in Fig. 1.1. Related equations are

$$\dot{x}(t) = A x(t) + \Gamma w(t), \quad (3.1)$$

$$z(t) = H x(t), \quad (3.2)$$

$$r(t) = z(t) + n_1(t), \quad (3.3)$$

$$y(t) = \int_{T_i}^{\infty} F(t, v) r(v) dv, \quad (3.4)$$

$$s(k) = y(k) + n_2(k), \quad (3.5)$$

$$d(t) = C x(t) \triangleq \text{the desired signal}, \quad (3.6)$$

and

$$d^0(t) = \sum_{k=\ell_i}^{\infty} G(t, kT) s(k), \quad (3.7)$$

where $k \triangleq kT$. (3.1), (3.2), and (3.6) are similar to (2.1), (2.2), and (2.4), respectively, except that A , Γ , H , and C are constrained to be constants in this case. x , w , z , y , and d are vectors (column matrices), and A , Γ , H , C , F , and G are matrices of appropriate dimensions. Prior statistics are

$$E[\dot{n}_2(k)] = E[n_1(t)] = E[x(0)] = E[w(t)] = 0, \quad (3.8)$$

$$R_w(\tau) = E[w(t) w(t-\tau)] = V_w(t) \delta(t-\tau), \quad (3.9)$$

$$R_{wx}(\tau) = E[w(t) x(t-\tau)] = 0 \quad \text{for } \tau > 0, \quad (3.10)$$

and

$$R_{wn_1}(\tau) = R_{wn_2}(\tau) = R_{n_1 n_2}(\tau) = 0 \quad (3.11)$$

where

$$R_{\xi\eta}(\tau) = E[\xi(t) \eta(t-\tau)] \quad (3.12)$$

and

$$R_{\xi}(\tau) = E[\xi(t) \xi(t-\tau)]. \quad (3.13)$$

Also, $R_x(0)$ and V_w are known. General constraints of this chapter are

Constraint 3.1. The signal and noises are stationary, and the filters are time-invariant.

Constraint 3.2. As shown by Eq. (3.11), $n_1(t)$, $n_2(kT)$, and $w(t)$ are uncorrelated with one another, where n_1 is the primary observation noise, and n_2 represents quantization noise or/and channel additive noise.

Constraint 3.3. Realizability of the filters is required.

Constraint 3.4. The sampling process is periodic.

Constraint 3.5. The system is at steady state; i.e., $\ell_i, T_i \rightarrow -\infty$ (29). With the postfilter given, the problem is to design an optimal prefilter for minimizing the mean square error between $d(t)$ and $d^o(t)$.

Refer to Fig. 1.1, there can be multiple sampled-data links between the two filters. By this we mean that multiple samplers, quantizers, and channels can be used to transmit $y(k)$, i.e., $y(k)$ can be a vector of dimension > 1 .

A prefilter essentially performs two functions:

1. To combat the loss of information due to sampling. The work of Chang (9) and Petersen et al. (42) essentially deals with this problem.

2. To combat the loss of information due to additive channel noise or/and quantization noise. The work of Hayes (20) and Costas et al. (10) belongs in this category.

Several other researchers, such as Kellog (31,32) and Chan and Donaldson (7,8) tried to treat both problems.

Derusso (12) suggested that the optimum prefilter in the case with a single sampled-data link could simply

be a scalar Wiener filter, which was later disapproved by Chang (9). Petersen (42) proposed that the prefilter of a jointly optimized pre- and postfilters pair can be represented by the combination of a continuous KSVF (Kalman State Vector Filter) followed by a coder which is simply a gain matrix. We shall prove later that this is not true in general except for the special case when $n_2=0$. But this structure of the prefilter is intuitively appealing. We shall show later that the optimal prefilter, not necessarily the jointly optimized one, can be represented by the cascade of a continuous KSVF followed by a coder where the impulse response of the coder is an impulse train in general instead of a single impulse at the origin, as it is the case in which the coder is a gain matrix.

Let

$\hat{\xi}(t/T_i, T_f, \eta(\tau))$ = the linear least-square estimate
of $\xi(t)$ based on the observation
of $\eta(\tau)$ for $T_i \leq \tau \leq T_f$.

The $\hat{\xi}(t/T_i, T_f, \eta(\tau))$ defined above is similar to the $\hat{\xi}(t/l_i, l_f, \eta(k))$ defined in Chapter 2 except that $\hat{\xi}(t/T_i, T_f, \eta(\tau))$ represents the optimal estimation based on continuous observation (continuous in time), whereas

$\hat{\xi}(t/l_i, l_f, \eta(k))$ represents the optimal estimation based on discrete estimation. We shall let $\hat{\xi}(t/T_f, \eta(\tau))$ denotes $\hat{\xi}(t/-\infty, T_f, \eta(\tau))$ and $\hat{\xi}(t/l_f, \eta(k))$ denotes $\hat{\xi}(t/-\infty, l_f, \eta(k))$ from here on.

The basic idea behind Petersen's work is the following. Suppose one is given $\hat{x}(t_0/t_0, r(\tau))$, the best estimate of $x(t)$ for $t \geq t_0$ is then given by (51, pp. 349)

$$\hat{x}(t/t_0, r(\tau)) = \Phi(t, t_0) \hat{x}(t_0/t_0, r(\tau)) \quad (3.14)$$

from (C.1) since $w(\tau)$ is independent of $r(v)$ for $\tau > v$, where Φ is the transition matrix of A . It is due to Eq. (3.14) that enables one to say that as far as estimating $x(t)$ for $t \geq t_0$ based on $r(\tau)$ for $\tau \leq t_0$ is concerned, $\hat{x}(t_0/t_0, r(\tau))$ provides all the useful information contained in $r(\tau)$ for $\tau \leq t_0$. Thus, if there is no constraint on the number of sampled-data links, and if $n_2=0$; one can use a continuous KSVF to obtain $y(t) = \hat{x}(t/t, r(\tau))$. Then transmit the whole vector $\hat{x}(kT/kT, r(\tau))$ after sampling through the channel. Finally, Eq. (3.14) is used to reconstruct $x(t)$ for $kT \leq t < (k+1)T$; i.e., since (refer to Fig. 1.1)

$$s(k) = y(k) = \hat{x}(kT/kT, r(\tau)), \quad (3.15)$$

we have

$$\begin{aligned}
 \hat{x}(t/m, s(K)) &= \hat{x}(t/m, y(k)) \\
 &= \hat{x}[t/m, \hat{x}(kT/kT, r(\tau))] \\
 &= \Phi(t, mT) \hat{x}(mT/mT, r(\tau)) \\
 &= \Phi(t, mT) s(m),
 \end{aligned} \tag{3.16}$$

from (3.14) and (3.15), where $m = \lfloor t/T \rfloor$. Notice that $\hat{x}(t/m, s(k))$ represents the optimal estimation based on the discrete observation of the $s(k)$'s, whereas $\hat{x}(mT/mT, r(\tau))$ represents the optimal estimation based on the continuous observation of the $r(t)$. In this case, the whole system acts as if it is an optimum fixed-interval (of interval T) prediction system in the absence of sampling. The only information that is lost at time t due to sampling is $r(v)$ for $mT < v < (m+1)T$; whereas the information about $r(v)$ for $v \leq mT$ is retained despite the periodic sampling process. It should be intuitively clear that in this case, the continuous KSVF used as the prefilter is optimum (jointly optimized with the postfilter). This point is implicit in Petersen's work, and we proceed to prove it analytically in Section 3.2.

If $n_2 \neq 0$, and the continuous KSVF is still used as a prefilter, then

$$s(k) = \hat{x}(kT/kT, r(\tau)) + n_2(k), \quad (3.17)$$

and

$$\begin{aligned} \hat{x}(t/m, s(k)) &= \hat{x}\{t/[m, (\hat{x}(kT/kT, r(\tau)) + n_2(k))]\} \\ &\neq \hat{x}(t/m, \hat{x}(kT/kT, r(\tau))). \end{aligned} \quad (3.18)$$

Thus, Eq. (3.14) is no longer applicable in this case. Hence, due to the disturbance of n_2 , $\hat{x}(kT/kT, r(\tau))$ might not represent all the useful information about $r(\tau)$ for $\tau \leq kT$. Instead, one tries to operate on $r(t)$ in such a way not only to protect the information contained in $r(t)$ against sampling, but also to protect it against the contamination of n_2 .

Conditions for the optimal prefilter with a given postfilter are given in Appendix E.

3.2 Jointly Optimized Pre- and Postfilters with Multiple Sampled-Data Links

Additional constraints in this section are:

Constraint 3.6. $n_2(k) = 0$.

Constraint 3.7. The number of sampled-data links is equal to n , the number of states.

With Constraints 3.1 through 3.7, the equation that specifies the optimal prefilter can be obtained from (E.31); the resulting equation with $d(t) = x(t)$ is

$$\begin{aligned}
& \left[\int_0^\infty \int_0^\infty R_{xr}(\beta+v) \delta F'(\beta) G'(v) dv d\beta - \right. \\
& \left. \sum_{\ell=-\infty}^{\infty} \int_0^\infty \int_0^\infty \int_0^\infty G(v-\ell T) F_0(\alpha) R_r(\ell T+\alpha-\beta) \delta F'(\beta) \right. \\
& \left. G'(v) dv d\alpha d\beta \right]_{dia} = 0
\end{aligned} \tag{3.19}$$

where $F_0(\alpha)$ denotes the optimum prefilter and $\delta F(\beta)$ is the variation defined by (E.11), and

$$0 \wedge \ell T = \begin{cases} \ell T & \text{if } \ell T \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{3.20}$$

Since $n_2=0$, if we let

$$G(t) = \begin{cases} \Phi(t) & \text{if } 0 \leq t < T \\ 0 & \text{otherwise,} \end{cases} \tag{3.21}$$

then $G(t)$ will be the optimum postfilter if

$y(t) = \hat{x}(t/t, r(\tau))$ (see Fig. 1.1_c, where $s(k) = y(k)$ since $n_2=0$). Thus, all we need to prove is that the $F_0(t)$ specified by (3.19) is the continuous KSVF. Then both the pre- and postfilters are jointly optimized. Substituting (3.21) into (3.19) to obtain

$$\left[\int_0^T \int_0^\infty \{ R_{xr}(\beta+v) - \Phi(v) \int_0^\infty F_0(\alpha) R_r(\beta-\alpha) d\alpha \} \delta F'(\beta) \Phi'(v) d\beta dv \right]_{dia} = 0. \quad (3.22)$$

From (3.1), it follows that

$$x(t) = \Phi(t-\tau) x(\tau) + \int_\tau^t \Phi(t-v) \Gamma w(v) dv. \quad (3.23)$$

Thus,

$$\begin{aligned} R_x(t-\tau) &= E[x(t) x(\tau)] \\ &= \Phi(t-\tau) R_x(0) + \int_\tau^t \Phi(t-v) \Gamma E[w(v) x(\tau)] dv. \end{aligned} \quad (3.24)$$

Since $w(v)$ is independent of $x(\tau)$ for $v > \tau$, $E[w(v) x(\tau)] = 0$ for $v > \tau$; and (3.24) becomes, for $t-\tau \geq 0$,

$$R_x(t-\tau) = \Phi(t-\tau) R_x(0). \quad (3.25)$$

Since $n_1(t)$ is independent of the signal,

$$\begin{aligned} R_{xr}(\beta+v) &= E[x(\beta+v) r(0)] \\ &= E[x(\beta+v) z(0)] + E[x(\beta+v) n_1(0)] \\ &= R_x(\beta+v) H'. \end{aligned} \quad (3.26)$$

Lemma 3.1 $\Phi(t) \Phi(v) = \Phi(v) \Phi(t) = \Phi(t+v)$.

Proof. This is simply one of the properties of a transition matrix (see Eq. (C.3)).

From Eq. (3.25) and Lemma 3.1, since $\beta, v > 0$,

$$\begin{aligned} R_X(\beta+v) &= \Phi(\beta+v) R_X(0) \\ &= \Phi(v) \Phi(\beta) R_X(0) \\ &= \Phi(v) R_X(\beta) \end{aligned} \quad (3.27)$$

Substituting (3.27) into (3.26) to obtain

$$R_{xr}(\beta+v) = \Phi(v) R_{xr}(\beta). \quad (3.28)$$

Substituting (3.28) into (3.22) yields

$$\left\{ \int_0^T \int_0^\infty \Phi(v) [R_{xr}(\beta) - \int_0^\infty F_O(\alpha) R_r(\beta-\alpha) d\alpha] \delta F'(\beta) \Phi'(v) d\beta dv \right\}_{dia} = 0. \quad (3.29)$$

It is obvious from (3.29) that it is sufficient to let

$$R_{xr}(\beta) - \int_0^\infty F_O(\alpha) R_r(\beta-\alpha) d\alpha = 0 \quad (3.30)$$

To illustrate that (3.30) is also a necessary condition for (3.39) to hold, let us look at the special case when the dimension of x is equal to 2 and $r(t)$ is a scalar.

In this case, let

$$\Phi(v) = \begin{bmatrix} \Phi_1(v) & \Phi_2(v) \\ \Phi_3(v) & \Phi_4(v) \end{bmatrix}, \quad (3.31)$$

$$\delta F = [\delta f_1, \delta f_2]', \quad (3.32)$$

and

$$R_{xr}(\beta) - \int_0^\infty F_O(\alpha) R_r(\alpha-\beta) d\alpha = [B_1(\beta), B_2(\beta)]'. \quad (3.33)$$

Substituting (3.31) through (3.33) into (3.29) to obtain

$$\int_0^T \int_0^\infty \begin{bmatrix} \Phi_1(v) & \Phi_2(v) \\ \Phi_3(v) & \Phi_4(v) \end{bmatrix} \begin{bmatrix} B_1(\beta) \\ B_2(\beta) \end{bmatrix} [\delta f_1(\beta), \delta f_2(\beta)] \begin{bmatrix} \Phi_1(v) & \Phi_2(v) \\ \Phi_3(v) & \Phi_4(v) \end{bmatrix} d\beta dv \stackrel{\text{dia}}{=} \\ = \begin{bmatrix} \int_0^\infty \{ \delta f_1(\beta) [B_1(\beta) (\int_0^T \Phi_1^2(v) dv) + B_2(\beta) (\int_0^T \Phi_1(v) \Phi_2(v) dv)] \\ + \delta f_2(\beta) [B_1(\beta) (\int_0^T \Phi_1(v) \Phi_2(v) dv) + B_2(\beta) (\int_0^T \Phi_2^2(v) dv)] \} d\beta \\ \int_0^\infty \{ \delta f_1(\beta) [B_1(\beta) (\int_0^T \Phi_3^2(v) dv) + B_2(\beta) (\int_0^T \Phi_3(v) \Phi_4(v) dv)] \\ + \delta f_2(\beta) [B_1(\beta) (\int_0^T \Phi_3(v) \Phi_4(v) dv) + B_2(\beta) (\int_0^T \Phi_4^2(v) dv)] \} d\beta \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (3.34)$$

Since $\delta f_1(\beta)$ and $\delta f_2(\beta)$ are arbitrary functions of β , the only possible solution of (3.34) is

$$B_1(\beta) = B_2(\beta) = 0 \quad (3.35)$$

Thus, (3.30) is also a necessary condition for (3.29) to hold. It is clear that a similar argument can be applied

to higher-order systems. Since (3.30) is the vector Wiener equation, we have $y(t) = \hat{x}(t/t, r(\tau))$, and the proof is completed. The block diagram of this system is shown in Fig. 3.1.

It is possible to prove that the above system with the number of sampled-data links equal to the dimension of the state vector is the optimum system in the sense that fewer sampled-data links with the jointly optimized pre- and postfilters will result in an increase of distortion, whereas increasing the number of sampled-data links will not reduce the error.

In the case in which there is only one sampled-data link available, the question is if one can, in some way, code the components of $\hat{x}(kT/kT, r(\tau))$ before rather than after sampling (as Eq. (3.16) does) so that the performance of this system is close to the system with multiple sampled-data links. In this case, at the receiving end, one will receive a scalar quantity in the form

$$s(j) = \int_{-\infty}^{jT} N(jT-v) \hat{x}(v/v, r(\tau)) dv, \quad (3.36)$$

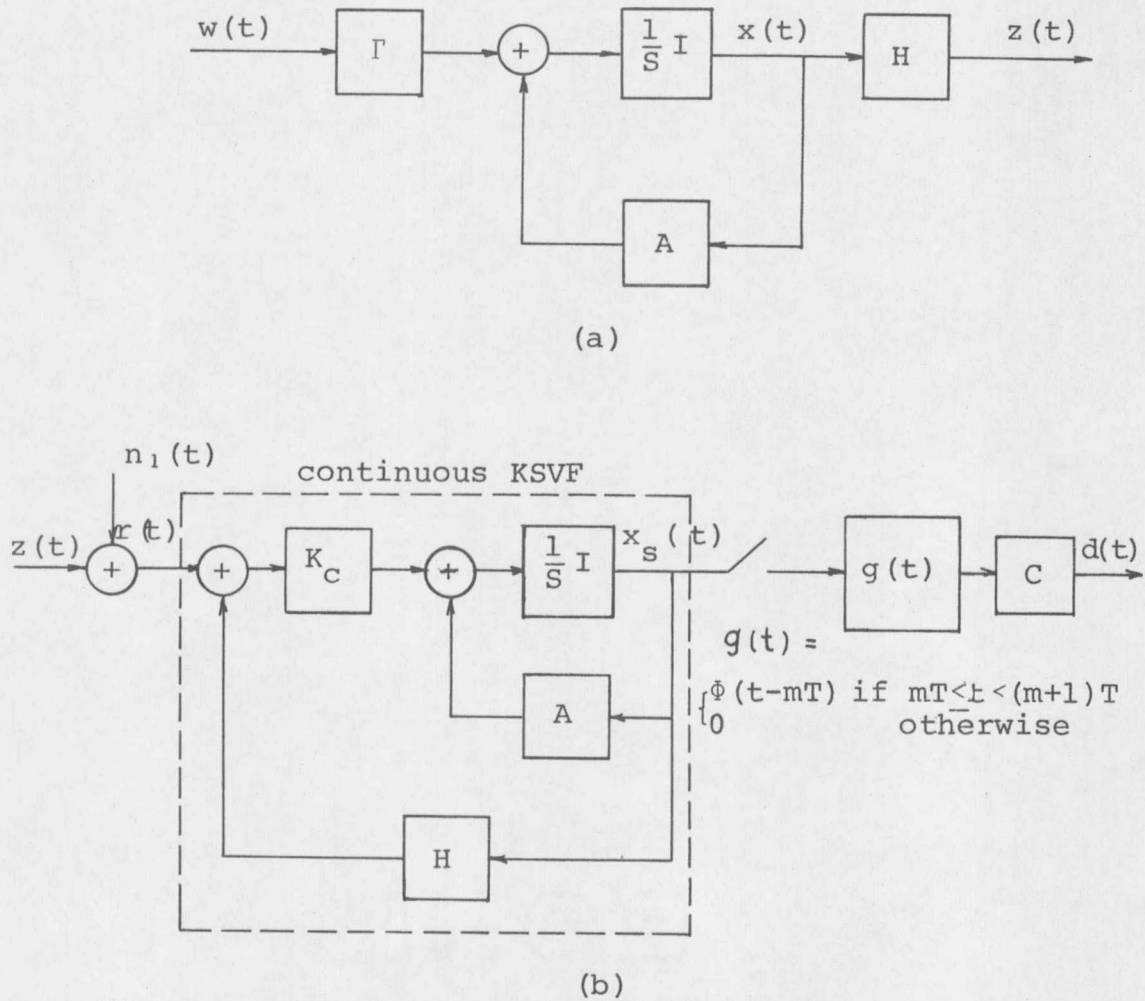


Figure 3.1 The jointly optimized pre- and postfilters with multiple sampled-data links and $n_2=0$: (a) the message model; (b) the system.

where $N(t)$ represents the coder inserted in the system before sampling. Notice that no matter whatever $N(t)$ one chose, $s(j)$ will always be a scalar and a function of the sampling instants only instead of a function of the continuous time. Hence, in this case, one loses the freedom to operate on $\hat{x}(kT/kT, r(\tau))$ in any way he wants; for example, to operate on $\hat{x}(kT/kT, r(\tau))$ as Eq. (3.16) does. In this sense, Petersen's pre- and postfilters structure is actually a suboptimum system (42).

3.3 The Optimum Prefilter with a Single Sampled-Data Link

In addition to Constraints 3.1 through 3.5, the following constraints are imposed in this section:

Constraint 3.8. There is only one sampled-data link available.

Constraint 3.9. $r(t)$ and $d(t)$ are scalars; i.e., H and C are row matrices.

Constraint 3.10. $n_2(k)$ can either be present or absent. If $n_2 \neq 0$, it might represent the quantization noise approximated by a discrete white noise or/and an independent channel additive noise. In the case in which the independent channel noise is present, a power constraint of

$$P_Y = \int_0^\infty \int_0^\infty f(\alpha) R_r(\alpha-\beta) f(\beta) d\alpha d\beta \quad (3.37)$$

is imposed on the prefilter. From here on, $f(\alpha)$ is used to denote $f_0(\alpha)$, the optimum prefilter, in order to simplify the notation.

The equation that specifies the optimal prefilter can be obtained from (E.43) or (E.50) depending on the characteristics of n_2 . The resulting equation with β in (E.43) or (E.50) replaced by t for easy presentation is (ignoring realizability constraint at first)

$$\int_{-\infty}^{\infty} R_{dr}(t+v) g(v) dv = \int_{-\infty}^{\infty} \left[\sum_{\ell=-\infty}^{\infty} \int_{-\infty}^{\infty} g(v-T) g(v) dv \right. \\ \left. R_r(\ell T+r-\alpha) + c_3 R_r(t-\alpha) \right] f(\alpha) d\alpha, \quad (3.38)$$

where $c_3 = T \cdot c_2 \cdot \int_0^\infty g^2(v) dv$ and c_2 is given by (E.48) if n_2 represents the quantization noise alone; $c_3 = \lambda T$, where λ is the Lagrange multiplier if a power constraint such as the one given by Eq. (3.37) is imposed on the prefilter; and $c_3 = 0$ if $n_2 = 0$. Rewriting (3.38) in such a form

$$\int_{-\infty}^{\infty} R_{dr}(t+v) g(v) dv = \int_{-\infty}^{\infty} f(\tau) \tilde{R}(t-\tau) d\tau, \quad (3.39)$$

where

$$\Xi(t-\tau) = \sum_{j=-\infty}^{\infty} \xi_j R_r(jT+t-\tau) \quad (3.40)$$

$$\xi_0 = \int_{-\infty}^{\infty} g^2(v) dv + c, \quad (3.41)$$

and

$$\xi_j = \int_{-\infty}^{\infty} g(v-jT) g(v) dv \quad \text{for } j \neq 0 \quad (3.42)$$

Now let $x(t)=[x_1(t), x_2(t), \dots, x_n(t)]'$, and assume that $g(t)$ is given a certain reconstruction filter. Then the equations specify the optimal prefilter for estimating the x_i 's for $i=1, 2, \dots, n$ with a single sampled-data link are

$$\int_{-\infty}^{\infty} g(v) R_{x_i r}(t+v) dv = \int_{-\infty}^{\infty} f_i(\tau) \Xi(t-\tau) d\tau, \quad (3.43)$$

for $i=1, 2, \dots, n$. In vector form,

$$\int_{-\infty}^{\infty} g(v) R_{xr}(t+v) dv = \int_{-\infty}^{\infty} F(\tau) \Xi(t-\tau) d\tau, \quad (3.44)$$

where $R_{xr}(\rho)=[R_{x_1 r}(\rho), R_{x_2 r}(\rho), \dots, R_{x_n r}(\rho)]'$, and $F(\tau)=[f_1(\tau), f_2(\tau), \dots, f_n(\tau)]'$. Notice that (3.44) is simply a vector representation of the disjoint equations of (3.43), it does not represent the equation for a multiple sampled-data links system, but merely enables one to

investigate various estimates at the same time (this is also evident from the fact that $g(t)$ in (3.44) is a scalar).

From here on, several lemmas and theorems are developed for solving the problem under consideration in this section. All the lemmas and theorems to be presented later have the constraints given by Constraints 3.1 through 3.5 and Constraints 3.8 through 3.10.

3.3.1 The structure of the optimal prefilter

Lemma 3.2 The optimal prefilter for estimating

$d(t)=C \cdot x(t)$ is $f(t)=C \cdot F(t)$, where $F(t)$ is specified by (3.44) and C is a row matrix.

Proof. Premultiplying by C on both sides of (3.44) to obtain

$$\int_{-\infty}^{\infty} C R_{xr}(t+v) g(v) dv = \int_{-\infty}^{\infty} [C F(\tau)] E(t-\tau) d\tau, \quad (3.45)$$

where

$$\begin{aligned} C R_{xr}(t+v) &= C \cdot E[x(t+v) r(0)] \\ &= E[C \cdot x(t+v) r(0)] \\ &= R_{dr}(t+v) \end{aligned} \quad (3.46)$$

Substituting (3.46) into (3.45) yields

$$\int_{-\infty}^{\infty} R_{dr}(t+v) g(v) dv = \int_{-\infty}^{\infty} [C F(\tau)] E(t-\tau) d\tau. \quad (3.47)$$

Comparing (3.47) with (3.39), we obtain

$$f(t) = C \cdot F(t), \quad (3.48)$$

where $F(t)$ is specified by (3.44)

We now introduce the innovations approach to the prefiltering problem. Instead of working with $r(t)$ directly, one first converts it to an innovation process $u(t)$ by a causal, and causally invertible filter f_u . One then finds the optimal prefilter f_y assuming that $u(t)$ is the observation process. As seen in Fig.3.2, the actual optimal prefilter, f , is the combination of f_u and f_y ; i.e.,

$$f(t) = f_u(t) \otimes f_y(t), \quad (3.49)$$

where $\otimes$ denotes convolution. It can be seen clearly from Fig. 3.2 that the insertion of the causally invertible f_u does not affect the optimality of the overall structure, since if f_u turns out to be undesirable, one can always include the inverse of f_u in f_y to cancel the effect of f_u . This argument is basically the same as the one used

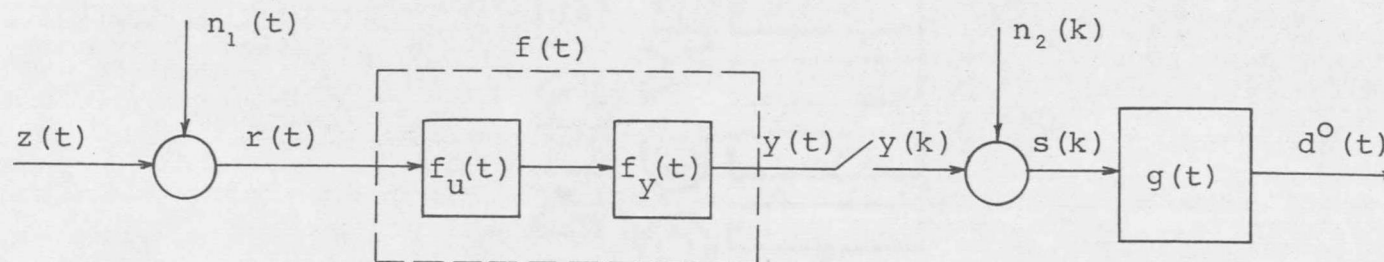


Figure 3.2 The innovations approach to the prefiltering problem.

in deriving the Wiener filter using the innovations approach. For a detail discussion of the above argument, see Section 1.4.3 of this thesis (58). The point is that the effect of the innovation filter put before sampling can be cancelled by the remainder of the optimal prefilter. If instead, one does not have the freedom in choosing the prefilter but has the freedom in choosing the postfilter, then the innovation filter f_u added before sampling will not be invertible by the postfilter. In this case, the addition of a pre-sampling filter will alter the performance of the optimum system.

When the observation process is $u(t)$ instead of $r(t)$, (3.44) becomes

$$\int_{-\infty}^{\infty} g(v) R_{xu}(t+v) dv = \int_{-\infty}^{\infty} F_Y(\tau) \sum_{j=-\infty}^{\infty} \xi_j R_u(jT+t-\tau) d\tau. \quad (3.50)$$

Since $u(t)$ is an innovation process,

$$\begin{aligned} R_u(t-\tau) &= E[u(t) u(\tau)] \\ &= V_u \delta(t-\tau), \end{aligned} \quad (3.51)$$

where V_u is the variance of u , which is a scalar, and is determined by $r(t)$ (see Appendix A, Equations (A.32) and (A.48), and the discussion about colored noise).

Substituting (3.51) into (3.50) yields

$$\int_{-\infty}^{\infty} g(v) R_{xu}(t+v) dv = \sum_{j=-\infty}^{\infty} V_u \xi_j F_Y(t+jT) \quad (3.52)$$

With the realizability constraint, $g(t)$ and $F_Y(t)$ are zero for $t < 0$. Thus, Eq. (3.52) becomes

$$\int_0^{\infty} g(v) R_{xu}(t+v) dv = \sum_{j=-\lceil t/T \rceil}^{\infty} V_u \xi_j F_Y(t+jT) \quad (3.53)$$

for $t, v \geq 0$, and (3.41) and (3.42) become

$$\xi_0 = \int_0^{\infty} g^2(v) dv + c_3 \quad (3.54)$$

$$\xi_j = \int_{0 \wedge jT}^{\infty} g(v-jT) g(v) dv \quad \text{for } j \neq 0 \quad (3.55)$$

where the symbol $0 \wedge jT$ is given by (3.20). In this section, the innovation process $u(t)$ of the observation process $r(t)$, and the variance of the innovation process, V_u , are assumed given. For a discussion of the innovation process, see Appendix A (23).

Theorem 3.1 The $F_Y(t)$ specified by (3.50) satisfies the differential equation

$$d F_Y(t)/dt = A F_Y(t) \quad (3.56)$$

for $t > 0$, and $t \neq kT$ for $k=0,1,\dots$

Proof. First of all, it is seen from (3.53) that the summation on the right-hand side of (3.53) changes its lower bound every time t passes through a sampling instant.

Thus, it is obvious that unless the ξ_j 's possess certain characteristic, $F_y(t)$ will in general be a piece-wise continuous function of t . Next, we have

$$\begin{aligned} \frac{d R_{xu}(t)}{dt} &= E \left[\frac{\partial x(t+v)}{\partial t} u(v) \right] \\ &= E \{ [A x(t+v) + \Gamma w(t+v)] u(v) \}. \end{aligned} \quad (3.57)$$

Since $u(v)$ is independent of $w(t+v)$ for $t > 0$ (see Appendix A),

$$E[w(t+v) u(v)] = 0 \quad \text{for } t > 0. \quad (3.58)$$

Thus,

$$\frac{d R_{xu}(t)}{dt} = A R_{xu}(t) \quad \text{for } t > 0; \quad (3.59)$$

i.e.,

$$R_{xu}(t) = \Phi(t-\tau) R_{xu}(\tau) \quad \text{for } t, \tau \geq 0, \quad (3.60)$$

where $\Phi(t)$ is the transition matrix of A . Since $t, v \geq 0$,

$$\begin{aligned} R_{xu}(t+v) &= \Phi(t+v-t) R_{xu}(t) \\ &= \Phi(v) R_{xu}(t). \end{aligned} \quad (3.61)$$

Substituting (3.61) into (3.53) yields, for $t \geq 0$,

$$\left[\int_0^\infty g(v) \Phi(v) dv \right] R_{xu}(t) = V_u \cdot \sum_{j=-\lceil t/T \rceil}^{\infty} \xi_j F_Y(t+jT). \quad (3.62)$$

Let

$$L = \int_0^\infty g(v) \Phi(v) dv, \quad (3.63)$$

where L is a $n \times n$ matrix, and is in general nonsingular (L is nonsingular if $\Phi(v)$ is nonsingular). Premultiplying by L^{-1} on both sides of (3.62) to obtain

$$R_{xu}(t) = V_u L^{-1} \sum_{j=-\lceil t/T \rceil}^{\infty} \xi_j F_Y(t+jT) \quad \text{for } t \geq 0. \quad (3.64)$$

Differentiating both sides of (3.64) with respect to t for $kT < t < (k+1)T$, where $k=0,1,\dots$, we have

$$dR_{xu}(t)/dt = V_u L^{-1} \sum_{j=-\lceil t/T \rceil}^{\infty} \xi_j \partial F_Y(t+jT)/\partial t. \quad (3.65)$$

Substituting (3.59) into (3.65) to obtain, for $kT < t < (k+1)T$, ($k=0,1,\dots$),

$$A R_{xu}(t) = V_u L^{-1} \sum_{j=-\lceil t/T \rceil}^{\infty} \xi_j \partial F_Y(t+jT)/\partial t. \quad (3.66)$$

Substituting (3.64) into (3.66) yields

$$V_u L^{-1} \left[\sum_{j=-\lceil t/T \rceil}^{\infty} \xi_j \left\{ \frac{\partial F_Y(t+jT)}{\partial t} - A F_Y(t+jT) \right\} \right] = 0. \quad (3.67)$$

Since (3.67) has to be satisfied for arbitrary $t > 0$, except at sampling instants, and since at least one $\xi_j \neq 0$ in general (ξ_0 for example); it is both sufficient and necessary that

$$\partial F_Y(t+jT)/\partial t = A F_Y(t+jT) \quad t+jT > 0, \quad t \neq kT \text{ for } k=0,1,\dots, \quad (3.68)$$

i.e.,

$$\frac{dF_Y(t)}{dt} = A F_Y(t), \quad \text{for } t > 0, \text{ and } t \neq kT \text{ for } k=0,1,\dots \quad (3.69)$$

Lemma 3.3 $L \Phi(t) = \Phi(t) L$

Proof. Since $L = \int_0^\infty g(v) \Phi(v) dv$, where $g(v)$ is a scalar, Lemma 3.3 follows directly from Lemma 3.1.

Theorem 3.2 The structure of the optimal realizable prefilter with any given realizable postfilter can be represented by the cascade of a continuous KSVF followed by a coder, where the impulse response of the coder is an impulse train in general.

In Chapter 4, we will prove a theorem which says that in the case in which $n_2(k)$ is absent, the coder of the jointly optimized prefilter is simply a gain matrix;

i.e., the impulse response of the coder in this case is simply an impulse function at the origin.

Proof. From Theorem 3.1, $F_Y(t)$ for $t \neq kT$ satisfies the equation

$$dF_Y(t)/dt = A F_Y(t),$$

whose solution is

$$F_Y(t) = \Phi(t-kT) F_Y(kT) \quad (3.70)$$

for $kT \leq t < (k+1)T$. Thus, the only problem left is to find $F_Y(kT)$ for $k=0,1,\dots$, etc. From (3.64), with $t=kT$ for $k=0,1,\dots$, we have

$$L R_{xu}(kT) = V_u \sum_{j=-k}^{\infty} \xi_j F_Y(kT+jT). \quad (3.71)$$

Through the use of (3.60) and Lemma 3.3, Eq. (3.71) can be written as

$$k=0, V_u^{-1} L R_{xu}(0) = \xi_0 F_Y(0) + \xi_1 F_Y(T) + \dots + \xi_i F_Y(iT) + \dots \quad (3.72-1)$$

$$k=1, V_u^{-1} \Phi(T) L R_{xu}(0) = \xi_{-1} F_Y(0) + \xi_0 F_Y(T) + \dots + \xi_{i-1} F_Y(iT) + \dots \quad (3.72-2)$$

$$\begin{aligned}
k=j, V_u^{-1} \Phi(jT) L R_{xu}(0) &= \xi_{-j} F_Y(0) + \xi_{-j+1} F_Y(T) + \dots \\
&+ \xi_{i-j} F_Y(iT) + \dots
\end{aligned}
\tag{372-j+1}$$

For a stable system, both ξ_k and $F_Y(kT)$ vanish as k increases. Thus, a solution of the $F_Y(kT)$'s in terms of a linear combination of the $[V_u^{-1} \times \Phi(jT) \cdot L \cdot R_{xu}(0)]$'s for $j=0,1,\dots$, etc., can be found from (3.72), for example, by reduction method. The solution can be written as

$$F_Y(kT) = V_u^{-1} \left[\sum_{i=0}^{\infty} a_{ki} \Phi(iT) \right] \cdot L \cdot R_{xu}(0), \tag{3.73}$$

where the a_{ki} 's are scalars, and are functions of the ξ_j 's for $j=0,1,\dots$, etc. From (3.73) and (3.70), it follows that for $mT \leq t < (m+1)T$,

$$\begin{aligned}
F_Y(t) &= \Phi(t-mT) F_Y(mT) \\
&= \left[\sum_{i=0}^{\infty} a_{mi} \Phi(iT) \right] L V_u^{-1} \Phi(t-mT) R_{xu}(0),
\end{aligned}
\tag{3.74}$$

where the second equality follows from Lemma 3.1 and Lemma 3.3. Since $\Phi(t) R_{xu}(0) = R_{xu}(t)$ for $t \geq 0$ from (3.60),

$$F_Y(t) = \left[\sum_{i=0}^{\infty} a_{mi} \Phi(iT) \right] L V_u^{-1} R_{xu}(t-mT) \tag{3.75}$$

for $mT \leq t < (m+1)T$. The optimal prefilter to operate on $u(t)$, the innovation process of $r(t)$, is

$$f_Y(t) = C F_Y(t) \quad (3.76)$$

from Lemma 3.2. Substituting (3.75) into (3.76) yields

$$f_Y(t) = C \left[\sum_{i=0}^{\infty} a_{mi} \Phi(iT) \right] L [V_u^{-1} R_{xu}(t-mT)]. \quad (3.77)$$

Let

$$\chi_m = C \left[\sum_{i=0}^{\infty} a_{mi} \Phi(iT) \right] L, \quad \text{for } m=0,1,\dots \quad (3.78)$$

then (3.77) becomes, for $mT \leq t < (m+1)T$

$$f_Y(t) = \chi_m [V_u^{-1} R_{xu}(t-mT)]. \quad (3.79)$$

The general form of $f_Y(t)$ for $t \geq 0$ is

$$f_Y(t) = \sum_{k=0}^{\infty} \chi_k V_u^{-1} R_{xu}(t-kT) [u_s(t-kT) - u_s(t-(k+1)T)] \quad (3.80)$$

from (3.79), where $u_s(t)$ is the unite step function and

χ_k is given by (3.78). Since $[R_{xu}(t-kT) u_s(t-(k+1)T)]$ is nonzero for $t \geq (k+1)T$ only,

$$R_{xu}(t-kT) u_s(t-(k+1)T) = \Phi(T) R_{xu}(t-(k+1)T) u_s(t-(k+1)T) \quad (3.81)$$

from (3.60). By the use of (3.81), (3.80) can be re-written in the form

$$f_Y(t) = \sum_{k=0}^{\infty} \chi_k V_u^{-1} R_{xu}(t-kT) u_s(t-kT) - \chi_k \Phi(T) \times \\ \times V_u^{-1} R_{xu}(t-(k+1)T) u_s(t-(k+1)T) \}$$

$$\begin{aligned}
&= \sum_{k=0}^{\infty} N_k V_u^{-1} R_{xu}(t-kT) u_s(t-kT) \\
&= \left[\sum_{k=0}^{\infty} N_k \delta(t-kT) \right] \otimes [V_u^{-1} R_{xu}(t) u_s(t)], \quad (3.82)
\end{aligned}$$

where $N_0 = \chi_0$ and $N_i = \chi_i - \chi_{i-1} \Phi(T)$ for $i \geq 1$, and $\otimes$ denotes convolution. From Fig. 3.2, the actual optimal prefilter is

$$f(t) = f_y(t) \otimes f_u(t) \quad (3.83)$$

where $f_u(t)$ represents the innovation filter. Substituting (3.82) into (3.83) yields

$$\begin{aligned}
f(t) &= \left[\sum_{k=0}^{\infty} N_k \delta(t-kT) \right] \otimes [V_u^{-1} R_{xu}(t) u_s(t)] \otimes f_u(t) \\
&= \left[\sum_{k=0}^{\infty} N_k \delta(t-kT) \right] \otimes [V_u^{-1} R_{xu}(t) u_s(t) \otimes f_u(t)]. \quad (3.84)
\end{aligned}$$

Recognizing that the continuous KSVF can be represented by the combination of an innovation filter followed by a filter characterized by its impulse response (23)

$$f_w(t) = V_u^{-1} R_{xu}(t) u_s(t); \quad (3.85)$$

i.e.,

$$\begin{aligned}
F_{ks}(t) &= \text{the continuous KSVF} \\
&= V_u^{-1} R_{xu}(t) u_s(t) \otimes f_u(t). \quad (3.86)
\end{aligned}$$

Thus, (3.84) becomes

$$\begin{aligned} f(t) &= \left[\sum_{k=0}^{\infty} N_k \delta(t-kT) \right] \otimes F_{ks}(t) \\ &= \sum_{k=0}^{\infty} N_k F_{ks}(t-kT) u_s(tk-T) \end{aligned} \quad (3.87)$$

Hence, the optimal prefilter can be represented by the combination of a continuous KSVF followed by a coder where the impulse response of the coder is an impulse train; i.e., let $N(t)$ denotes the impulse response of the coder, then

$$N(t) = \sum_{i=0}^{\infty} N_i \delta(t-iT). \quad (3.88)$$

Thus, the proof is completed. The block diagram of the structure of the optimal prefilter is shown in Fig. 3.3.

The coder contained in the optimal prefilter corresponds to the periodic factor contained in the transfer function of the optimal prefilter obtained from a frequency-domain analysis. As we shall see in Chapter 4, the optimal coder of the jointly optimized pre- and postfilters becomes a constant row matrix when n_2 in Fig. 1.1 is absent.

Several examples with suboptimal postfilters are worked out next. Notice that in this section, nothing has

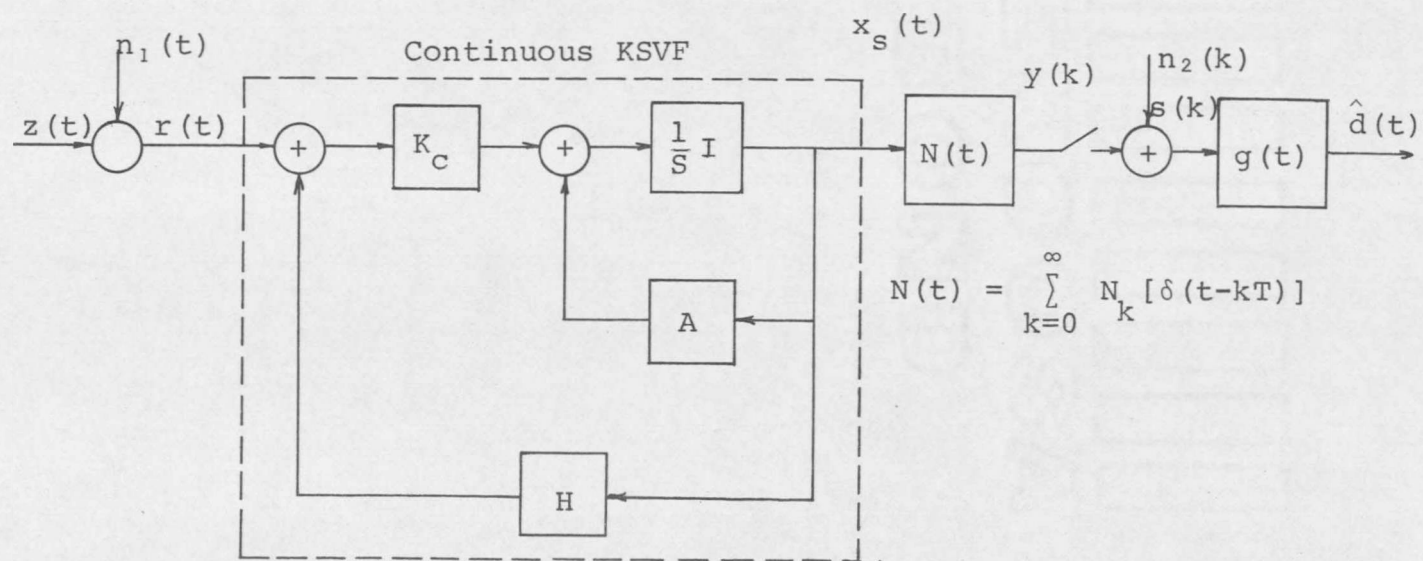


Figure 3.3 The optimum prefilter with a single sampled-data link and a given postfilter.

been said about what kind of postfilter is being used; i.e., the postfilter can be of any reconstruction filter.

3.3.2 Examples

Example 3.1

The impulse response of the postfilter is given by

$$g(t) = \begin{cases} g_e(t) & \text{if } 0 \leq t < T \\ 0 & \text{otherwise,} \end{cases}$$

where $g_e(t)$ can be of any continuous function of t . In the case in which $g(t)$ represents a zero-order hold, $g_e(t)$ becomes a constant. From (3.54) and (3.55), assuming that $n_2(k)=0$; we have

$$\xi_k = \begin{cases} \int_0^T g_e^2(t) dt & \text{if } k=0 \\ 0 & \text{otherwise.} \end{cases}$$

Then from (3.64), it follows that

$$L R_{xu}(t) = V_u \xi_0 F_Y(t),$$

where L depends on $g_e(t)$ and the input signal, and is given by (3.63).

Thus,

$$F_Y(t) = (V_u \xi_0)^{-1} L R_{xu}(t) \cdot u_s(t)$$

From (3.76) and (3.83), the optimal prefilter for estimating $d(t)=C x(t)$ is given by

$$\begin{aligned} f(t) &= \xi_0^{-1} C L [V_u^{-1} R_{xu}(t) \cdot u_s(t) \otimes f_u(t)] \\ &= \xi_0^{-1} C L F_{ks}(t) \end{aligned} \quad (3.89)$$

from (3.86). Thus,

$$N_0 = \xi_0^{-1} C L \quad (3.90)$$

and

$$N_j = 0 \quad \text{for } j \neq 0 \quad (3.91)$$

by comparing (3.89) with (3.87). It is seen from this example that if $g(t)$ is constrained to be of finite time duration, the optimal prefilter can usually be found without too much trouble.

Example 3.2

The impulse response of the postfilter is given by

$$g(t) = \exp(-t) \quad \text{for } 0 \leq t. \quad (3.92)$$

Again, assume $n_2=0$; then from (3.54) and (3.55), we have

$$\xi_j = (1/2) \exp(-|j|T) \quad (3.93)$$

for $j \geq 0$. Substituting (3.93) into (3.72) yields,

$$V_u^{-1} L R_{xu}(0) = (1/2) [F_Y(0) + e^{-T} F_Y(T) + \dots + e^{-iT} F_Y(iT) + \dots] \quad (3.94-1)$$

$$V_u^{-1} L \Phi(T) R_{xu}(0) = (1/2) [e^{-T} F_Y(0) + F_Y(T) + e^{-T} F_Y(2T) + \dots + e^{-(i-1)T} F_Y(iT) + \dots] \quad (3.94-2)$$

$$V_u^{-1} L \Phi(kT) R_{xu}(0) = (1/2) [e^{-kT} F_Y(0) + e^{-(k-1)T} F_Y(T) + \dots + F_Y(kT) + e^{-T} F_Y((k+1)T) + \dots] \quad (3.94-k+1)$$

Since e^{-jT} vanishes when $j \rightarrow \infty$, subtracting $[e^{-T} \cdot (3.94-2)]$ from (3.94-1) yields

$$L V_u^{-1} [I - e^{-T} \Phi(T)] R_{xu}(0) = (1/2) (1 - e^{-2T}) F_Y(0);$$

i.e.,

$$F_Y(0) = 2 (1 - e^{-2T})^{-1} [I - e^{-T} \Phi(T)] L V_u^{-1} R_{xu}(0).$$

Similarly, by the method of mathematical induction, it is found that

$$F_Y(kT) = \Phi((k-1)T) [\Phi(T) - I \cdot e^{-T}] F_Y(0) \quad (3.95)$$

From (3.70), (3.60), and (3.95),

$$F_Y(t) = \begin{cases} 2(1-e^{-2T}) L [I-e^{-T}\Phi(T)] V_u^{-1} R_{xu}(t) & \text{for } 0 \leq t \leq T \\ [\Phi(T)-Ie^{-T}] 2(1-e^{-2T}) \cdot L \cdot [I-e^{-T}\Phi(T)] V_u^{-1} R_{xu}(t-T) & \text{for } t > T. \end{cases} \quad (3.96)$$

From (3.76), (3.83), (3.86), and (3.96), it follows that the impulse response of the optimal prefilter is given by

$$f(t) = N_0 F_{ks}(t) u_s(t) + N_1 F_{ks}(t-T) u_s(t-T) \quad (3.97)$$

where

$$N_0 = 2(1-e^{-2T}) C L [I - e^{-T} \Phi(T)]$$

and

$$N_1 = e^{-T} N_0.$$

By comparing (3.97) with (3.87), it is found that $N_j = 0$ for $j \geq 2$.

Chapter 4

JOINT OPTIMIZATION OF THE PRE- AND POSTFILTERS

4.1 Problem Statement and Preliminary Remarks

In this chapter, the joint optimization of the pre- and postfilters of the system shown in Fig. 1.1 is considered. System equations and prior statistics are given by Eqs. (3.1) through (3.11). The constraints of the system being considered are given by Constraints 3.1 through 3.5 and Constraints 3.8 through 3.10. The central result of this chapter is the derivation of the structure of the jointly optimized pre- and postfilters when n_2 in Fig. 1.1 is absent.

We have mentioned in Section 3.3.1 that the coder contained in the optimum prefilter corresponds to the periodic factor contained in the transfer function* of the optimum prefilter obtained from the frequency-domain approach. When $n_2=0$ (see Fig. 1.1_c), this periodic factor can be shifted to the receiving end to be included in the postfilter without influencing the performance of the

A periodic transfer function $H^(S)$ has the property that $H^*(S) = H^*(S + jn\frac{2\pi}{T})$ for $n=0,1,2,\dots$, where T is the period.

system, as pointed out by Robbins (49). Thus, the impulse response of the optimal prefilter becomes a continuous function instead of a piece-wise continuous function in this case. That is, the coder contained in the optimal prefilter becomes a constant matrix (a gain matrix). A proof of this point from the state-space approach is presented in Section 4.3. The proof, together with the development leading to the proof serves several purposes:

1. Degeneration of the joint optimization problem.
2. To illustrate the consistency between the state-space approach and the frequency-domain analysis.
3. To derive the structure of the jointly optimized pre- and postfilters.

4.2 Degeneration of the Joint Optimization of the Pre- and Postfilters

The basic background of the degeneration of the joint optimization problem is provided by the following theorem.

Theorem 4.1 Let

$x_s(t) = \hat{x}(t/-\infty, t, r(\tau))$ = the linear least-square estimate of $x(t)$ based on $r(\tau)$ for $-\infty < \tau \leq t$, where x and r are given by (3.1) through (3.3).

$x_s(t)$ is the best zero-lag estimate of $x(t)$ in the absence of sampling. That is, $x_s(t)$ is the output of the continuous KSVF. Also let $v(k) \triangleq$ the sum of a linear combination of the $x_s(j)$'s for $j \leq k$ and an independent noise $n_p(k)$; i.e.,

$$v(k) = \sum_{j=-\infty}^k h_{k-j} x_s(j) + n_p(k). \quad (4.1)$$

Then the theorem says that

$$\hat{x}_s(t/-\infty, m, v(k)) = \hat{x}(t/-\infty, m, v(k)),$$

where the quantity $\hat{\xi}(t/l_i, l_f, n(k))$ has been defined in Chapter 2 and $m = \lfloor t/T \rfloor$,

Furthermore,

$$\begin{aligned} E\{[x(t) - \hat{x}(t/-\infty, m, v(k))]^2\} &= E\{[x(t) - x_s(t)]^2\} + \\ &E\{[x_s(t) - \hat{x}_s(t/-\infty, m, v(k))]^2\}. \end{aligned}$$

From here on, $\hat{\xi}$ will be used to denote $\hat{\xi}(t/-\infty, m, v(k))$, and ξ will be used to denote $\xi(t)$, where ξ can be of any related signals (such as x and x_s , etc.), in order to simplify notations.

Proof. From the orthogonality principle, it follows that for $k \leq m$,

$$E[(x_s - \hat{x}_s) v'(k)] = 0. \quad (4.2)$$

We want to prove that for $k \leq m$,

$$E[(x - \hat{x}_s) v'(k)] = 0.$$

Since $x - \hat{x}_s = (x - x_s) + (x_s - \hat{x}_s)$,

$$E[(x - \hat{x}_s) v'(k)] = E[(x - x_s) v'(k)] + E[(x_s - \hat{x}_s) v'(k)]. \quad (4.3)$$

The second term on the right-hand side of (4.3) is zero from (4.2). Substituting (4.1) into the first term on the right-hand side of (4.3) yields for $k \leq m$

$$E[(x - x_s) v'(k)] = \sum_{j=-\infty}^k E[(x - x_s) x_s'(j)] h'_{k-j} + E[(x - x_s) n_p'(k)]. \quad (4.4)$$

The first term on the right-hand side of (4.4) is zero since $x_s(j)$ is a linear combination of $r(\tau)$ for $\tau \leq jT$. The second term is also zero since $n_p(k)$ is assumed to be an independent noise. Thus, the first term on the right-hand side of (4.3) is also zero, and we have, for $k \leq m$

$$E[(x - \hat{x}_s) v(k)] = 0; \quad (4.5)$$

i.e.,

$$\hat{x}(t/-\infty, m, v(k)) = \hat{x}_s(t/-\infty, m, v(k)). \quad (4.6)$$

The error covariance of $[x(t) - \hat{x}(t/-\infty, m, v(k))]$ can be written as

$$\begin{aligned} E[(x - \hat{x})^2] &= E[(x - \hat{x}_s)^2] \\ &= E\{[(x - x_s) + (x_s - \hat{x}_s)]^2\} \\ &= E[(x - x_s)^2] + E[(x_s - \hat{x}_s)^2] + E[(x - x_s)(x_s - \hat{x}_s)'] \\ &\quad + E[(x_s - \hat{x}_s)(x - x_s)'], \end{aligned}$$

where $E[\xi^2] = E[\xi\xi']$. Since x_s is a linear combination of $v(k) = [\sum_{j=-\infty}^k h_{k-j} x_s(j) + n_p(k)]$,

$$E[(x - x_s)(x_s - \hat{x}_s)'] = E[(x_s - \hat{x}_s)(x - x_s)'] = 0 \quad (4.7)$$

Thus,

$$E[(x - \hat{x})^2] = E[(x - x_s)^2] + E[(x_s - \hat{x}_s)^2]. \quad (4.8)$$

Eq. (4.8) completes the proof.

Since $d(t) = C x(t)$, the error covariance of estimating $d(t)$ based on the observation of $s(k)$ (see Fig. 1.1) for $-\infty < k \leq m$ is

$$\begin{aligned} E\{[d(t) - \hat{d}(t/-\infty, m, s(k))]^2\} &= C E\{[x(t) - x_s(t)]^2\} C' + \\ &\quad C E\{[x_s(t) - \hat{x}_s(t/-\infty, m, s(k))]^2\} C'. \quad (4.9) \end{aligned}$$

It is seen from (4.9) that the overall error consists of two terms. The first term on the right-hand side of (4.9) represents the error due to the noise $n_1(t)$ (see Fig. 1.1), and is independent of the coder N_k 's. The second term corresponds to the error due to sampling and $n_2(k)$, and is dependent on the coder.

Lemma 4.1 The sampled value of the output of the optimum prefilter, $y(j)$, can be represented by a linear combination of the $x_s(k)$'s for $k \leq j$.

Proof. From (3.87), it follows that

$$\begin{aligned} y(t) &= \int_0^{\infty} f(v) r(t-v) dv \\ &= \int_0^{\infty} \sum_{k=0}^{\infty} N_k F_{ks}(v-kT) u_s(v-kT) r(t-v) dv \\ &= \sum_{k=0}^{\infty} N_k \int_{kT}^{\infty} F_{ks}(v-kT) r(t-v) dv. \end{aligned} \quad (4.10)$$

Let $\rho = v - kT$; then $v = \rho + kT$ and $dv = d\rho$, and (4.10) becomes

$$\begin{aligned} y(t) &= \sum_{k=0}^{\infty} N_k \int_0^{\infty} F_{ks}(\rho) r[(t-kT)-\rho] d\rho \\ &= \sum_{k=0}^{\infty} N_k x_s(t-kT) \end{aligned} \quad (4.11)$$

Let $t=jT$; then (4.11) becomes

$$y(jT) = \sum_{k=0}^{\infty} N_k x_s(jT-kT) \quad (4.12)$$

Let $i=j-k$; (4.12) can be written as

$$y(j) = \sum_{i=-\infty}^j N_{j-i} x_s(i). \quad (4.13)$$

Since n_2 in Fig. 1.1 is assumed to be uncorrelated with the input signal, we have the following conclusion drawn from Theorem 4.1 and Lemma 4.1:

Theorem 4.2 The joint optimization of the pre- and post-filters shown in Fig. 1.1 can be degenerated to the problem of finding the optimum postfilter $G(t)$ and the optimum coder N_k for estimating $x_s(t)$ from the $s(j)$'s for $j \leq \langle t/T \rangle$, where

$$s(j) = y(j) + n_2(j), \quad (4.14)$$

$y(j)$ is given by (4.13), and $x_s(t)$ is specified by the state equation (from (A.33) or (A.49) depending on the characteristic of $n_1(t)$).

$$\dot{x}_s(t) = A x_s(t) + K_c u(t), \quad (4.15)$$

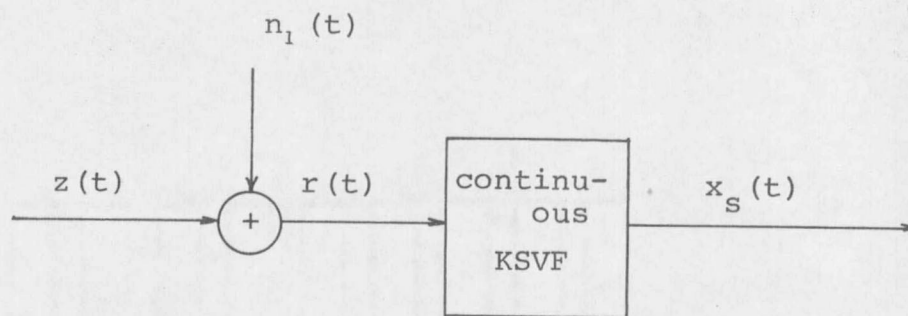
where K_c is the continuous KSVF gain given by (A.35) or (A.51), and $u(t)$ is the innovation process of $r(t)$ given

by (A.34) or (A.50) depending on the characteristic of $n_1(t)$, the primary observation noise.

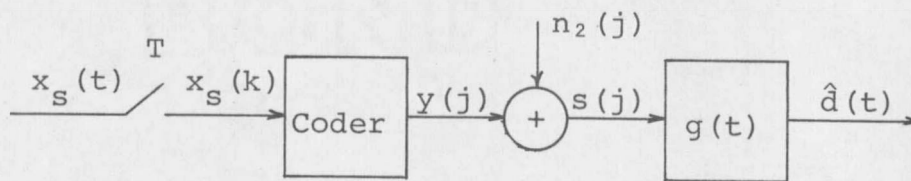
It is seen from (4.13) that the coder contained in the optimum prefilter can be considered as a discrete filter operating on the $x_s(j)$'s. That is, one can assume that there are n samplers which sample $x_s(t)$ to give the $x_s(j)$'s, where n equals the dimension of the state vector. Then the $x_s(j)$'s are operated on by the coder to yield $y(k)$, which is then quantized and transmitted through the channel. The block diagram of this degenerated system is shown in Fig. 4.1.

Since both the first- and second-order statistics of $x_s(t)$ are known, the system shown in Fig. 4.1 is similar to the system considered by Hayes (20; also see Fig. 1.2_h) except that the input to the coder shown in Fig. 4.1 is a vector instead of a scalar.

If the output of the coder can be characterized by state equations, one can use the method developed in Chapter 2 to solve for the optimal postfilter. The case in which $n_2(k)=0$ is investigated in the next section.



(a)



$$y(j) = \sum_{k=-\infty}^j N_{j-k} x_s(k)$$

(b)

Figure 4.1 Degeneration of the joint optimization problem: (a) generation of $x_s(t)$; (b) the degenerated system.

4.3 Jointly Optimized Pre- and Postfilters with $n_2(k)=0$

The structure of the optimum filters will be given by Theorem 4.3. Several Lemmas are given before proving Theorem 4.3.

Lemma 4.2

$$R_{x_s}(t) = E[x_s(t+\tau) x_s'(\tau)] = \begin{cases} \Phi(t) R_{x_s}(0) & \text{if } t \geq 0 \\ R_{x_s}(0) \Phi'(-t) & \text{if } t < 0 \end{cases}$$

Proof. From (4.15), it follows that for $t \geq 0$

$$x_s(t+\tau) = \Phi(t) x_s(\tau) + \int_{\tau}^t \Phi(t-\mu) K_C u(\mu) d\mu. \quad (4.16)$$

Thus, for $t \geq 0$

$$\begin{aligned} E[x_s(t+\tau) x_s'(\tau)] &= \Phi(t) E[x_s(\tau) x_s'(\tau)] \\ &\quad + \int_{\tau}^t \Phi(t-\mu) K_C E[u(\mu) x_s'(\tau)] d\mu. \end{aligned} \quad (4.17)$$

From (A.9) (where $T_i = -\infty$ and $T_f = t$) and the fact that $u(t)$ is a white noise, it follows that for $\mu > \tau$

$$E[u(\mu) x_s'(\tau)] = 0. \quad (4.18)$$

Substituting (4.18) into (4.17) to obtain

$$R_{x_s}(t) = \Phi(t) R_{x_s}(0) \quad \text{for } t \geq 0 \quad (4.19)$$

For $t < 0$,

$$\begin{aligned}
 R_{x_s}(t) &= E[x_s(t+\tau) x_s'(\tau)] \\
 &= \{E[x_s(\tau) x_s'(t+\tau)]\}, \\
 &= \{\Phi(-t) R_{x_s}(0)\},
 \end{aligned} \tag{4.20}$$

from (4.19). Thus,

$$R_{x_s}(t) = R_{x_s}(0) \Phi'(-t). \quad \text{for } t < 0 \tag{4.21}$$

The proof is completed by combining (4.19) and (4.21).

Lemma 4.3 $\xi_j = \xi_{-j}$, where ξ_j is given by (3.54) and (3.55).

Proof. Since $\xi_j = \int_{0 \wedge jT}^{\infty} g(t) g(t-jT) \cdot dt$ for $j \neq 0$, the result follows by letting $\mu = t - jT$.

Before we proceed with Theorem 4.3, let us assume that $n_2(k)$ might not be zero, and investigate what in general will happen if one insists on using a prefilter whose impulse response is continuous in time. If the $F_y(t)$ specified by (3.53) is continuous in t , then

$$F_y(kT) = \Phi(kT) F_y(0) \quad \text{for } k \geq 0 \tag{4.22}$$

from Theorem 3.1. Substituting (4.22) into (3.72) to obtain

$$V_u^{-1} L R_{xu}(0) = [\xi_0 I + \xi_1 \Phi(T) + \dots + \xi_i \Phi(iT) + \dots] F_Y(0), \quad (4.23-1)$$

$$V_u^{-1} \Phi(T) L R_{xu}(0) = [\xi_{-1} I + \xi_0 \Phi(T) + \dots + \xi_{i-1} \Phi(iT) + \dots] F_Y(0), \quad (4.23-2)$$

$$V_u^{-1} \Phi(jT) L R_{xu}(0) = [\xi_{-j} I + \xi_{-j+1} \Phi(T) + \dots + \xi_{i-j} \Phi(iT) + \dots] F_Y(0) \quad (4.23-j+1)$$

Premultiplying both sides of (4.23-2) by $\Phi(-T)$ to obtain

$$\begin{aligned} V_u^{-1} L R_{xu}(0) &= \xi_{-1} \Phi(-T) F_Y(0) + [\xi_0 I + \xi_1 \Phi(T) + \dots] F_Y(0) \\ &= \xi_{-1} \Phi(-T) F_Y(0) + V_u^{-1} L R_{xu}(0), \end{aligned} \quad (4.24)$$

where the first equality follows from the application of Lemma 3.1, and the second equality follows from the substitution of (4.23-1). After cancelling the common terms on both sides of (4.24), we have

$$\xi_{-1} \Phi(-T) F_Y(0) = 0 \quad (4.25)$$

Since $\Phi(-T)$ and $F_Y(0)$ are nonzero in general, (4.25) implies that

$$\xi_{-1} = 0.$$

Similarly, one can obtain the equation

$$\xi_j = 0 \quad \text{for all } j < 0.$$

Since $\xi_j = \xi_{-j}$ from Lemma 4.3, we have, for any $j \neq 0$

$$\xi_j = 0. \quad (4.26)$$

In other words, the constraint that $\xi_j = 0$ for any $j \neq 0$ has to be imposed on the postfilter no matter whether the postfilter is optimum or not if one insists that the impulse response of the optimum prefilter be continuous. In the case in which a jointly optimized solution is desired, this constraint has to be taken into consideration in deriving necessary conditions for the jointly optimized pre- and postfilters.

Theorem 4.3 With Constraints 3.1 through 3.5 and Constraints 3.8 through 3.10, with $n_2(k) = 0$, the optimal prefilter of the jointly optimized pre- and postfilters can be represented by the combination of a continuous KSVF followed by a constant coder N , where

$$N = \xi_0^{-1} C L, \quad (4.27)$$

$$\xi_0 = \int_0^\infty g^2(t) dt, \quad (4.28)$$

$$L = \int_0^{\infty} g(v) \Phi(v) dv, \quad (4.29)$$

and C is known (see (3.6)). The $g(t)$ that appears in (4.28) and (4.29) represents the impulse response of the optimum postfilter for estimating $d(t)=C \cdot x(t)$, and is given by

$$g(t) = C G(t), \quad (4.30)$$

where $G(t)$ is the impulse response of the optimal postfilter for estimating $x_s(t)$ from the $y(j)$'s. $y(j)$ is given by

$$y(j) = N x_s(j), \quad (4.31)$$

and $x_s(t)$ is specified by (4.15).

The state equation and the observation model given by (4.15) and (4.31) are similar to those given by (2.1) through (2.3) with $x(t)$ replaced by $x_s(t)$, H replaced by N , and $w(t)$ replaced by $u(t)$. $u(t)$ is a white noise uncorrelated with $x_s(\tau)$ for $\tau < t$. The model given by (4.15) and (4.31) fits in the standard model used in Chapter 2 with $n(k)=0$. Thus, the optimum postfilter can be obtained by the analysis we have in Chapter 2, assuming N is known.

Proof. Since $G(t)$ is the impulse response of the optimum postfilter for estimating $x_s(t)$ from the $y(jT)$'s, where $y(j)$ is the sampled value of the output of the prefilter and is a linear combination of $x_s(j)$ (see Eq. (4.31)), $G(t)$ is the impulse response of the optimum postfilter for estimating $x(t)$ with the given prefilter from Theorem 4.1. Also, since $d(t) = C x(t)$, $g(t) = C G(t)$ is the impulse response of the optimum postfilter for estimating $d(t)$ with the given prefilter. Thus, all we need to prove is that the prefilter specified above also satisfies the necessary and sufficient condition (Eq. (3.38) with $C_3=0$) for it to be an optimum prefilter.

Since the impulse response of the prefilter specified above is a continuous function of time, the postfilter must satisfy the equation

$$\xi_j = 0$$

for all $j \neq 0$, where ξ_j is given by (3.55), in order for the prefilter to be optimum, as follows from the discussion prior to the statement of Theorem 4.3. Let

$$\hat{\xi}(t/k) = \hat{\xi}(t/-\infty, k, y(i)) = \text{the linear least-square estimate of } \xi(t) \text{ based on the observation of } y(i) \text{ for } -\infty < i \leq k,$$

and

$$\hat{\xi}(t) = \hat{\xi}(t/m),$$

where $m = \langle t/T \rangle$. Then

$$\hat{x}_s(t) = \sum_{i=-\infty}^m G(t-iT) y(i) \quad (4.32)$$

From (2.27) and (4.15), $x_s(t)$ can also be expressed as

$$\hat{x}_s(t) = \Phi(t-mT) \hat{x}_s(m) \quad (4.33)$$

for $mT \leq t < (m+1)T$. $\hat{x}_s(m)$ can be obtained from (B.9) through (B.16) with $V_{wn} = 0$ (i.e., $K_0 = 0$), $V_n = 0$, and V_w obtained from the discretized state equation of (4.15); and with x replaced by x_s , H replaced by N , and r replaced by y . Furthermore, assume that the estimation process is at steady-state condition. Then

$$\hat{x}_s(m) = \Phi(T) \hat{x}_s(m-1) + K_d [y(m) - N(m) \Phi(T) \hat{x}_s(m-1)], \quad (4.34)$$

where we have used the fact that $\Phi(m, m-1) = \Phi(T)$ (Φ is the transition matrix of A). After combining terms, (4.34) becomes

$$\hat{x}_s(m) = [I - K_d N] \Phi(T) \hat{x}_s(m-1) + K_d y(m), \quad (4.35)$$

where I is the identity matrix. By repeatedly using

(4.35), we obtain

$$\hat{x}_s(m) = \sum_{i=-\infty}^m \mathcal{V}^{m-i} K_d y(i), \quad (4.36)$$

where

$$\mathcal{V} = [I - K_d N] \Phi(T). \quad (4.37)$$

Substituting (4.36) into (4.33) to obtain

$$\hat{x}_s(t) = \sum_{i=-\infty}^m \Phi(t-mT) \mathcal{V}^{m-i} K_d y(i) \quad (4.38)$$

for $mT \leq t < (m+1)T$. Comparing (4.38) with (4.32) yields, for $mT \leq t < (m+1)T$,

$$G(t-iT) = \Phi(t-mT) \mathcal{V}^{m-i} K_d \quad (4.39)$$

Let $v = t - iT$ and $\ell = m - i$. Then for $mT \leq t < (m+1)T$,

$$\begin{aligned} v &= t - iT = (t - mT) + (m - i)T \\ &= (t - mT) + \ell T. \end{aligned} \quad (4.40)$$

Since $mT \leq t < (m+1)T$ in (4.39), (4.39) becomes (after letting $v = t - iT$ and $\ell = m - i$)

$$G(v) = \Phi(v - \ell T) \mathcal{V}^{\ell} K_d \quad \text{for } \ell T \leq v < (\ell+1)T \quad (4.41)$$

where is given by (4.37). Next, let us investigate the value of ξ_1 . From (3.55),

$$\begin{aligned}\xi_1 &= \int_0^\infty g(v) g(v+T) dv \\ * &= \int_0^T g(v) g(v+T) dv + \int_T^{2T} g(v) g(v+T) dv + \dots \\ &\quad + \left[\int_T^{(\ell+1)T} g(v) g(v+T) dv + \dots \right] \quad (4.42)\end{aligned}$$

From (4.41) and (4.30), (4.42) becomes

$$\begin{aligned}\xi_1 &= \int_0^T g(v) C\Phi(v) K_d dv + \int_T^{2T} g(v) C\Phi(v-T) \mathcal{V}^2 K_d + \dots + \\ &\quad \int_{(\ell-1)T}^{\ell T} g(v) C\Phi(v-(\ell-1)T) \mathcal{V}^\ell K_d dv + \\ &\quad \int_{\ell T}^{(\ell+1)T} g(v) C\Phi(v-\ell T) \mathcal{V}^{\ell+1} K_d dv + \dots \quad (4.43)\end{aligned}$$

Substituting (4.37) into the second term on the right-hand side of (4.43) and using Lemma 3.1 to obtain

$$\begin{aligned}\int_T^{2T} g(v) C\Phi(v-T) \mathcal{V}^2 K_d dv &= \int_T^{2T} g(v) \cdot C\Phi(v) \Phi(-T) [I - K_d N] \times \\ &\quad \Phi(T) \mathcal{V} K_d dv\end{aligned}$$

*Since $g(v)$ is a piece-wise continuous function.

$$\begin{aligned}
&= \int_T^{2T} g(v) C\Phi(v) \mathcal{V} K_d dv - C \int_T^{2T} g(v) \Phi(v-T) K_d N\Phi(T) \mathcal{V} \\
&\quad \times K_d dv \quad (4.44)
\end{aligned}$$

The integrand of the first term on the right-hand side of (4.44) is in the same form as the integrand of the first term on the right-hand side of (4.43). By letting $\mu = v - T$, the second term on the right-hand side of (4.44) becomes

$$\begin{aligned}
&\int_T^{2T} g(v) C\Phi(v-T) K_d N\Phi(T) \mathcal{V} K_d dv \\
&= \int_0^T g(\mu+T) C\Phi(\mu) K_d N\Phi(T) \mathcal{V} K_d d\mu \\
&= \int_0^T g(\mu+T) g(\mu) N\Phi(T) \mathcal{V} K_d d\mu \\
&= \int_0^T C\Phi(\mu) \mathcal{V} g(\mu) N\Phi(T) \mathcal{V} K_d d\mu \\
&= C \left[\int_0^T g(\mu) \Phi(\mu) d\mu \right] [\mathcal{V} N\Phi(T) \mathcal{V} K_d], \quad (4.45)
\end{aligned}$$

where the second and third equality follow from the substitution of (4.41), and the last equality follows from the fact that $g(\mu)$ is a scalar. Similarly, the $(\ell+1)$ th term on the right-hand side of (4.43) can be written as

$$\int_{\ell T}^{(\ell+1)T} g(v) C\Phi(v-\ell T) \mathcal{V}^{\ell+1} K_d dv$$

$$\begin{aligned}
&= \int_{\ell T}^{(\ell+1)T} g(v) C\Phi(v) \Phi(-\ell T) \cdot [I - K_d N] \Phi(T) \mathcal{V}^\ell K_d dv \\
&= \int_{\ell T}^{(\ell+1)T} g(v) C\Phi(v - (\ell-1)T) \mathcal{V}^\ell K_d dv - \\
&\quad \int_{\ell T}^{(\ell+1)T} g(v) C\Phi(v - \ell T) K_d N \Phi(T) \mathcal{V}^\ell K_d dv.
\end{aligned}
\tag{4.46}$$

The integrand of the first term on the right-hand side of (4.46) is in the same form as the integrand of the ℓ th term on the right-hand side of (4.43). By letting $\mu = v - \ell T$, the second term on the right-hand side of (4.46) can be written as

$$\begin{aligned}
&\int_{\ell T}^{(\ell+1)T} g(v) C\Phi(v - \ell T) K_d N \Phi(T) \mathcal{V}^\ell K_d dv \\
&= \int_0^T g(\mu + \ell T) [C\Phi(\mu) K_d] N \Phi(T) \mathcal{V}^\ell K_d d\mu \\
&= \int_0^T g(\mu + \ell T) g(\mu) N \Phi(T) \mathcal{V}^\ell K_d d\mu \\
&= \int_0^T C\Phi(\mu) \mathcal{V}^\ell g(\mu) N \Phi(T) \mathcal{V}^\ell K_d d\mu \\
&= C \left[\int_0^T g(\mu) \Phi(\mu) d\mu \right] [\mathcal{V}^\ell N \Phi(T) \mathcal{V}^\ell K_d]
\end{aligned}
\tag{4.47}$$

Through the use of Eqs. (4.44) through (4.47), Eq. (4.43) becomes

$$\begin{aligned}
 \xi_1 &= C \left[\int_0^T g(v) \Phi(v) dv + \int_T^{2T} g(v) \Phi(v) dv + \dots \right] \mathcal{V} K_d - \\
 &\quad C \left[\int_0^T g(v) \Phi(v) dv + \int_T^{2T} g(v) \Phi(v) dv + \dots \right] \times \\
 &\quad [\mathcal{V} N \Phi(T) \mathcal{V} K_d] - \dots - \\
 &\quad C \left[\int_0^T g(v) \Phi(v) dv + \int_T^{2T} g(v) \Phi(v) dv + \dots \right] \times \\
 &\quad [\mathcal{V}^\ell N \Phi(T) \mathcal{V}^\ell K_d] - \dots \\
 &= C \left[\int_0^\infty g(v) \Phi(v) dv \right] \mathcal{V} [I - N \Phi(T) - \dots - \mathcal{V}^{\ell-1} \times \\
 &\quad N \Phi(T) \mathcal{V}^\ell - \dots] K_d \\
 &= C L \mathcal{V} [I - N \Phi(T) - \dots - \mathcal{V}^{\ell-1} N \Phi(T) - \dots] K_d, \quad (4.48)
 \end{aligned}$$

where the last equality follows from the substitution of (3.63). Since $N = \xi_0^{-1} C L$ from (4.27),

$$C L \mathcal{V} = \xi_0 N \mathcal{V}. \quad (4.49)$$

From (B.13_a) and (B.14), it follows that with $V_n = 0$ (where H is replaced by N),

$$\begin{aligned}
 N K_d &= N V_X^{\mathcal{V}}(j/j-1) N' [N V_X^{\mathcal{V}}(j/j-1) N']^{-1} \\
 &= 1
 \end{aligned} \quad (4.50)$$

Thus,

$$\begin{aligned}
 \xi_0 N \mathcal{V} &= \xi_0 N [I - K_d N] \Phi(T) \\
 &= \xi_0 [N - N K_d N] \Phi(T) \\
 &= \xi_0 [N - N] \Phi(T) = 0
 \end{aligned} \tag{4.51}$$

Furthermore, for a stable system, $\mathcal{V}^\ell \rightarrow 0$ as $\ell \rightarrow \infty$. Through the use of (4.49) and (4.51), we finally obtain

$$\begin{aligned}
 \xi_1 &= \xi_0 N \mathcal{V} [I - N\Phi(T) - \dots - \mathcal{V}^{\ell-1} N\Phi(T) - \dots] K_d \\
 &= 0 \cdot [I - N\Phi(T) - \dots - \mathcal{V}^{\ell-1} N\Phi(T) - \dots] \cdot K_d \\
 &= 0.
 \end{aligned} \tag{4.52}$$

Similarly, one can obtain

$$\begin{aligned}
 \xi_j &= C L \mathcal{V} [I - N\Phi(T) - \dots - \mathcal{V}^{\ell-1} N\Phi(T) - \dots] \mathcal{V}^{j-1} K_d \\
 &= 0
 \end{aligned} \tag{4.53}$$

for all $j \geq 1$. Since $\xi_j = \xi_{-j}$ from Lemma 4.3, we conclude that

$$\xi_j = 0 \quad \text{for all } j \neq 0 \tag{4.54}$$

from (4.53).

The rest of the proof can be done easily. Since the prefilter is assumed to be the combination of a continuous KSVF followed by a constant coder N ,

$$\begin{aligned}
 f(t) &= N F_{ks}(t) \\
 &= C [\xi_0^{-1} L V_u^{-1} R_{xu}(t) u_s(t) \otimes f_u(t)] \quad (4.55)
 \end{aligned}$$

where the second equality follows from the substitution of Eqs. (3.86) and (4.27). Let

$$F_{yr}(t) = \xi_0^{-1} L V_u^{-1} R_{xu}(t) u_s(t) \quad (4.56_a)$$

and

$$f_{yr}(t) = C F_{yr}(t). \quad (4.56_b)$$

Substituting (4.56) into (4.55) to obtain

$$f(t) = C [F_{yr}(t) \otimes f_u(t)] = f_{yr}(t) \otimes f_u(t) \quad (4.57)$$

Since $\xi_j=0$ for $j \neq 0$ from (4.54), (3.70) and (3.71) can be satisfied by letting $F_y(t)=F_{yr}(t)$. Since (3.70) and (3.71) completely specify the optimum $F_y(t)$, the $f(t)$ given by (4.57) is the optimum prefilter by comparing (4.56_b) with (3.76) and (4.57) with (3.83). Thus, the proof is completed.

We have seen from this theorem that with the optimum prefilter, one tries to estimate x_s from the $y(j)$'s, where the quantities are modeled by (4.15) and (4.31). Without the optimum prefilter, one tries to estimate x from the $r(j)$'s, where the quantities are modeled by (2.1)

through (2.3). As far as the postfilter is concerned, the models are quite similar to each other (with and without the optimum prefilter). The only difference is that in the case with the optimum prefilter, one has a certain freedom in choosing the coder N in order to reduce the error. Without the optimum prefilter, the observation matrix H is fixed.

The block diagram of the (degenerated) jointly optimized system with $n_2=0$ is shown in Fig. 4.2.

With the background we have so far, the only problem left is to find the coder N , which we solve in the next section by the nonlinear programming techniques.

4.4 The Application of Nonlinear Programming

In this section, the prefilter is restricted to the class of filters that can be represented by the combination of a continuous KSVF followed by a constant coder. If n_2 in Fig. 1.1 is absent, this structure of the prefilter is optimum as proven in the previous section. Otherwise, it becomes a suboptimum system.

4.4.1 Outline of procedure

Since the coder N is a constant matrix, the output of the prefilter is given by (4.31). The observation

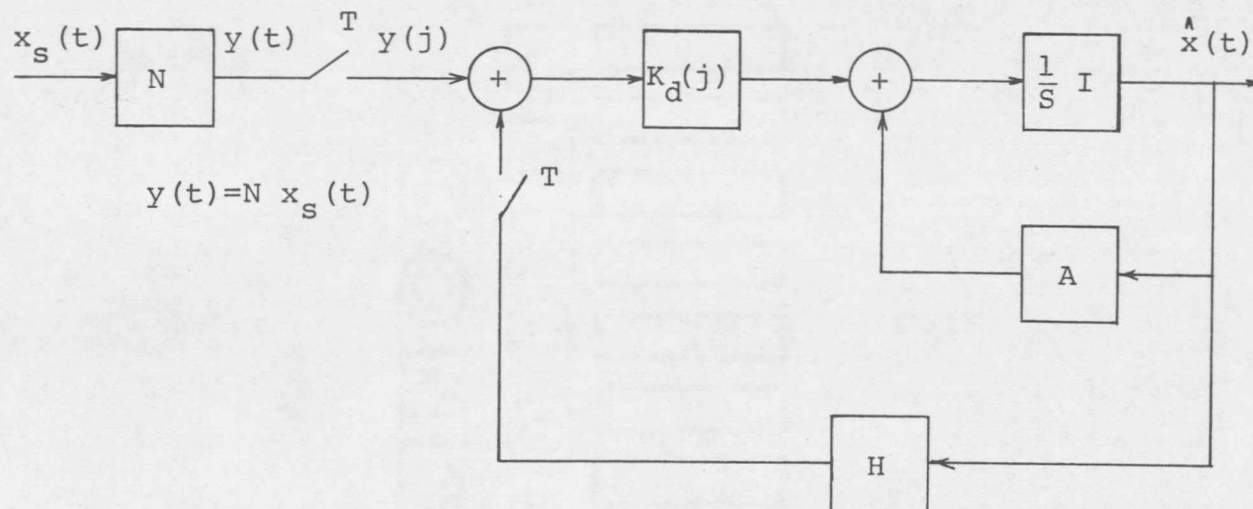


Figure 4.2 Degeneration of the joint optimization problem when $n_2=0$.

process at the receiving end, $s(j)$, is given by (3.5); and $x_s(t)$ is specified by (4.15). The problem of finding $\hat{x}_s(t/-\infty, m, s(j))$ has been solved in Chapter 2 assuming the coder N is known. Thus, the problem is to find the coder N . Let $\hat{\xi}(t) = \hat{\xi}(t/-\infty, m, s(j))$. Then from Theorem 4.1, it follows that

$$E\{[x(t) - \hat{x}(t)]^2\} = E\{[x(t) - \hat{x}_s(t)]^2\} + E\{[x_s(t) - \hat{x}_s(t)]^2\}. \quad (4.58)$$

Since $d(t) = C x(t)$,

$$E\{[d(t) - \hat{d}(t)]^2\} = C E\{[x(t)]^2\} C' + C E\{[x_s(t) - \hat{x}_s(t)]^2\} C'. \quad (4.59)$$

The first term on the right-hand side of (4.59) is independent of the coder N . For the second term, write

$$\begin{aligned} E\{[x_s(t) - \hat{x}_s(t)]^2\} &= V_{x_s}^u(t) \\ &= \Phi(t-mT) V_{x_s}^u(m) \Phi'(t-mT) + \\ &\quad \int_{mT}^t \Phi(t-v) K_c V_u K_c' \Phi'(t-v) dv. \end{aligned} \quad (4.60)$$

for $mT \leq t < (m+1)T$ from (D.37). The second term on the right-hand side of (4.60) is again independent of the coder N .

Integrating the first term on the right-hand side of (4.60) from $t=0$ to $t=T$, then dividing it by T to obtain ($m=0$ for $0 \leq t < T$)

$$J_S = (1/T) \int_0^T \Phi(t) V_{x_S}^{\sim}(0) \Phi'(t) dt \quad (4.61)$$

Premultiplying by C and then postmultiplying by C' on both sides of (4.61), the function that we want to minimize with respect to the coder N is

$$J_N = C J_S C' = (1/T) C \cdot \int_0^T \Phi(t) V_{x_S}^{\sim}(0) \Phi'(t) dt C'. \quad (4.62)$$

J_N corresponds to the part of the error that depends on the coder N , and J_N corresponds to the error with a penalty function given by (D.16). Since steady-state condition is assumed (Constraint 3.5), $V_{x_S}^{\sim}(0) = V_{x_S}^{\sim}$ is specified by the following equations

$$K_d = V_{x_S}^{\sim} N' [N V_{x_S}^{\sim} N' + V_{n_2}]^{-1}, \quad (4.63)$$

$$V_{x_S}^{\sim} = \Phi(T) V_{x_S}^{\sim} \Phi'(T) + V_{w_d}, \quad (4.64)$$

and

$$V_{x_S}^{\sim} = [I - K_d N] V_{x_S}^{\sim}, \quad (4.65)$$

where

$$V_{w_d} = \int_0^T \Phi(t) K_c V_u K_c' \Phi'(t) dt, \quad (4.66)$$

where V_u is the variance of the innovation process $u(t)$ of $r(t)$, and K_c is the continuous KSVF gain for estimating the $x(t)$ from the $r(t)$ (see Eqs. (3.1) through (3.3)). Eqs. (4.63) through (4.65) are obtained from Eqs. (B.13) through (B.16) with $V_{wn}=0$, $V_w=V_{w_d}$; with x replaced by x_s , H replaced by N , and V_n replaced by V_{n_2} , where n_2 is specified by (3.5); and with

$$V_{x_s}^{\sim}(j) = V_{x_s}^{\sim}(j-1) = V_{x_s}^{\sim}, \quad (4.67)$$

$$V_{x_s}^{\sim}(j/j-1) = V_{x_s}^{\sim}(j-1/j-2) = V_{x_s}^{\sim}, \quad (4.68)$$

and

$$K_d(j) = K_d(j-1). \quad (4.69)$$

Eq. (4.66) is obtained from (C.9) with $t=(k+1)T-\tau$; and with $\Gamma(\tau)$ replaced by K_c , and $V_w(\tau)$ replaced by V_u . Let

$$V_{x_s}^{\sim} = \begin{bmatrix} e_1 & e_2 & \cdot & \cdot & \cdot & e_n \\ e_2 & e_{n+1} & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & e_l \end{bmatrix} \quad (4.70)$$

where $\ell = n + \binom{n}{2}$. n is the dimension of the state vector.

Also, let

$$N = [n_1, n_2, \dots, n_n]. \quad (4.71)$$

A brute-force method to find the optimal coder is described at first as follows: Eqs. (4.63) through (4.65) consist of a set of Riccati equations that enables one to solve for the $V_{x_s}^v$ in terms of such parameters as N , V_{n_2} and V_{w_d} , etc. One then substitutes the expression of the $V_{x_s}^v$ found above into (4.62). Finally, the n_i 's can be found by the ordinary calculus. In simple cases such as when the input signal is a first-order Gauss-Markov process, this brute-force method can be applied easily. But in more complicated cases, the Riccati equations of (4.63) through (4.65) are rather hard to solve. In these cases, the nonlinear programming techniques can be applied to find the optimal coder N . Substituting (4.70) and (4.71) into (4.63) through (4.65), we obtain a set of equations in the n_i 's and e_i 's in the form

$$FE_i[e_1, e_2, \dots, e_i; n_1, n_2, \dots, n_n] = \sum_{k=1}^{m_i} \psi_{ik} \left[\prod_{t=1}^{\ell} (e_t)^{p_{tk}} \right] \times \left[\sum_{j=1}^n (n_j)^{o_{jk}} \right] = 0 \quad (4.72)$$

where all the quantities are scalars. Substituting (4.70) into (4.62) to obtain

$$J_N = \sum_{j=1}^{m_0} b_j \left[\sum_{k=1}^{\ell} (e_k)^{i_{jk}} \right]. \quad (4.73)$$

Consider both the e_i 's and n_i 's as variables; the problem is to minimize J_N subject to the equality constraints of (4.72), and the constraints that the diagonal elements of the $V_{x_s}^v$ (see (4.70)) be nonnegative. This is a typical problem that can be handled by the nonlinear programming and search techniques. An algorithm developed by Pierre (45) has been applied to this problem. Examples will be given in Section 4.4.2, and in Chapter 5.

4.4.2 Examples

Example 4.1: A second-order Gauss-Markov process as the input signal.

Referring to Fig. 1.1, the signal and observation models at the sending end are given by the following equations:

$$\dot{x}(t) = \begin{bmatrix} -a & 0 \\ 0 & -b \end{bmatrix} x(t) + w(t). \quad (4.74)$$

and

$$r(t) = [1, 1] x(t) + n_1(t), \quad (4.75)$$

where

$$R_w(t-\tau) = \begin{bmatrix} 2a\sigma_a & 0 \\ 0 & 2b\sigma_b \end{bmatrix} \delta(t-\tau),$$

and

$$R_{n_1}(t-\tau) = V_n \delta(t-\tau).$$

At the receiving end, $n_2=0$. Since the gain of the pre-filter is nonunique in the absence of n_2 (see the discussion after (E.33) in Appendix E), one can let

$$N = [1, n_2].$$

From the procedure outlined in Section 4.4.1, it is found that $e_2 = -n_2 e_3$ and $e_1 = n_2^2 e_3$, where the e_i 's are defined by (4.70) with $\ell=3$. Letting $e=e_3$, the function to be minimized can be expressed in the form

$$J_N(e, n_2) = [n_2 \cdot e \cdot (1 - \alpha_a \alpha_b) - n_2^2 e \cdot (1 - \alpha_a^2) - e \cdot (1 - \alpha_b^2) / 3.] / 2T$$

subject to the equality constraint

$$FE(e, n_2) = \varphi_c e^2 + e (\varphi_d - \varphi_a) + \varphi_b,$$

where $\alpha_a = \exp(-aT)$ and $\alpha_b = \exp(-bT)$, and

$$\psi_a = k_1 \alpha_b^2 + k_3 n_2^2 \alpha_a^2 + 2 k_2 n_2 \alpha_a \alpha_b,$$

$$\psi_b = k_2^2 - k_1 k_3,$$

$$\psi_c = n_2^2 (\alpha_a - \alpha_b)^2,$$

$$\psi_d = k_1 + 2 n_2 k_2 + n_2^2 k_3,$$

$$\begin{bmatrix} k_1 & k_2 \\ k_2 & k_3 \end{bmatrix} = \int_0^T \Phi(t) K_C V_u K_C' \Phi'(t) dt,$$

and

K_C = the continuous KSVF gain for estimating the $x(t)$ from the $r(\tau)$ for $-\infty < \tau \leq t$, where the $x(t)$ and $r(\tau)$ are specified by (4.74) and (4.75).

Furthermore $e \geq 0$. The result found by the nonlinear programming technique for the case $a=-1$, $b=-3$, $\sigma_a=0.1875$, $\sigma_b=0.1042$, $V_n=1$, $T=0.1$, and $d(t)=[1, 1] x(t)$ is

$$n_2 = 0.91.$$

The steady-state prefilter is specified by the following equations:

$$\dot{x}_s(t) = A x_s(t) + K_C u(t), \quad (4.76)$$

and

$$y(t) = [1, 0.91] x_s(t) \quad (4.77)$$

where

$$u(t) = r(t) - [1, 1] \cdot x_s(t),$$

$$A = \begin{bmatrix} -1 & 0 \\ 0 & -3 \end{bmatrix},$$

and

$$K_c = \begin{bmatrix} 0.1690 \\ 0.0984 \end{bmatrix}$$

The steady-state optimal postfilter is specified by the following equations:

$$\dot{\hat{x}}(t) = A \hat{x}(t) + K_d \sum_{k=-\infty}^{\infty} u_d(i) \delta(t-iT) \quad (4.78)$$

and

$$d^0(t) = [1, 1] x(t), \quad (4.79)$$

where

$$u_d(i) = y(iT) - N \cdot \begin{bmatrix} -\alpha_a & 0 \\ 0 & -\alpha_b \end{bmatrix} \hat{x}(iT),$$

and

$$K_d = [0.6747, 0.3574].$$

To check the validity of Eq. (4.27) (Theorem 4.3), one has to find the impulse response of the optimal post-filter $g(t)$, which is found to be

$$g(t) = [g_a(t) + g_b(t)] u_p(t, 0) + \sum_{i=1}^{\infty} (0.7942)^{i-1} \\ \cdot [(0.7942 - \alpha_b) g_a(t - iT) + (0.7942 - \alpha_a) g_b(t - iT)] \\ \cdot u_p(t, iT),$$

where

$$g_a(t) = 0.6750 \exp(-t),$$

$$g_b(t) = 0.3575 \exp(-3t),$$

$$u_p(t, iT) = u_s(t - iT) - u_s(t - (i+1)T),$$

and

$$u_s(t) \triangleq \text{the unit step function.}$$

From (3.54), since $n_2=0$, $c_3=0$ and

$$\xi_0 = \int_0^T g^2(t) dt$$

$$\begin{aligned}
&= \int_0^T [g_a(t) + g_b(t)]^2 dt + \sum_{i=1}^{\infty} (0.7942)^{2(i-1)} \int_{iT}^{(i+1)T} \\
&\quad [(0.7942 - \alpha_b) g_a(t - iT) + (0.7942 - \alpha_a) g_b(t - iT)]^2 dt \\
&= 0.09069.
\end{aligned}$$

Also, since

$$\Phi(t) = \begin{bmatrix} \exp(-t) & 0 \\ 0 & \exp(-3t) \end{bmatrix},$$

from (4.29),

$$\begin{aligned}
L &= \int_0^{\infty} g(t) \Phi(t) dt \\
&= [0.0907, 0.0825].
\end{aligned}$$

We thus have

$$\begin{aligned}
\xi_0^{-1} C L &= [0.9996, 0.9097] \\
&\approx [1, 0.91] = N.
\end{aligned}$$

Hence, except for a very slight round-off error, (4.27) has been verified to be correct.

We used a diagonal matrix in this example in order to simplify computation. The same procedure can be used to handle the problem with a nondiagonal matrix A .

Example 4.2: A second-order Butterworth process as the input signal.

The signal and observation models are

$$\dot{x}(t) = \begin{bmatrix} 0 & 1 \\ -\omega_c^2 & -\omega_c\sqrt{2} \end{bmatrix} x(t) + w(t)$$

and

$$r(t) = [1, 0] x(t) + n_1(t),$$

where

$$R_w(t-\tau) = \begin{bmatrix} 0 & 0 \\ 0 & 2\sqrt{2}\omega_c^3 \end{bmatrix} \delta(t-\tau)$$

and

$$R_{n_1}(t-\tau) = (2.45/\omega_c) \delta(t-\tau).$$

Again, $n_2=0$. By an argument similar to that used in Example 4.1, one can let $N=[1, n_2]$. The procedure for solving this example is similar to that of Example 4.1, except that the expressions are more complicated. The details are omitted. The result found by nonlinear programming for the case in which $\omega_c=1$, $T=\pi$, and $d(t)=[1, 0] \cdot x(t)$ is

$$n_2 = 0.4239.$$

This result checks with the one found by Petersen (42).

These two examples primarily serve as a check for the correctness of the development we have here. More examples with physical interpretations will be given in Chapter 5.

Chapter 5

SYSTEM PERFORMANCE

5.1 Problem Statement and Preliminary Remarks

The design of the optimum realizable pre- and postfilters, and the design of the fixed-lag postfilter have been studied using the state-space approach in previous Chapters. It is of interest to investigate the performance of various systems in this Chapter. We will study the physical interpretations of this problem such as the primary function of the optimum realizable prefilter; how much error reduction is achieved by the addition of a prefilter; how much complication this creates in implementation; how the addition of a prefilter compares with allowing delay at the receiving end, i.e., with having a finite-lag optimum postfilter in the absence of a prefilter; how close the performance of the optimum realizable filter is to that of the optimum unrealizable filters; and how close the performances of certain suboptimum systems are to that of the optimum systems. Unless specifically mentioned, all the systems considered in this chapter have a single sampled-data link.

Since only stationary inputs, time-invariant systems, and steady-state conditions will be considered in this chapter, it is helpful to look at the frequency-domain approach to this problem briefly first. This will serve to check the results obtained in the previous chapter of this thesis.

5.2 The Frequency-Domain Approach to the Filtering Problem

Most of the solutions obtained using the frequency-domain approach to filtering problem are based on the solution of Wiener's problem, where the function to be minimized with respect to $g(t)$ (or $S_g(f)$, the fourier transform of $g(t)$) is

$$J = R_d(0) - 2 \int_{-\infty}^{\infty} R_{rd}(\tau + \alpha) g(\tau) d\tau + \int_{-\infty}^{\infty} g(\rho) \int_{-\infty}^{\infty} R_r(\rho - \tau) g(\tau) d\rho d\tau. \quad (5.1)$$

In the frequency domain,

$$J = R_d(0) - 2 \int_{-\infty}^{\infty} S_g(f) S_{rd}(-f) df + \int_{-\infty}^{\infty} |S_g(f)|^2 S_r(f) df, \quad (5.2)$$

where $S.(f)$ is the fourier transform of $R.(t)$.

The solution for the optimum unrealizable filter obtained from (5.2) is

$$S_g(f) = S_{dr}(f)/S_r(f), \quad (5.3_a)$$

and the optimum realizable filter is

$$S_g(f) = (1/S_r^+(f)) \cdot [S_{dr}(f)/S_r^-(f)]_+, \quad (5.3_b)$$

where $\xi(f) = \xi^+(f) \cdot \xi^-(f)$ with $\xi^+(f)$ ($\xi^-(f)$) containing upper (lower) half plane poles and zeros only, and $\xi(f) = [\xi(f)]_+ + [\xi(f)]_-$ with

$$[\xi(f)]_+ = \int_0^\infty \xi(t) \exp(-j2\pi ft) dt, \quad (5.4)$$

and

$$[\xi(f)]_- = \int_{-\infty}^0 \xi(t) \exp(-j2\pi ft) dt, \quad (5.5)$$

where j denotes the imaginary number. $\xi(f)$ is the fourier transform of any real time function.

Let us consider the system shown in Fig. 1.1_c with the following assumptions: the number of sampled-data link is one; the inputs are stationary; the filters are time-invariant; n_2 is an independent noise. Furthermore, a power constraint is imposed on the prefilter. Then the function we want to minimize with respect to the

pre- and postfilters is given by (E.42) with all the quantities being scalars:

$$\begin{aligned}
 J = R_d(0) - (2/T) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} R_{dr}(\beta+v) f(\beta) g(v) dv d\beta + (1/T) \\
 \times \left\{ \int_{-\infty}^{\infty} g(v-lT) g(v) dv \left[\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(\alpha) R_r(lT+\beta-\alpha) \right. \right. \\
 \times f(\beta) d\alpha d\beta + R_{n_2}(l) \} + \lambda \cdot \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(\alpha) R_r(\beta-\alpha) \\
 \times f(\beta) d\alpha d\beta
 \end{aligned} \quad (5.6)$$

with the power constraint being given by (E.40). In the frequency domain,

$$\begin{aligned}
 J = R_d(0) - (2/T) \int_{-\infty}^{\infty} S_g(f) S_f(f) S_{dr}(-f) df \\
 + \int_{-\infty}^{\infty} |S_g(f)|^2 [|S_f(f)|^2 S_r(f) + S_{n_2}(f)]^* df \\
 + \lambda \cdot \int_{-\infty}^{\infty} |S_f(f)|^2 S_r(f) df,
 \end{aligned} \quad (5.7)$$

where we have used the relations

$$\int_{-\infty}^{\infty} a(t) b(t) dt = \int_{-\infty}^{\infty} S_a(f) S_b(-f) df, \quad (5.8)$$

and

$$\begin{aligned}
 [S_a(f)]^* &= \sum_{k=-\infty}^{\infty} a(kT) \exp(-j2\pi f kT) \\
 &= (1/T) \sum_{k=-\infty}^{\infty} S_a(f - (k/T))
 \end{aligned}$$

$\triangleq$ the spectrum of the sampled version of $a(t)$. (5.9)

where $S_a(f)$ and $S_b(f)$ are the fourier transforms of real time functions $a(t)$ and $b(t)$, respectively. The last term on the right-hand side of (5.7) is independent of $S_g(f)$. Comparing (5.7) with (5.2), the transfer function of the optimum unrealizable postfilter is seen to be

$$S_g(f) = S_f(-f) S_{dr}(f)/D(f) \quad (5.10)$$

from (5.3_a), where

$$D(f) = [|S_f(f)|^2 S_r(f) + S_{n_2}(f)]^*; \quad (5.11)$$

and the transfer function of the optimum realizable postfilter is

$$S_g(f) = (1/D^+(f)) [S_f(-f) S_{dr}(f)/D^-(f)]_+. \quad (5.12)$$

from (5.3_b). In terms of the Laplace transform variable s , the optimum realizable postfilter can be written as

(by letting $f=s/2\pi j$, where j is the imaginary number)

$$S_g(s) = (1/D^+(s)) [S_f(-s) S_{dr}(s)/D^-(s)]_+, \quad (5.13)$$

where

$$D(s) = [|S_f(s)|^2 S_r(s) + S_{n_2}(s)]^* = [|S_f(s)|^2 S_r(s)]^* + [S_{n_2}^*(s)]. \quad (5.14)$$

For the prefilter, since

$$\int_{-\infty}^{\infty} S_a^*(f) S_b(f) df = \int_{-\infty}^{\infty} S_a(f) S_b^*(f) df \quad (\text{reference number 9}) \quad (5.15)$$

(5.7) becomes

$$\begin{aligned} J = R_d(0) - (2/T) \int_{-\infty}^{\infty} S_f(f) S_g(f) S_{dr}(-f) df + (1/T) \\ \times \int_{-\infty}^{\infty} [|S_f(f)|^2 S_r(f) + S_{n_2}(f)] [|S_g(f)|^2]^* df \\ + \lambda \cdot \int_{-\infty}^{\infty} |S_f(f)|^2 S_r(f) df; \end{aligned} \quad (5.16)$$

i.e.,

$$\begin{aligned} J = R_d(0) - (2/T) \int_{-\infty}^{\infty} S_f(f) S_g(f) S_{dr}(-f) df \\ + (1/T) \int_{-\infty}^{\infty} |S_f(f)|^2 S_r(f) [(|S_g(f)|^2)^* + \lambda] df \\ + (1/T) \int_{-\infty}^{\infty} S_{n_2}(f) [|S_g(f)|^2]^* df. \end{aligned} \quad (5.17)$$

The last term on the right-hand side of (5.17) is independent of the prefilter. Comparing (5.17) with (5.2), we find that the optimum unrealizable prefilter is given by

$$S_f(f) = S_g(-f) S_{dr}(f) / S_e(f), \quad (5.18)$$

where

$$S_e(f) = \{[|S_g(f)|^2]^* + \lambda\} S_r(f); \quad (5.19)$$

from (5.3_a); and the optimum realizable prefilter is given by

$$S_f(f) = (1/S_e^+(f)) [S_g(-f) S_{dr}(f) / S_e^+(-f)]_+. \quad (5.20)$$

from (5.3_b). In terms of the Laplace transform variable s , the optimum realizable prefilter can be expressed as (by letting $f=s/2\pi j$ in (5.20), where j denotes the imaginary number)

$$S_f(s) = (1/S_e^+(s)) [S_g(-s) S_{dr}(s) / S_e^+(-s)]_+ \quad (5.21)$$

where

$$S_e(s) = \{[|S_g(s)|^2]^* + \lambda\} S_r(s). \quad (5.22)$$

5.2.1 Jointly optimized unrealizable pre- and postfilters

It has been shown by Chan and Donaldson (7) that the jointly optimized unrealizable pre- and postfilters in the case in which $n_1=0$ are ideal bandpass filters with identical passband. The width of the passband is $f_s=1/T$, where T is the sampling period. Thus, the jointly optimized pre- and postfilters eliminate aliasing. The same situation exists when $n_1 \neq 0$. The extension is straightforward and will not be proved here.

Brown (5) has shown that when $n_2=0$ ($n_1 \neq 0$; refer to Fig. 1.1_c), the passband is determined by

$$S_f(f) \text{ [or } S_g(f)] \begin{cases} \neq 0 & \text{if } |S_{dr}(f)|^2/S_r(f) > \\ & |S_{dr}(f-k/T)|^2/S_r(f-k/T) \\ & \text{for any } k \neq 0 \\ = 0 & \text{otherwise.} \end{cases} \quad (5.23)$$

Inside the passband, the filters satisfy the equation

$$S_f(f) S_g(f) = T S_{dr}(f)/S_r(f), \quad (5.24)$$

where $S_{dr}(f)/S_r(f)$ is the unreliable Wiener filter in the absence of sampling. The mean-square error between the desired signal $d(t)$ and the output of this system is given by

$$J = \int_{f_s} [S_d(f) - |S_{dr}(f)|^2/S_r(f)] df + \int_{f'_s} S_d(f) df, \quad (5.25)$$

where f'_s denotes the complement of the passband f_s . The first term on the right-hand side of (5.25) represents the error due to the primary observation noise n_1 (see Fig. 1.1), and the second term represents the error due to sampling. Recognize that

$$\int_{-\infty}^{\infty} [S_d(f) - S_{dr}(f)^2/S_r(f)] df \triangleq \text{the mean-square error}$$

caused by using the unrealizable Wiener filter in the absence of sampling.

Thus, inside the passband, the optimum unrealizable pre- and postfilters try to filter out the noise just as the Wiener filter does (32). But outside the passband, the system simply ignore both the signal and the noise.

Since $f_s = 1/T$, as T increases, f_s decreases. This means that as T increases, the optimum unrealizable pre- and postfilters become more and more ignorant of the effect of $n_1(t)$, and concentrate their efforts in dealing with the aliasing effect.

It is interesting to compare (5.25) with (4.9). The first term on the right-hand side of (4.9) represents the error due to n_1 by using the optimum realizable zero-lag pre- and postfilters. Since this is the error produced by using the Wiener filter in the absence of sampling, it is seen from (4.9) that this part of the error (due to n_1) remains unchanged when sampling is introduced in the case of optimum zero-lag filtering.

5.2.2 Jointly optimized realizable pre- and postfilters

When there is no delay factor, the only poles of the bracketed expressions on the right-hand side of (5.13) and (5.21) are the ones due to $S_{dr}(s)$. Let these poles be denoted by λ_i for $i=1,2,\dots,n_p$; and let

$$S_a(s) = [S_g(-s) S_{dr}(s)/S_e^-(s)]_+ = \sum_{i=1}^n \alpha_i / (s + \gamma_i), \quad (5.26)$$

and

$$S_b(s) = [S_f(-s) S_{dr}(s)/D^-(s)]_+ + \sum_{i=1}^n \beta_i / (s + \gamma_i), \quad (5.27)$$

where the λ_i 's are known, and the α_i and β_i 's are to be determined.

Substituting (5.27) into (5.13) to obtain

$$S_g(s) = (1/D^+(s)) \left[\sum_{i=1}^n \beta_i / (s + \gamma_i) \right]. \quad (5.28)$$

Thus,

$$\begin{aligned} [S_g(s) S_g(-s)]^* &= [1/D(s) \left[\sum_{i=1}^n \beta_i / (s + \gamma_i) \right] \left[\sum_{i=1}^n \beta_i / (-s + \gamma_i) \right]]^* \\ &= \frac{1}{[S_f^2(s) S_r(s)]^* + S_{n_2}^*(s)} \\ &\quad \left\{ \left[\sum_{i=1}^n \frac{\beta_i}{(s + \gamma_i)} \right] \cdot \left[\sum_{i=1}^n \frac{\beta_i}{(-s + \gamma_i)} \right] \right\}^*, \end{aligned} \quad (5.29)$$

where we have used the relation

$$[\xi^*(s) \eta(s)]^* = \xi^*(s) \eta^*(s) \quad (\text{reference 9}), \quad (5.30)$$

and $\xi(s)$ and $\eta(s)$ are the Laplace transforms of any real time functions. Similarly, by substituting (5.26) into (5.21), we obtain

$$[S_f(s) S_f(-s) S_r(s)]^* = \frac{1}{[S_g^2(s)]^* + \lambda} \left\{ \sum_{i=1}^n \frac{\alpha_i}{(s + \gamma_i)} \cdot \sum_{i=1}^n \frac{\alpha_i}{(-s + \gamma_i)} \right\}^*. \quad (5.31)$$

Let $X = [|S_g(s)|^2]^*$ and $Y = [|S_f(s)|^2 S_r(s)]^*$, then one can solve for X and Y from (5.29) and (5.31) in terms of

α_i , β_i , γ_i , and λ . By substituting the resulting X and Y into (5.13) and (5.21), one can find expressions for $S_g(s)$ and $S_f(s)$ in terms of α_i , β_i , etc. By substituting the resulting $S_g(s)$ and $S_f(s)$ into (5.26) (or (5.27)), the α_i and β_i 's can be found by equating coefficients. Finally, λ can be found from the power constraint given by (E.40) (with the quantities transformed into the s domain).

This method is in general rather cumbersome, even for special cases such as that in which $n_2=0$ (see Fig. 1.1_C). This special case has been studied by Chang (9) and Robbins (49). The result is summarized as follows. The optimum prefilter can be expressed as

$$S_f(s) = K_o S_a(s)/S_r^+(s), \quad (5.32)$$

and the optimum postfilter is given by

$$S_g(s) = S_b(s)/(K_o S_\phi(s)) \quad (5.33)$$

where K_o is an arbitrary constant,

$$S_a(s) = [S_b(-s) S_{dr}(s)/(S_\phi^-(s) S_r^-(s))]_+, \quad (5.34)$$

$$S_b(s) = [S_a(-s) S_{dr}(s)/S_\phi^-(s) S_r^-(s)]_+, \quad (5.35)$$

and

$$S_{\phi}(s) = [S_a(s) S_a(-s)]^* = [S_b(s) S_b(-s)]^* \quad (5.36)$$

Clearly,

$$S_a(s) S_a(-s) = S_b(s) S_b(-s). \quad (5.37)$$

The solution can, again, be started by assuming that

$$S_a(s) = \sum_{i=1}^n \alpha_i / (s + \gamma_i).$$

We will compare the result obtained in Chapter 4 with the above equations using an example. That is, we will show that the result obtained in Example 4.1 satisfies the above equations. It follows from (5.32) that $K_O S_a(s)$ represents the part of the prefilter that operates on the normalized innovation process of $r(t)$ --the primary observation process. Thus (refer to Fig. 1.1), the power spectrum of the output of the optimum prefilter is:

$$S_y(s) = K_O^2 S_a(s) S_a(-s). \quad (5.38)$$

From (4.76) and (4.77), since $u(t)$ is a white noise, we have

$$S_y(s) = [1, 0.91] (sI - A)^{-1} K_C V_u K_C' (-sI - A)^{-1} [1, 0.91]'. \quad (5.39)$$

Comparing (5.38) with (5.39) to obtain

$$\begin{aligned} K_o S_a(s) &= [1, 0.91] (sI-A)^{-1} K_c \sqrt{V_u} \\ &= \sqrt{V_u} 0.26 [(s+2.3075)/((s+1)(s+3))]. \end{aligned} \quad (5.40)$$

Similarly, $S_b(s)$ can be found from (4.78) and (4.79).

Since $V_{u_d} = 0.0057$ ($\triangleq$ the variance of the innovation process of $y(k)$), we have

$$S_b(s) = 0.0778 (s+2.3074)/[(s+1) \cdot (s+3)]. \quad (5.41)$$

First, it can be seen that with a properly chosen

K_o ,

$$S_a(s) S_a(-s) = S_b(s) S_b(-s).$$

Furthermore, we want to check to see whether (5.35) and (5.34) can also be satisfied by the $S_a(s)$ and $S_b(s)$ given by (5.40) and (5.41). From (5.36) and (5.40), it follows that

$$\begin{aligned} S_\phi(s) &= [S_a(s) S_a(-s)]^* \\ &= \left[\frac{V_{u_d} \cdot 0.26}{K_o} \right] \frac{\sqrt{\frac{x_b}{x_a}} \cdot (1-x_a z^{-1})}{(1-\alpha_a z^{-1})(1-\alpha_b z^{-1})} \cdot \frac{\sqrt{\frac{x_b}{x_a}} \cdot (1-x_a z)}{(1-\alpha_a z)(1-\alpha_b z)}, \end{aligned} \quad (5.42)$$

where $x_a=0.79$, $\alpha_a=0.9$, $\alpha_b=0.74$, and $x_b=0.067$. The $S_b(s)$ found by (5.35) with $S_a(s)$ given by (5.40) is (notice that the constant K_0 does not affect the result since it is cancelled by the numerator and denominator in (5.35)):

$$S_b(s) = 0.0779 (s+2.3077)/[(s+1)(s+3)]. \quad (5.43)$$

(5.41) represents the result found directly from the state-space approach, whereas (5.43) represents the result found using the equations derived from the frequency-domain approach assuming $S_a(s)$ is known. Comparing (5.41) with (5.43), we see that except for round-off error, (5.35) is satisfied by the result obtained in Example 4.1. It is obvious that (5.34) can also be satisfied by the $S_a(s)$ and $S_b(s)$ given by (5.40) and (5.41). Thus, except for a slight round-off error, the result found using the state-space approach satisfies the equations derived from the frequency-domain analysis.

The fact that the periodic factor of the transfer function of a prefilter can be shifted to the receiving end in the case in which $n_2=0$ can be seen from (5.30). Assume that $S_f(s)=S_{f_c}(s) S_{f_p}^*(s)$, where $S_{f_p}^*(s)$ represents the periodic factor of the prefilter. Then

$$\begin{aligned}
S_Y^*(s) &= [S_r(s) S_{f_c}(s) S_{f_c}(-s) [S_{f_p}(s) S_{f_p}(-s)]^*]^* \\
&= [S_r(s) S_{f_c}(s) S_{f_c}(-s)]^* [S_{f_p}(s) S_{f_p}(-s)]^*. \quad (5.44)
\end{aligned}$$

Thus, if y_1 is the output of f_c , one can sample y_1 first; transmit it through the channel; and then pass it through f_p . The resulting output will have power spectral density $S_Y^*(s)$.

5.3 Error Analysis.

The performances of various systems will be investigated by means of several examples in this section. Instead of providing the details about how those examples were worked out, we will outline the general procedures in the following subsection, and emphasize the results and their physical interpretations in the remainder of this chapter.

5.3.1 Outline of the general procedure used in working the examples

In the case in which no consideration is given to the design of the prefilter, expressions of the error covariance of the system with the optimum postfilter has already been given in Appendix D. With stationary inputs,

time-invariant systems, and with the assumption that the estimation process has reached a steady-state condition, then the $g_d(t)$ given by (D.14) and (D.15) can also be found by the frequency-domain approach (57). The result is:

$$g_d(t) = \mathcal{L}^{-1} [S_{dr}(s)/S_r^{*-}(s)], \quad (5.45)$$

where $\mathcal{L}^{-1}$ denotes the inverse Laplace transform, and $S_r^{*-}(s) = [S_r^-(s)]^*$. Eqs. (D.13) and (5.45) enables one to find the steady-state error when $S_r^*(s)$ and $S_{dr}(s)$ are known.

In the case in which the optimum prefilter is present but the system is constrained by Constraints 3.1 through 3.5 and Constraints 3.9 and 3.10, then Theorems 4.1 through 4.3, and Eq. (4.9) provide the basic tools for finding the error. The first term on the right-hand side of (4.9) represents the error resulting from using the Wiener filter (or the Kalman-Bucy filter) in the absence of sampling, which is a well known quantity. The second term on the right-hand side of (4.9) is dependent on whether n_2 is present or absent, and on the number of the sampled-data links that are available.

When there is only one sampled-data link available and $n_2=0$, we have seen from Theorem 4.3 that once the constant coder N is known, the joint optimization problem is reduced to the problem of designing the postfilter without considering the prefilter. Thus, our procedure is first to find the optimum coder using the state-space approach (Section 4.4), and then Eqs. (D.13) and (5.45) (where r is replaced by y , the output of the optimum prefilter; see Eq. (4.31)) are used to find the expression of the second term on the right-hand side of (4.9).

When there are n sampled-data links available, where n is the dimension of the state vector which characterizes the input signal, and $n_2=0$, an expression of the second term on the right-hand side of (4.9) can be found as follows. From the results obtained in Section 3.2, the signal observed at the receiving end is the $x_s(k)$'s (the sampled version of $x_s(t)$ which is the output of the continuous KSVF, and is specified by (4.15)). From (D.37), it follows that

$$E \{ [x_s(t) - \hat{x}_s(t)]^2 \} = \Phi(t-mT) V_{x_s} V_{x_s}'(m) \Phi'(t-mT) +$$

$$\int_{mT}^t \Phi(t-v) K_c V_u K_c' \Phi'(t-v) dv, m = \langle t/T \rangle. \quad (5.46)$$

Since the observation process is $x_s(kT)$, $V_{x_s}^v(m)=0$, and (5.46) becomes

$$E\{[x_s(t) - \hat{x}_s(t)]^2\} = \int_{mT}^t \Phi(t-v) K_C V_u K_C' \Phi'(t-v) dv. \quad (5.47)$$

where K_C is the steady-state continuous KSVF gain, and V_u is the innovation process of the primary observation process $r(t)$. Both K_C and V_u can be found by the known results of Kalman filtering. Finally, the steady-state error can be found by integrating both sides of (4.9) from $t=0$ to $t=T$, and then dividing by T on both sides of the resulting equation.

The steady-state error of the system with the jointly optimized unrealizable pre- and postfilters can be found from (5.25) (with $n_2=0$). Thus, we have outlined the general procedure used in calculating the steady-state error of the optimum systems. Error calculations for suboptimum systems with a ZOH (zero order hold) as the postfilter will be explained in the examples.

Before proceeding, let us define the following quantities as the normalized steady-state mean-square error between the desired signal and the output of the systems specified thereafter:

- EUNFG: with the jointly optimized unrealizable pre- and postfilters.
- EUNG: with the optimum unrealizable postfilter and without a prefilter.
- EG: with the optimum realizable postfilter and without a prefilter.
- EFG: with the jointly optimized realizable pre- and postfilters.
- EFM: with the jointly optimized pre- and postfilters, and with multiple sampled-data links.
- EFZ: with a ZOH (or the modified ZOH with a gain factor) as the postfilter, and with the optimum prefilter.
- EZ: with a ZOH as the postfilter and without a prefilter.
- EMZ: with the modified ZOH (with a gain factor) as the postfilter and without a prefilter.

and

- EFGQ: with the cascade of a continuous KSVF followed by a constant coder as the prefilter, and with the optimum postfilter in the case in which $n_2 \neq 0$.

Also define

$$ED = -EFG + EG$$

$$EDM = -EFM + EG$$

and

$$EDZ = EMZ - EFZ.$$

Furthermore, assume that σ is the variance of the desired signal. Define a given system as being asymptotically stable if the resulting mean-square error between the output of the system and the desired signal does not exceed σ as $T \rightarrow \infty$.

$n_2=0$ in all of the following subsections except 5.3.6 (refer to Fig. 1.1_c). By Theorem 4.3, the optimum realizable prefilter in the absence of n_2 is simply a cascade of a Vector Wiener filter (continuous KSVF) followed by a constant coder. It is well known that the function of the Wiener filter is to try to filter out the noise; this fact suggests that in the absence of n_2 , the essential function of the prefilter is to try to reduce the aliasing of the noise n_1 . The validity of this assumption will be investigated in the following subsections.

5.3.2 Case 1: A first-order Gauss-Markov process with an uncorrelated colored noise

The models of the signal and observation processes are given by the following equations:

$$\dot{x}(t) = -a x(t) + w(t), \quad (5.48)$$

$$d(t) = x(t), \quad (5.49)$$

$$\dot{n}_1(t) = -b n_1(t) + w_n(t), \quad (5.50)$$

and

$$r(t) = x(t) + n_1(t), \quad (5.51)$$

with prior statistics given by

$$R_w(t-\tau) = 2a \sigma \delta(t-\tau) \quad (5.52)$$

$$R_{w_n}(t-\tau) = 2b V_n \delta(t-\tau) \quad (5.53)$$

and

$$R_{ww_n}(t-\tau) = 0 \quad (5.54)$$

The performances of various systems for this example are shown in Figs. 5.1 and 5.3 through 5.9, and Tables 5.1 through 5.3.

In Fig. 5.1, we plot EFG and ED versus aT with $V_n/\sigma=1$, and with b varying from $5a$ to $14a$. As shown in Fig. 5.1, the error reduction achieved by the addition of the optimum prefilter in contrast to the case with the optimum postfilter only shows a peak value at a certain sampling period. This phenomenon can be explained as follows. Since the noise has a broader spectrum than the signal, the aliasing of the noise is expected to get worse

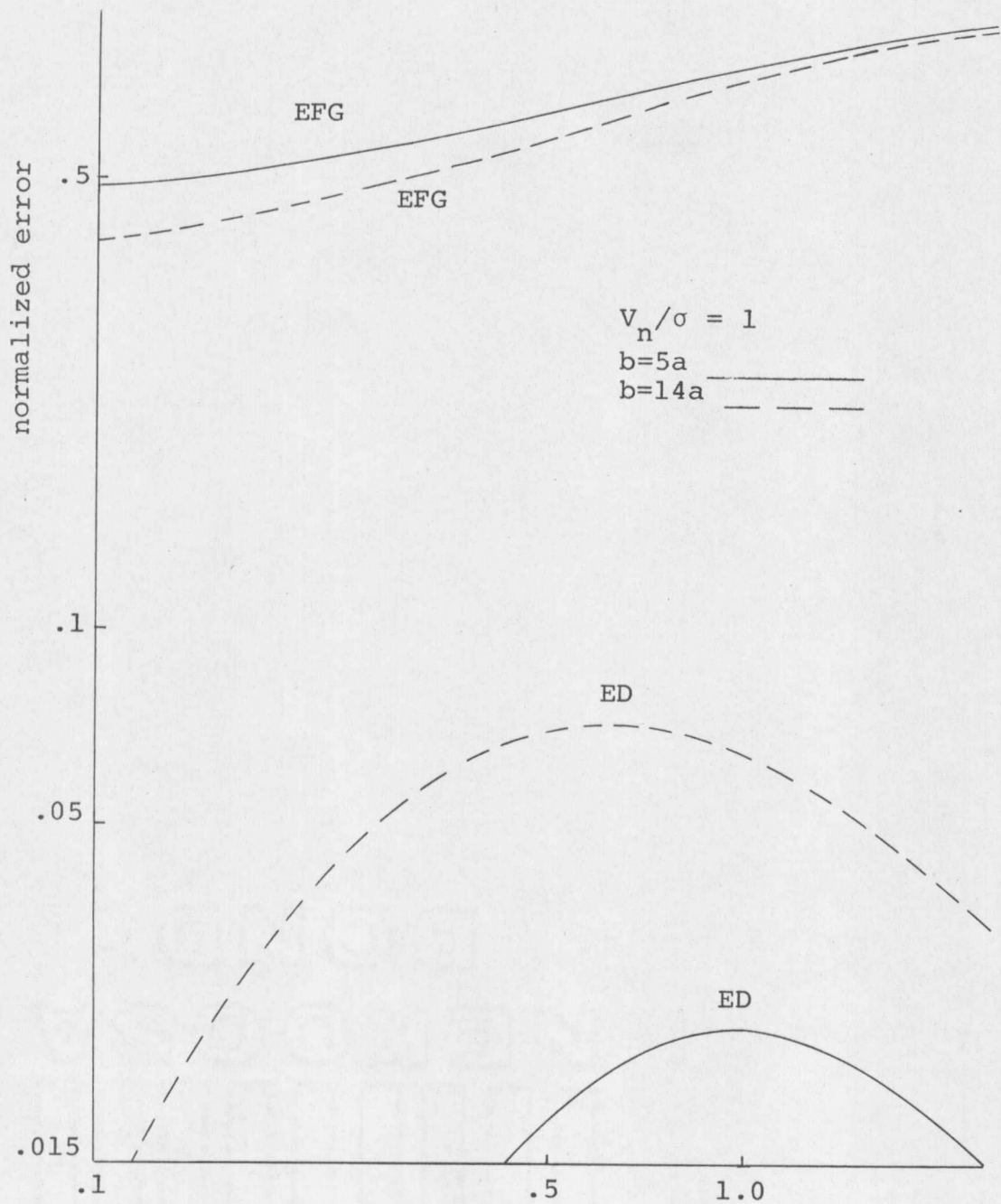


Figure 5.1 EFG and ED versus aT for Case 1.

Table 5.1 Normalized errors of the optimum systems with,
and without a prefilter for Case 1, $V_n/\sigma=1$.

b	aT	EFG	EG	ED	ED/EG %
14a	0.1	0.3953	0.4080	0.0127	3.19
	.2	.4501	.4858	.0357	7.35
	.3	.4983	.5521	.0538	9.75
	.4	.5407	.6053	.0645	10.65
	.5	.5782	.6479	.0696	10.75
	.6	.6115	.6828	.0713	10.45
	.7	.6409	.7129	.0709	10.20
	.8	.6672	.7365	.0693	9.96
	.9	.6906	.7576	.0670	8.84
	1.0	.7115	.7760	.0644	8.30
	2.	.8362	.8787	.0404	4.68
	3.	.8891	.9168	.0277	3.02
5a	0.1	.4807	.4822	.0014	0.30
	.2	.5277	.5328	.0050	.94
	.3	.5691	.5786	.0094	1.63
	.4	.6056	.6194	.0138	2.22
	.5	.6378	.6553	.0174	2.66
	.6	.6663	.6866	.0203	2.95
	.7	.6917	.7139	.0222	3.10
	.8	.7142	.7376	.0233	3.16
	.9	.7343	.7582	.0238	3.14
	1.0	.7523	.7762	.0239	3.09
	2.	.8594	.8767	.0173	1.97
	3.	.9047	.9168	.0121	1.32

at a sampling period smaller than the one where the aliasing of the signal becomes serious. To illustrate this point further, refer to Fig. 2, where it is shown that at a certain sampling period, the aliasing of the noise is rather severe (Fig. 5.2_b) but the aliasing of the signal is not (Fig. 5.2_d). The sampling periods at which the aliasing of the noise becomes serious whereas the aliasing of the signal is not are the periods for which the function of the realizable prefilter comes into play.

These are the sampling periods around the point where the peak error reduction occurs in Fig. 5.1. At other sampling periods, either the aliasing of both the signal and the noise are insignificant or the aliasing of the signal becomes so severe that nothing much can be done by the optimum realizable zero-lag prefilter. Also shown in Fig. 5.1 is the fact that when b increases, more error reduction is achieved by using the optimum prefilter.

This phenomenon is further illustrated in Fig. 5.3, where the error reduction efficiency ED/EG is plotted versus aT for various b ($V_n/\sigma=1$). It is seen from Fig. 5.3 that as b increases, ED/EG also increases. Since as b increases, the spectrum of the noise becomes broader and the aliasing

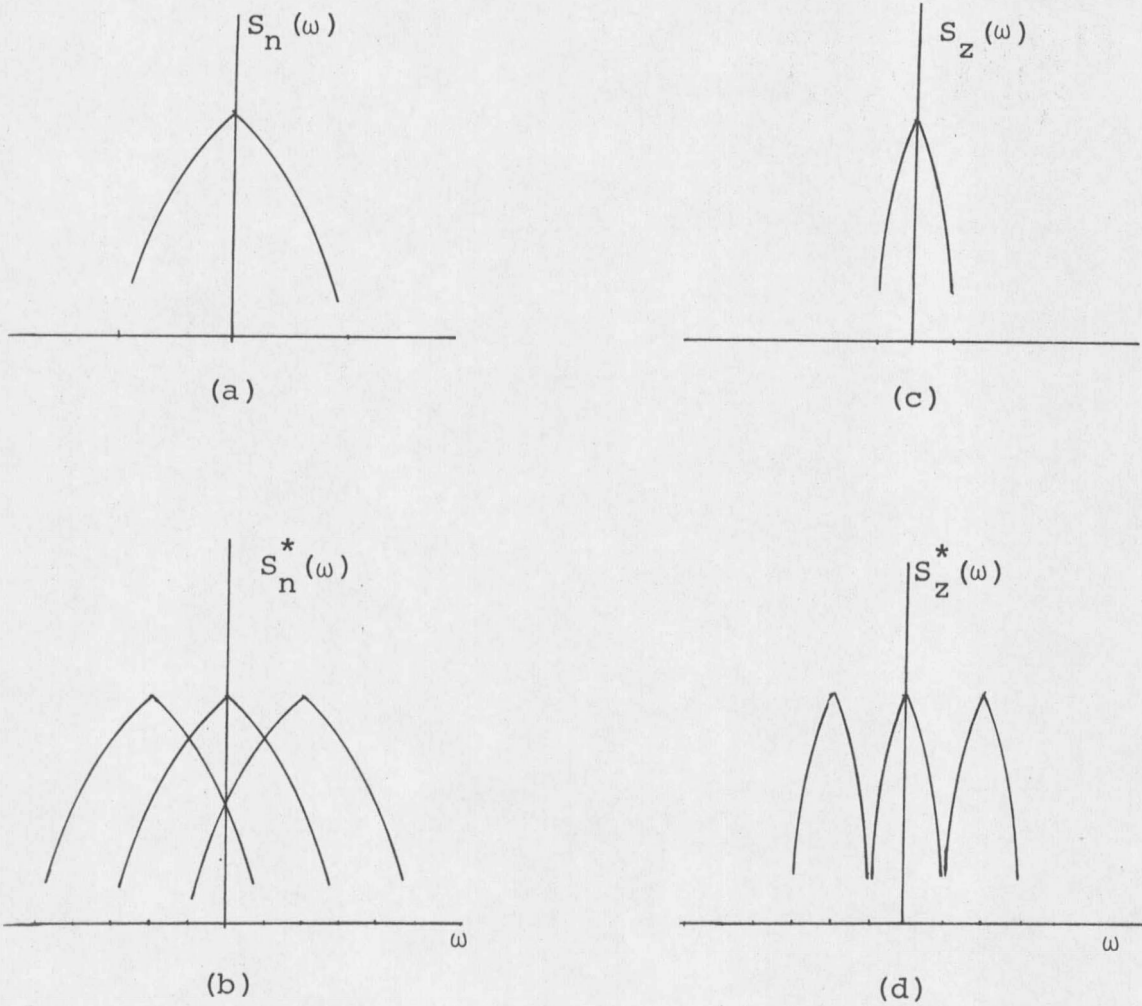


Figure 5.2 The aliasing of random processes with different spectra (after sampling).

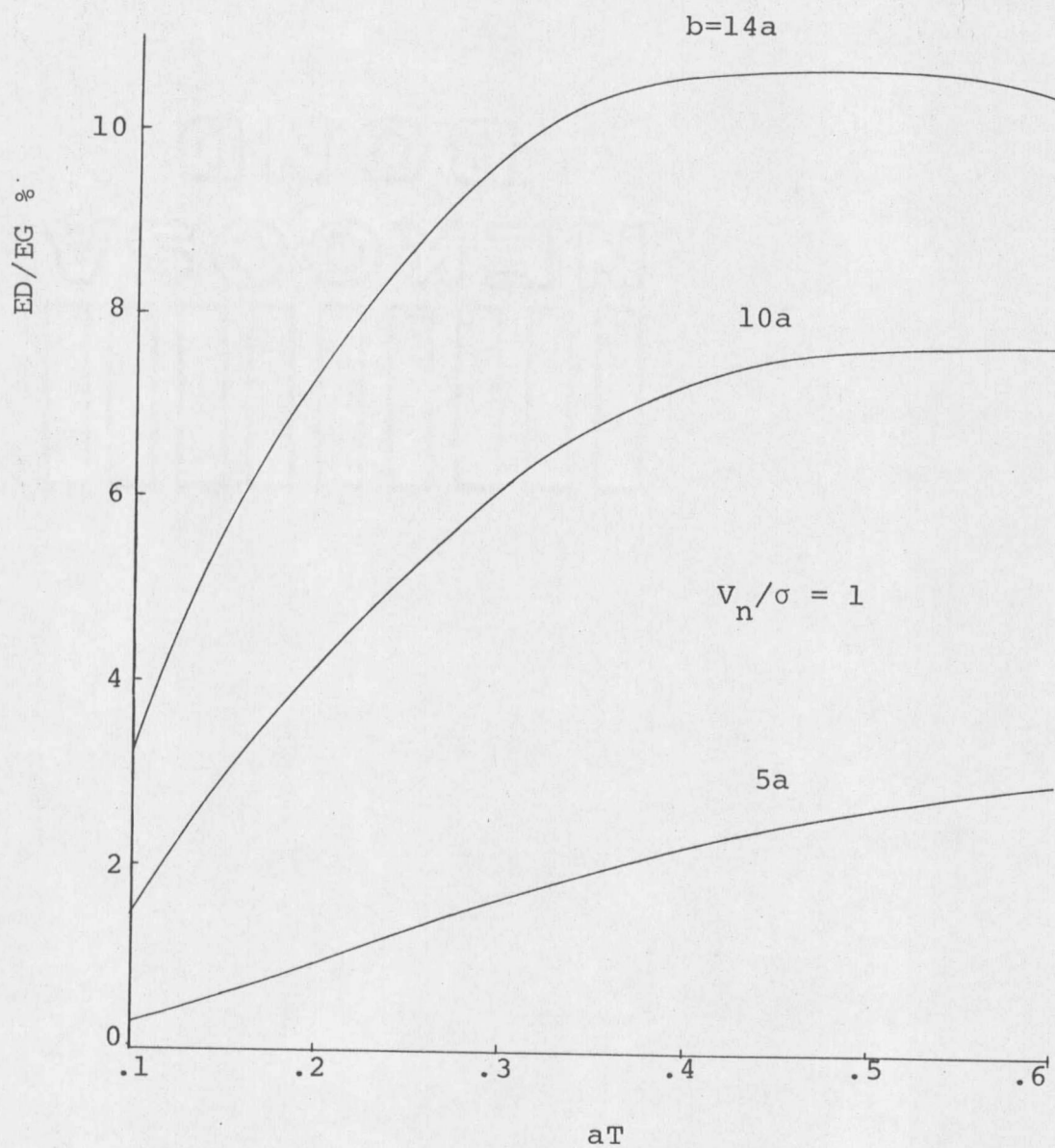


Figure 5.3 The error-reduction efficiency ED/EG versus aT for Case 1.

becomes worse, the results of Figs. 5.1 and 5.3 are as expected. Also for the same reason, as b increases, the sampling period at which the aliasing of the noise becomes serious gets smaller. Thus, the peak error reduction occurs at a smaller sampling period as b increases. This phenomenon is illustrated in Fig. 5.4. In Fig. 5.5, we plot ED/EG versus aT with V_n/σ varying from 0.1 to 1. We see from this figure that ED/EG increases with V_n/σ . Since increasing V_n simply means that the noise becomes more severe, this result is, again, as expected.

Thus, we have shown that the essential function of the optimum realizable zero-lag prefilter is to try to reduce the aliasing of the noise n_1 in this example. Similar results hold for higher-order signals, as we shall see later.

The inability of the optimum realizable zero-lag prefilter to deal with the aliasing of the signal can also be seen from the fact that in the absence of the noises n_1 and n_2 (refer to Fig. 1.1_C), the optimum realizable zero-lag prefilter for the case of a first-order Gauss-Markov process is simply a straight line; i.e., it simply does not do anything in this case. For higher-order

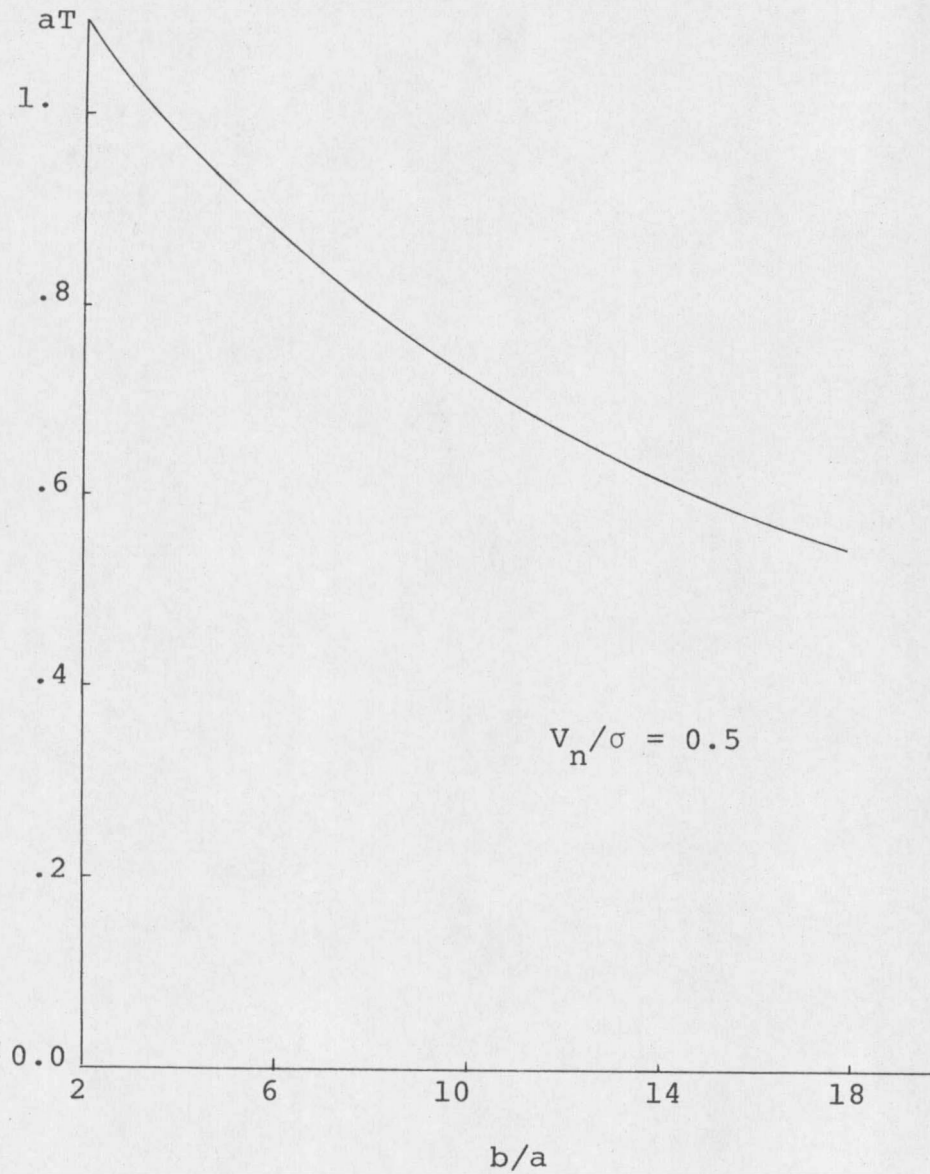


Figure 5.4 The sampling period at which peak error reduction occurs by using the optimum prefilter for Case 1.

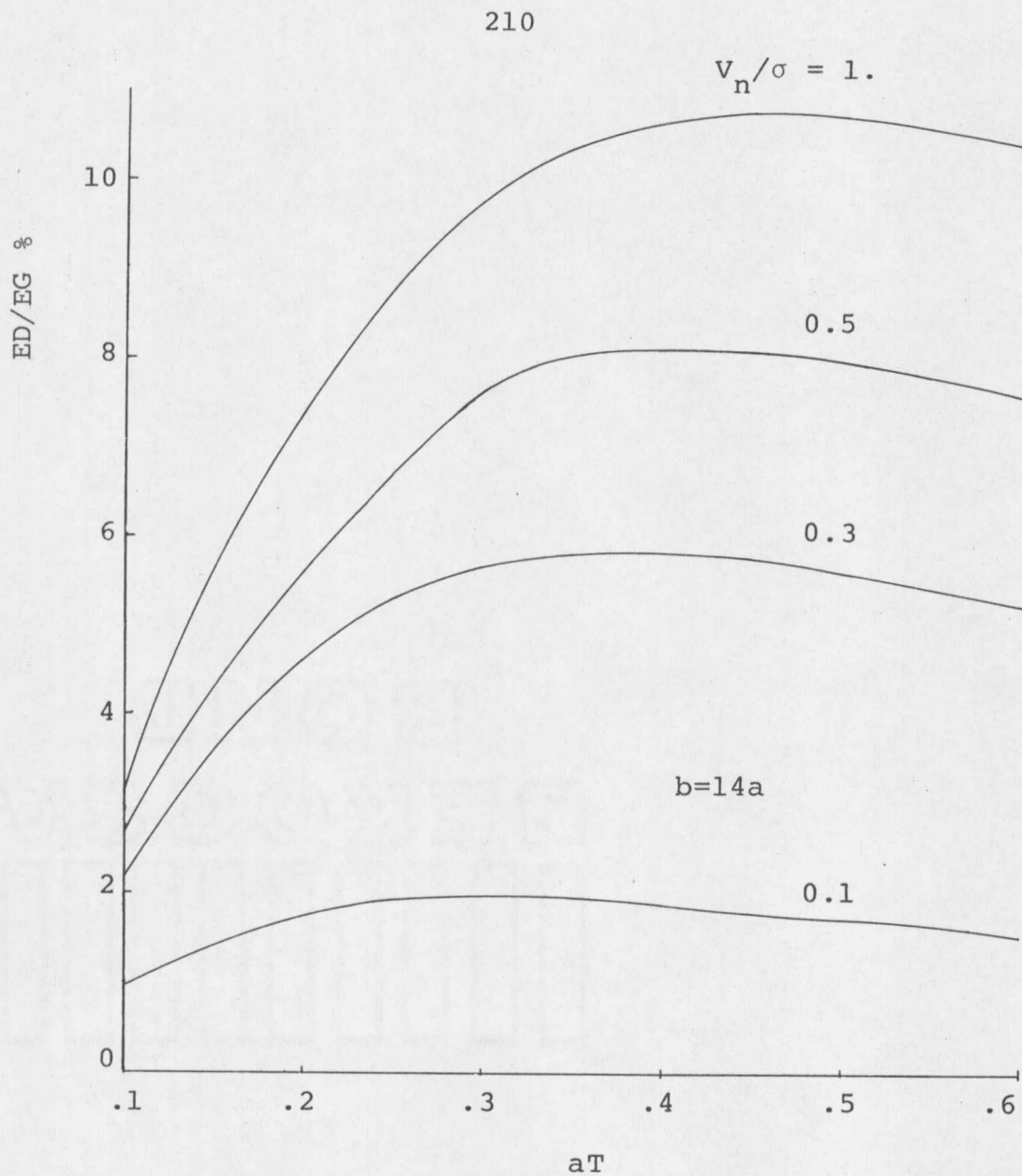


Figure 5.5 The error reduction efficiency ED/EG versus aT for Case 1.

signals, since the optimum system must estimate the states $x_i(t)$ first, as far as estimating one state component ($x_1(t)$ for example) is concerned, other components ($x_j(t)$ for $j \neq 1$) behave as if they are noises. Thus, the optimum realizable zero-lag prefilter can still do something even if the observation noises are absent in the case with higher-order signals. This point will be further illustrated in Section 5.3.4.

When the postfilter is the ZOH, the optimum prefilter can be found as follows. Since

$$g(t) = \begin{cases} 1 & \text{if } 0 \leq t < T, \text{ where } T \text{ is the sampling period} \\ 0 & \text{otherwise.} \end{cases} \quad (5.55)$$

this is just a special case of Example 3.1. Thus, the optimum zero-lag prefilter in this case is

$$f(t) = n_a \cdot f_{ks}(t), \quad (5.56)$$

where $f_{ks}(t)$ is the impulse of the Wiener filter, and n_a can be found by (3.90) to be (where $n_a = N_0$)

$$n_a = [1 - \exp(-aT)]/aT. \quad (5.57)$$

One can also use the ordinary calculus to find the coder n_a . From Example 3.1, the optimum prefilter in this

case is the cascade of continuous Wiener filter followed by a gain factor. Denoting this gain by k_a , we see that the error is given by

$$\begin{aligned} EFZ &= (1/T) \int_0^T E\{[d(t) - k_a d_s(0)]^2\} dt \\ &= \sigma - 2 k_a \cdot n_a \cdot R_{d_s}(0) + k_a^2 R_{d_s}(0), \end{aligned} \quad (5.58)$$

where n_a is given by (5.57), and $d_s(t)$ is the output of the Wiener filter. Differentiating EFZ with respect to k_a and equating the result to zero to obtain $k_a = n_a$, which is the result obtained by (3.90). The steady-state error then becomes

$$EFZ = \sigma - n_a^2 R_{d_s}(0), \quad (5.59)$$

where n_a is given by (5.57) and $d_s(t)$ is the output of the Wiener filter.

The steady-state error EZ (of the system with a ZOH as the postfilter and without a prefilter) is

$$\begin{aligned} EZ &= (1/T) \int_0^T E\{[d(t) - r(0)]^2\} dt \\ &= 2\sigma \cdot (1 - n_a) + V_n, \end{aligned} \quad (5.60)$$

where n_a is given by (5.57).

Plots of EFZ and EZ versus aT with $b=14a$, and with V_n/σ varying from 0.2 to 0.7 are shown in Fig. 5.6 with numerical data given in Table 5.2. Note that as T increases, EZ becomes larger than σ . Hence, the system with a ZOH as the postfilter and without a prefilter is not an asymptotically stable system (see the definition of an asymptotically stable system in Section 5.3.1). One can add a gain factor k_z to improve the performance of this system. The error with this modified ZOH is

$$\begin{aligned} \text{EMZ} &= (1/T) \int_0^T E \{ [d(t) - k_z r(0)]^2 \} dt \\ &= \sigma - 2 k_z n_a \sigma + k_z^2 (\sigma + V_n). \end{aligned} \quad (5.61)$$

It is easy to find that the optimum $k_z = n_a \sigma / (\sigma + V_n)$, and EMZ becomes

$$\text{EMZ} = [1 - n_a^2 \sigma / (\sigma + V_n)]. \quad (5.62)$$

Since the difference between the modified ZOH and the ZOH is simply a gain factor k_z , the optimum prefilter with the modified ZOH as the postfilter is simply $1/k_z$ of the one with the ZOH as the postfilter, and the combination of the pre- and postfilters are the same for the two systems (the system with the ZOH as the postfilter and with

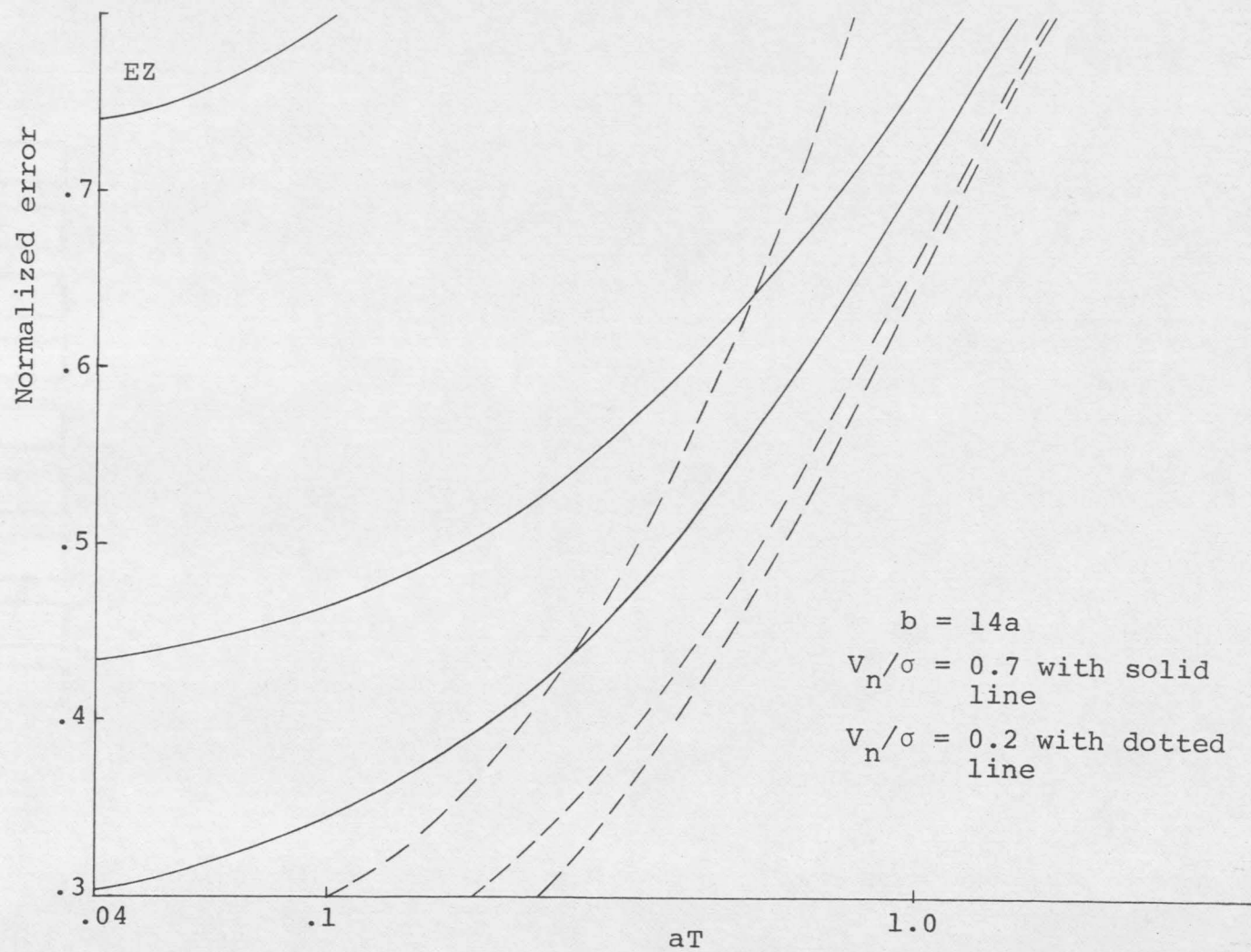


Figure 5.6 EFZ, EMZ, and EZ versus aT for Case 1.

Table 5.2 Normalized errors of the suboptimum systems with ZOH as the postfilter, and with, and without an optimum prefilter for Case 1, $b=14a$.

V_n/σ	aT	EFZ	EMZ	EDZ	EDZ/EMZ %	EZ
0.7	0.04	0.3045	0.4348	0.1302	29.94	0.7394
	.06	.3182	.4459	.1276	28.63	.7588
	.08	.3315	.4567	.1251	27.40	.7799
	.10	.3446	.4672	.1227	26.26	.7967
	.3	.4598	.5609	.1001	18.03	.9721
	.5	.5518	.6357	.0839	13.20	1.1261
	.7	.6257	.6957	.0700	10.07	
	.9	.6853	.7443	.0589	7.92	
	1.0	.7108	.7650	.0541	7.07	
	2.	.8647	.8900	.0253	2.84	
	3.	.9274	.9410	.0153	1.44	
.2	.04	.1606	.1992	.0386	19.40	.2395
	.06	.1707	.2150	.0378	17.62	.2588
	.08	.1932	.2303	.0371	16.13	.2779
	.10	.2089	.2453	.0364	14.84	.2967
	.3	.3480	.3780	.0300	7.94	.4721
	.5	.4582	.4839	.0249	5.14	.6261
	.7	.5482	.5680	.0207	3.65	.7617
	.9	.6202	.6377	.0174	2.74	.8813
	1.0	.6509	.6670	.0160	2.41	.9358
	2.	.8367	.8442	.0075	0.89	1.3353
	3.	.9124	.9163	.0040	0.44	

the optimum prefilter and the system with the modified ZOH as the postfilter and with the corresponding optimum prefilter). Plots of EMZ versus aT with $b=14a$ and with V_n/σ varying from 0.2 to 0.7 are also shown in Fig 5.6 with numerical data given in Table 5.2. Numerical data for $EDZ=EMZ-EFZ$ and the error reduction efficiency EDZ/EMZ versus aT are also given in Table 5.2. EDZ/EMZ versus aT is plotted in Fig. 5.7.

Interestingly, both EDZ and EDZ/EMZ are monotonically decreasing function of the sampling period T . Similar results are obtained for higher-order signals. This phenomenon can be explained as follows. The output of the modified ZOH in the absence of the prefilter is

$$\begin{aligned} d^0(t) &= n_a r(mT) \\ &= n_a [d(mT) + n_1(mT)] \end{aligned} \quad (5.63)$$

for $mT \leq t < (m+1)T$, where n_a is given by (5.57). Since n_a is a monotonically decreasing function of the sampling period T , $d^0(t)$ has a smaller weight on both the signal and noise as T increases. Thus, with the modified ZOH, the effect of the presence of the noise becomes less and less prominent as T increases. This can also be seen

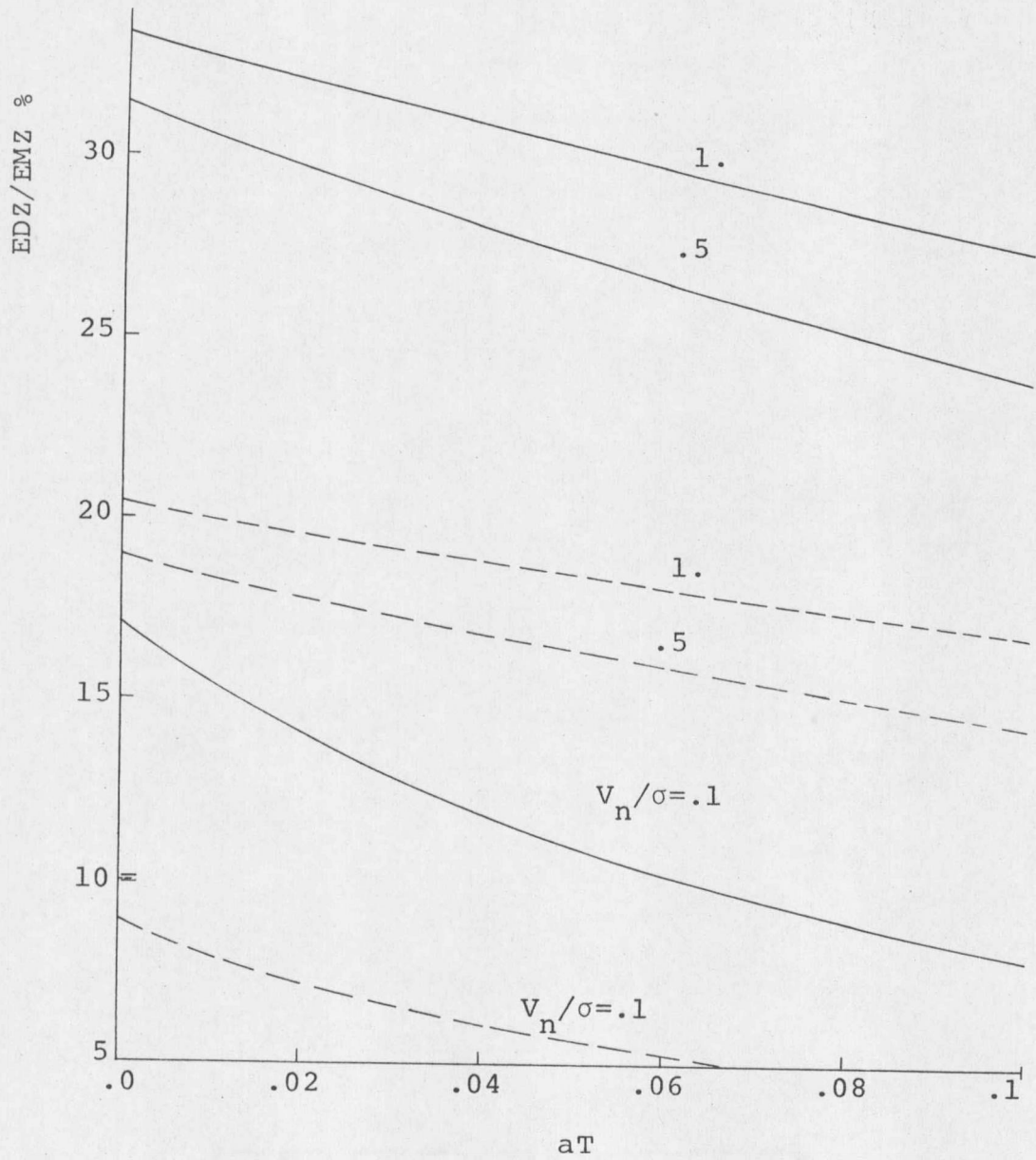


Figure 5.7 EDZ/EMZ versus aT for Case 1.

from (5.62), where EMZ depends on the factor $n_a^2 \sigma / (\sigma + V_n)$. As T increases, n_a decreases, and $n_a^2 \sigma / (\sigma + V_n)$ becomes less and less an indication of how severe the noise is. Recall that the essential function of the prefilter is to deal with the aliasing of the noise n_1 . If the noise becomes less significant, the function of the prefilter also becomes less prominent. In the limit when $n_1=0$, $EFZ=EMZ$ in this example. Since the noise appears to be less significant as T increases with the modified ZOH, EMZ approaches EFZ as T increases. This argument is similar to the one we made before about the function of the jointly optimized unrealizable pre- and postfilters where as T increases, the filters gradually ignore the effect of the noise n but concentrate their efforts in dealing with the aliasing of the signal.

A plot of EUNFG, together with its component ERRSA (the part of the error due to sampling) and ERRN (the part of the error due to noise n_1), versus aT with $V_n/\sigma=1$ and $b=14a$, is shown in Fig. 5.8. It can be seen from Fig. 5.8 that as T increases, ERRSA gradually comes to dominate the overall error.

Finally, in Fig. 5.9, we plot, from bottom to top, EUNFG, EUNG, EFG, EFZ, and EG versus aT with $V_n/\sigma=0.8$ and

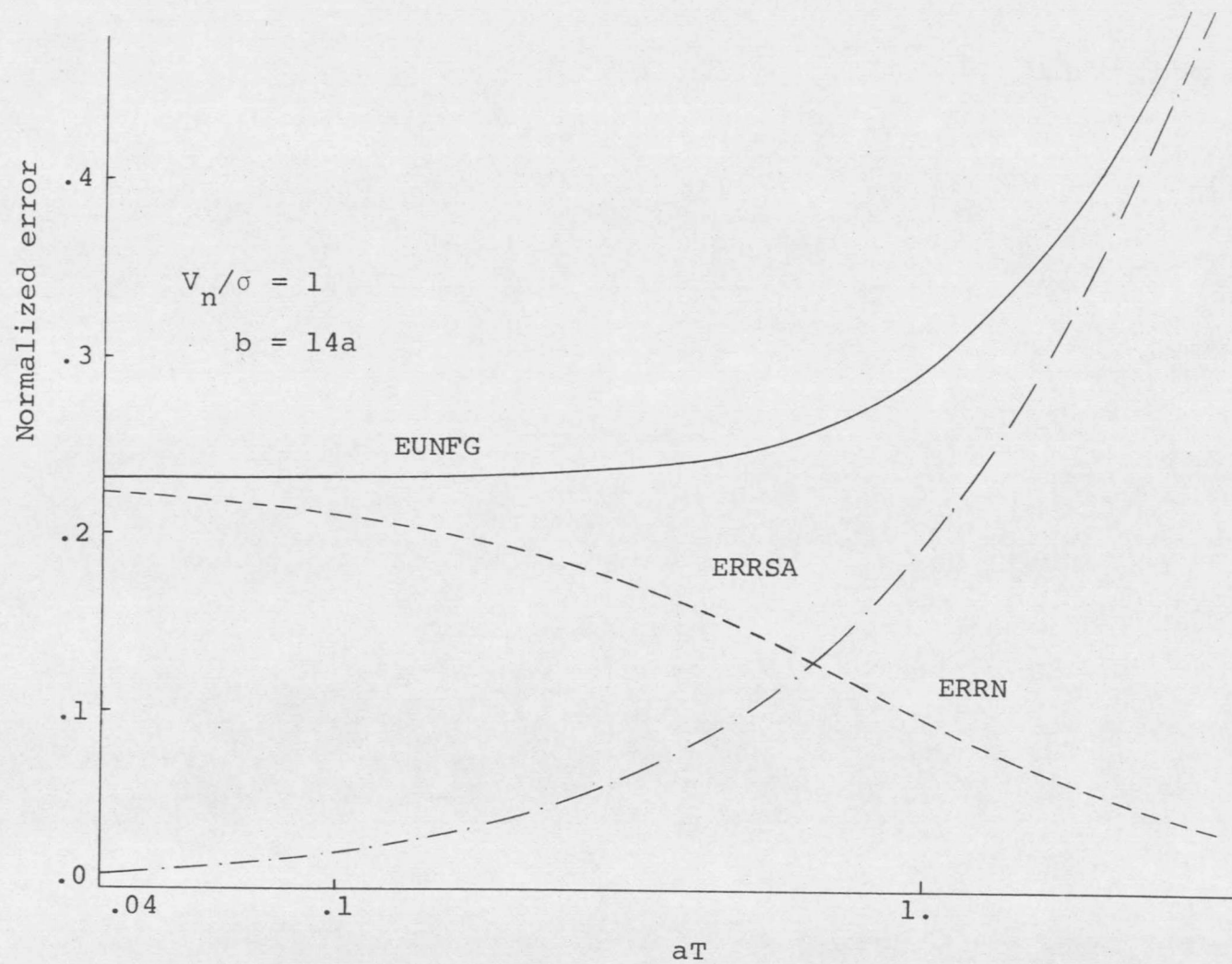


Figure 5.8 Normalized ERRSA, ERRN, and EUNFG versus aT for Case 1.

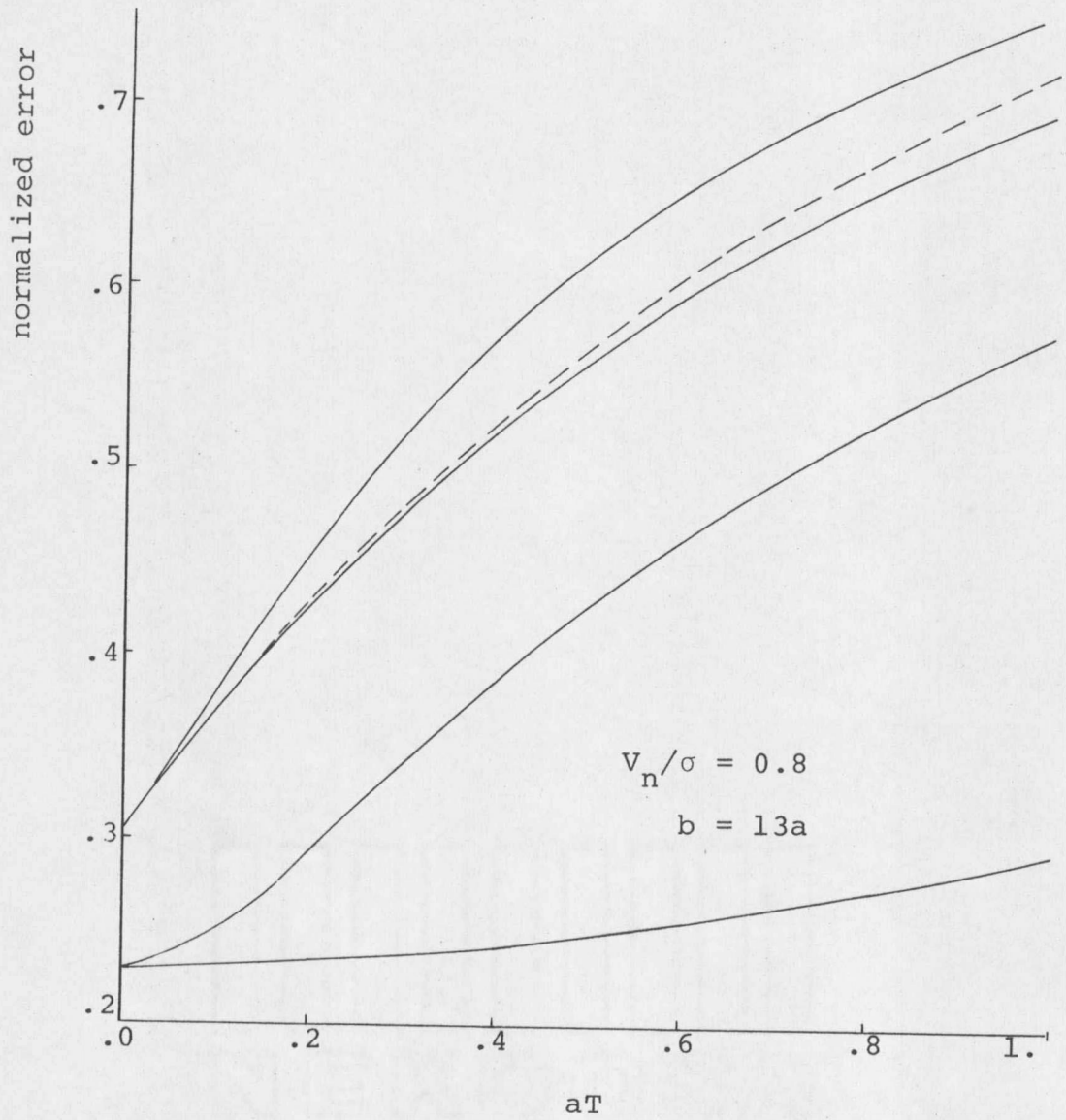


Figure 5.9 From bottom to top, EUNFG, EUNG, EFG, EFZ, and EG versus aT for Case 1.

$b=13a$ in order to compare the performances of various systems. Numerical data for Fig. 5.9 are given in Table 5.3. More interpretations will be given in Section 5.3.6.

5.3.3 Case 2: A second-order Butterworth process with an uncorrelated white noise

The models of the signal and noise, etc., are the same as those given in Example 4.2 with $\omega_c=1$,

$$R_w(t-\tau) = \begin{bmatrix} 0 & 0 \\ 0 & 200\sqrt{2} \end{bmatrix} \delta(t-\tau), \quad (5.64)$$

$$R_{n_1}(t-\tau) = 10 \delta(t-\tau), \quad (5.65)$$

and

$$R_{wn}(t-\tau) = 0. \quad (5.66)$$

Plots of EUNFG and EFG versus aT are shown in Fig. 5.10.

It is of interest to investigate how much error reduction is achieved by using multiple sampled-data links in contrast to a single sampled-data link in this example. The result is shown in Fig. 5.11, where it is seen that (EFG-EFM) is small compared with EFG (or EFM); i.e., the error reduction achieved is poor. Similar

Table 5.3 Performances of various systems for Case 1, $V_n/\sigma = 0.8$
and $b=1.3a$.

aT	EUNFG	EUNG	EFG	EFZ	EG	EMZ	EZ
0.02	0.1982	0.2313	0.2735	0.2906	0.2739	0.3874	0.7199
.04	.1986	.2342	.2878	.3046	.2893	.3994	.7394
.06	.1992	.2384	.3017	.3182	.3049	.4112	.7588
.08	.1996	.2438	.3152	.3316	.3207	.4227	.7779
.10	.2002	.2504	.3283	.3446	.3366	.4340	.7967
.2	.2032	.2931	.3892	.4055	.4128	.4865	.8873
.3	.2074	.3411	.4427	.4598	.4790	.5335	.9721
.4	.2131	.3865	.4899	.5084	.5338	.5754	1.0516
.5	.2202	.4271	.5316	.5518	.5793	.6130	
.6	.2288	.4632	.5685	.5907	.6167	.6466	
.7	.2385	.4956	.6012	.6257	.6502	.6768	
.8	.2493	.5247	.6303	.6571	.6783	.7039	
.9	.2609	.5513	.6564	.6853	.7028	.7283	
1.0	.2731	.5755	.6796	.7108	.7242	.7503	
2.	.4033	.7362	.8181	.8647	.8462	.8832	
3.	.5142	.8166	.8766	.2974	.8961	.9373	

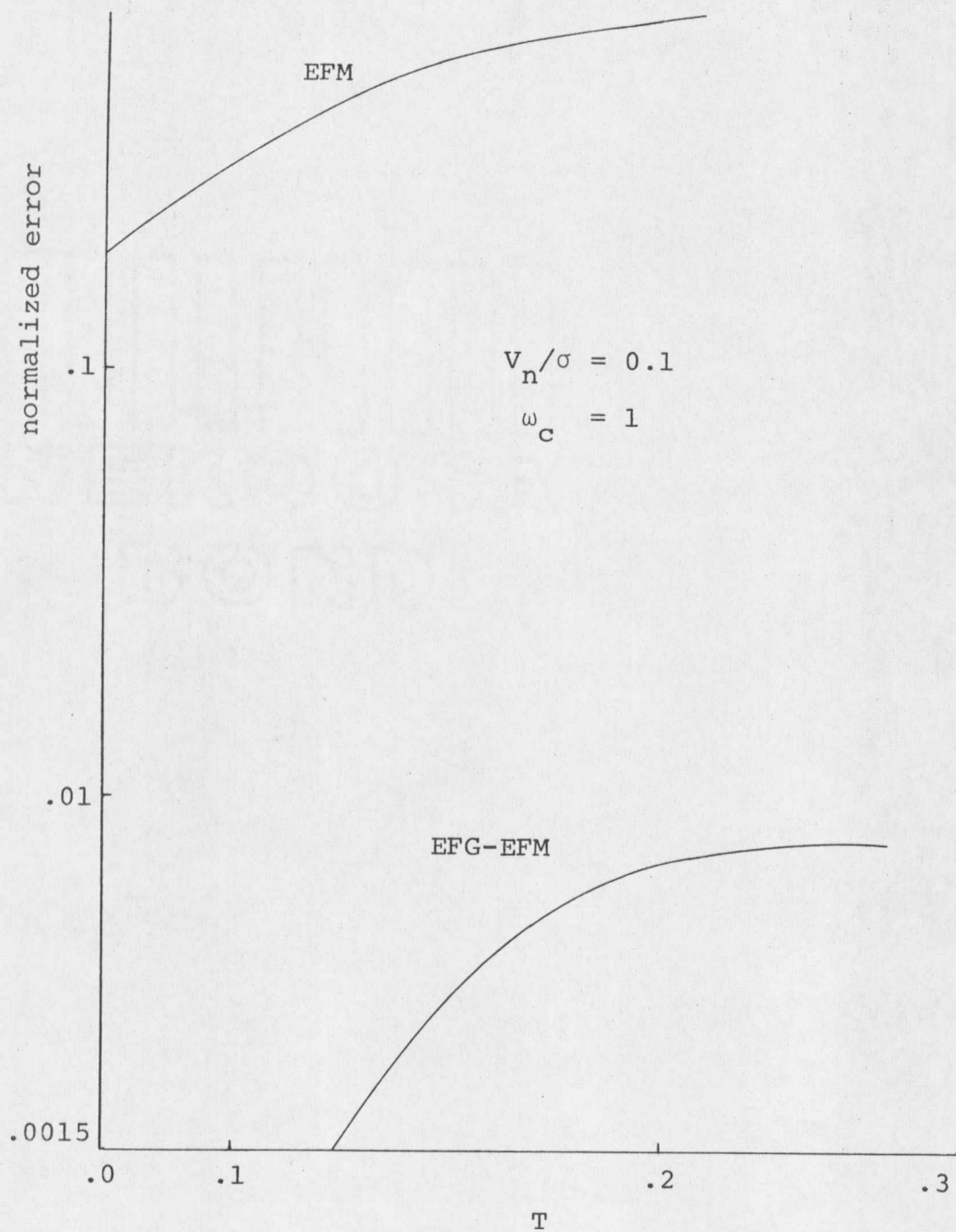


Figure 5.10 EFM, and (EFG - EFM) versus T for Case 2.

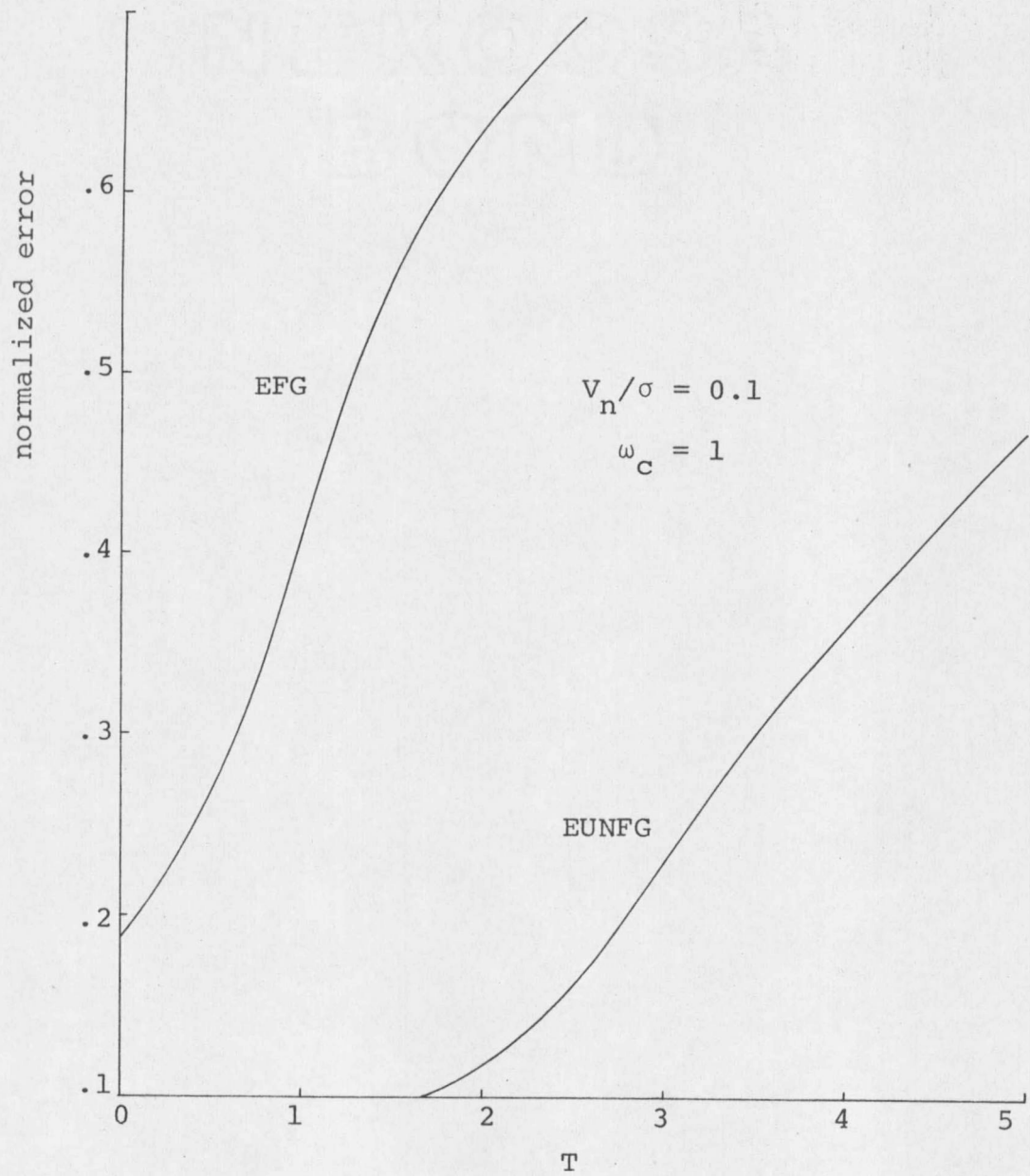


Figure 5.11 EUNFG, and EFG versus T for Case 2.

results holds for a second-order Gauss-Markov process either with or without noise. Whether the same situation exists for higher-order signals and for more complicated noises remains to be checked. Since there is no coder included in the optimum prefilter with multiple sampled-data links when $n_2=0$, it is in general much easier to investigate the performance of this system than that of the system with a single sampled-data link. The investigation of the performance of the system with multiple sampled-data links gives a clear indication of how well the performance of the system with a single sampled-data link will be. Thus, most of the insight into the prefiltering problem with a single sampled-data link can be obtained by looking at the system with multiple sampled-data links. We will take this shortcut in the following subsections.

5.3.4 Case 3: A second-order Gauss-Markov process without noise

The signal and observation models are given in Example 4.1 with $n_1=0$, $b=10a$, and $\sigma_b=0.6\sigma_a$.

Since this is a second-order system, as far as estimating one of the state components ($x_1(t)$ for example) is concerned, the other component ($x_2(t)$) behaves like noise.

Thus, the optimum realizable prefilter can still do something in this case (in the absence of n_1), as follows from the discussion in Section 5.3.2.

Plots of EUNFG, EUNG, EFM, and EG (dotted line) versus aT are shown in Fig. 5.12, with numerical data given in Table 5.4. It is seen that the error reduction achieved by adding the optimum realizable zero-lag prefilter with multiple sampled-data links in contrast to the system with optimum realizable zero-lag postfilter only and with single sampled-data link is rather small in this case. Also, since (EFM-EG) is small, it is expected that (EFG-EFM) can not be large.

5.3.5 Case 4: Second-order Gauss-Markov processes with uncorrelated colored noises

The signal and the observation models are the same as those given in Example 4.1 except the noise is now specified by

$$\dot{n}_1(t) = -d_n n_1(t) + w_n(t) \quad (5.67)$$

with

$$R_{w_n}(t-\tau) = 2 d_n V_n \delta(t-\tau) \quad (5.68)$$

and

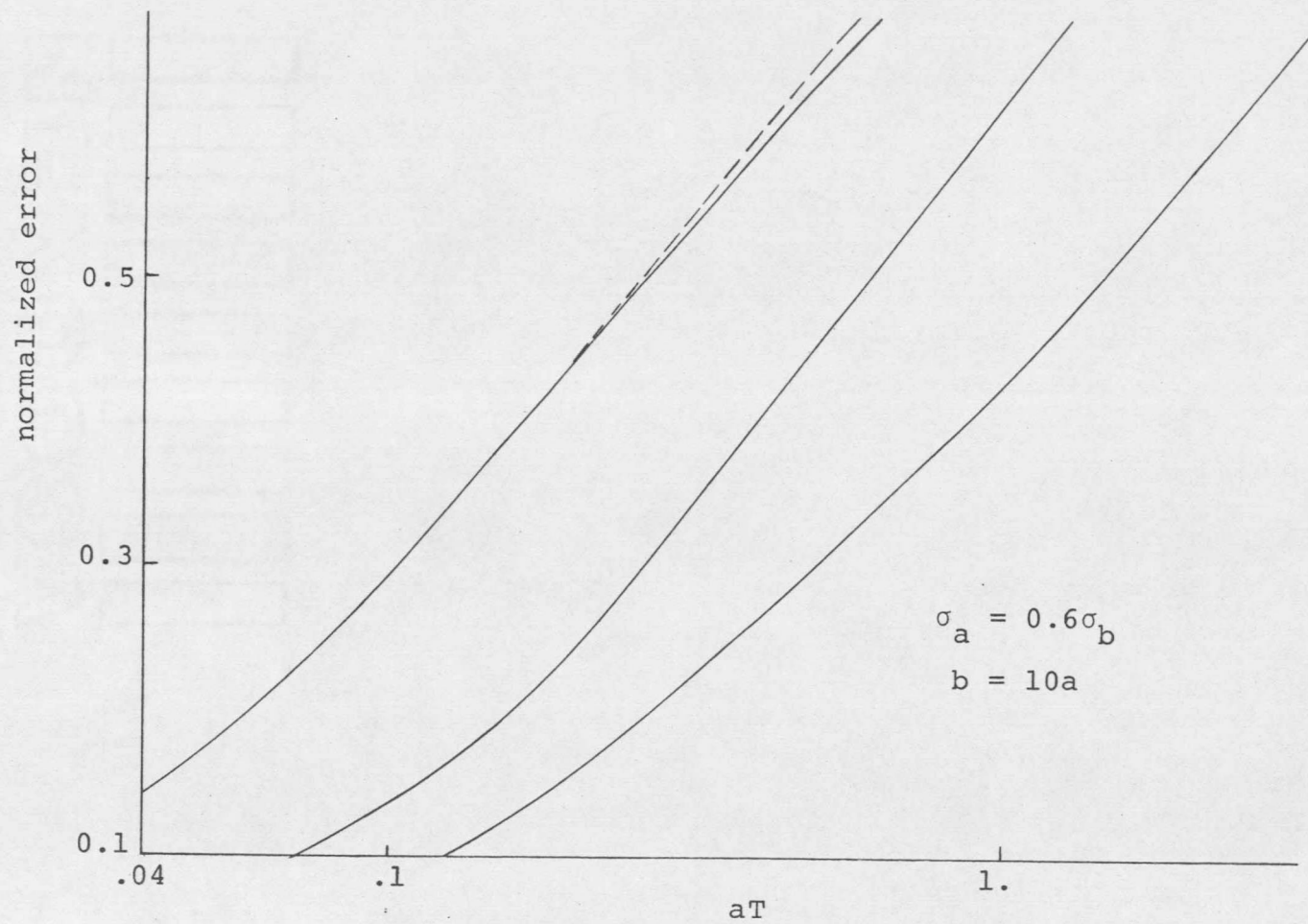


Figure 5.12 From right to left, EUNFG, EUNG, EFM, and EG (dotted line) versus aT for Case 3.

Table 5.4 Performances of various systems for Case 3,
 $\sigma_a = 0.6\sigma_b$ and $b=10a$.

aT	EUNFG	EUNG	EFM	EG
0.02	0.0177	0.0238	0.0798	0.0799
.04	.0353	.0612	.1466	.1467
.06	.0526	.0875	.2031	.2032
.08	.0697	.1144	.2514	.2517
.10	.0862	.1391	.2931	.2936
.2	.1606	.2403	.4318	.4412
.3	.2211	.3349	.5257	.5327
.4	.2664	.3986	.5861	.5968
.5	.3039	.4519	.6316	.6449
.6	.3349	.4906	.6677	.6828
.7	.3615	.5312	.6975	.7134
.8	.3849	.5594	.7227	.7389
.9	.4058	.5900	.7443	.7606
1.0	.4249	.6113	.7631	.7790
2.	.5634	.7556	.8683	.8791
3.	.6534	.8281	.9110	.9185

$$R_{w_n w}(t-\tau) = 0 \quad (5.69)$$

Plots of EFM and EDM are shown in Fig. 13, and plot of EDM/EG is shown in Fig. 5.14, with numerical data given in Table 5.5. It is seen from these figures that EDM and EDM/EG again exhibit peak values at a certain sampling period. Plots of EUNFG, EFM, and EG versus αT are shown in Fig. 5.15, with numerical data given in Table 5.6.

Similar to the case of a first-order Gauss-Markov process with uncorrelated colored noise, if a ZOH is used as the postfilter, the corresponding optimum prefilter can be found to be

$$f(t) = [n_a, n_b] F_{ks}(t),$$

where $F_{ks}(t)$ is the impulse response of the continuous KSVF, n_a is given by (5.57), and

$$n_b = [1 - \exp(-bT)]/bT.$$

The optimal coder $N=[n_a, n_b]$ can either be found from (3.90) or by ordinary calculus as we did in Section 5.3.2. The optimum gain factor of the modified ZOH can be found

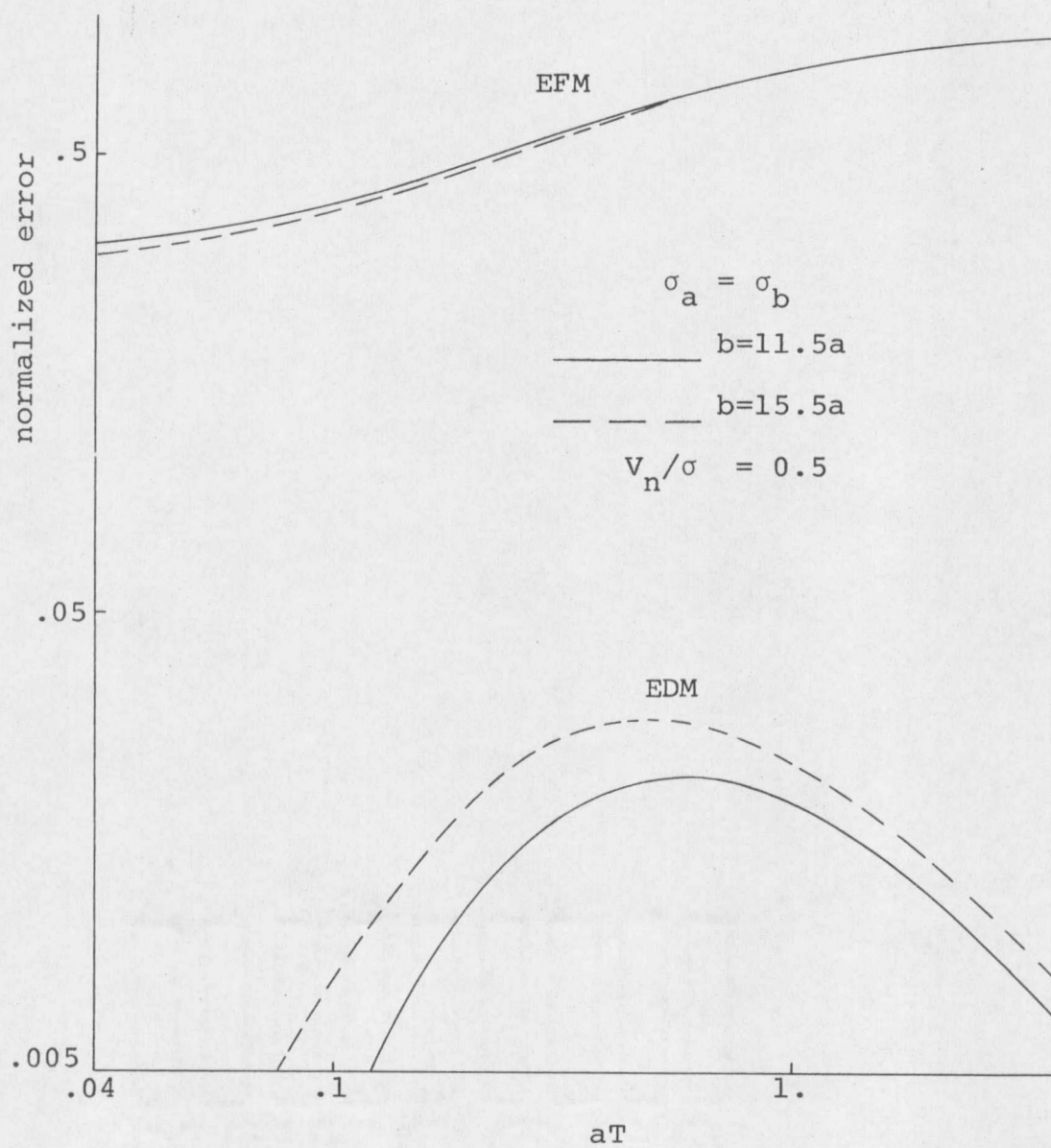


Figure 5.13 EFM and EDM versus aT for Case 4.

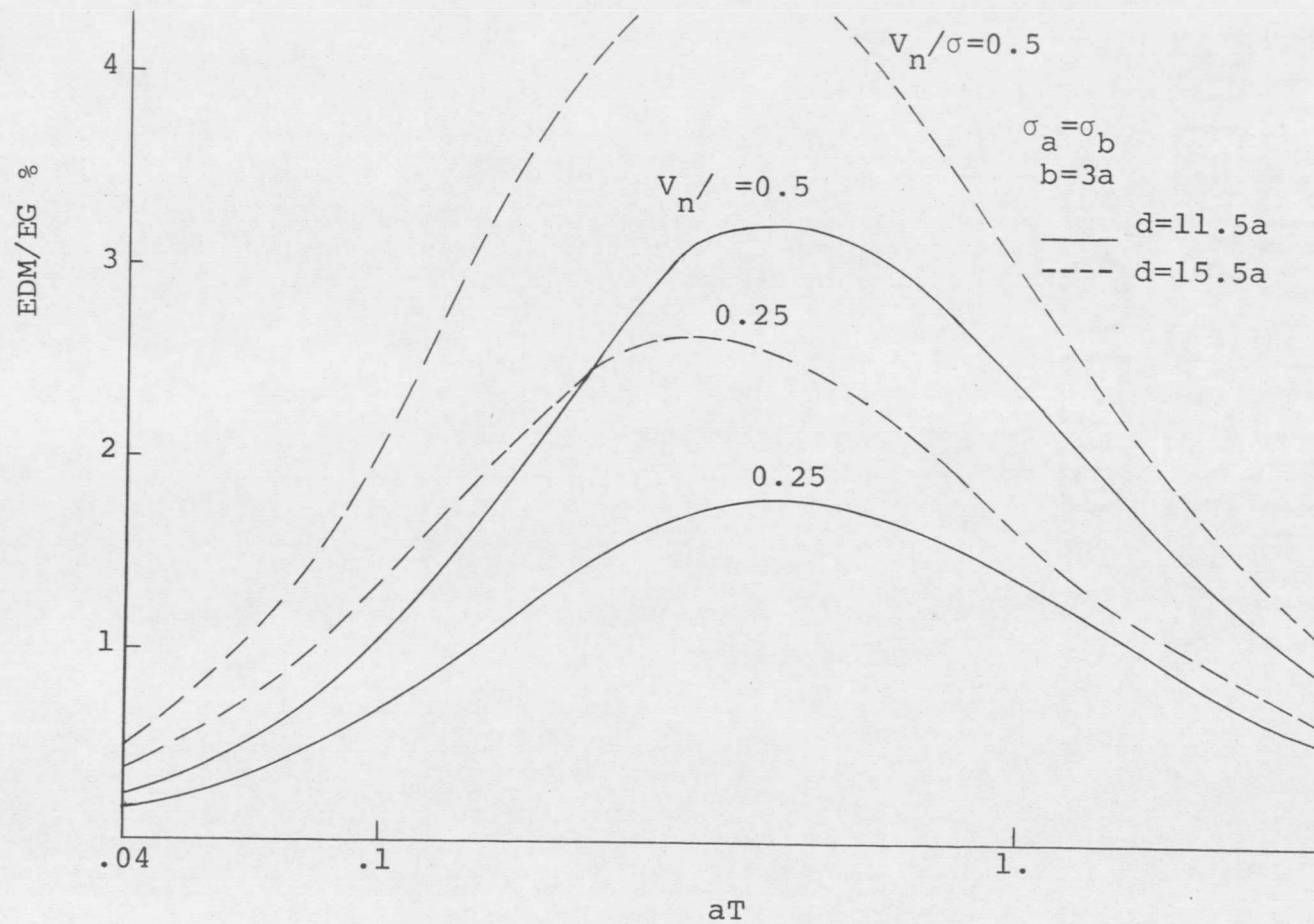


Figure 5.14 The error-reduction efficiency EDM/EG versus aT for Case 4.

Table 5.5 Normalized errors of the optimum systems with, and without a prefilter for Case 4, $V_n/\sigma=0.5$, $\sigma_a=\sigma_b$, and $b=3a$.

d	aT	EFM	EG	EDM	EDM/EG %
11.5a	0.04	0.3306	0.3314	0.0008	0.24
	.06	.3550	.3567	.0017	0.48
	.08	.3779	.3808	.0029	0.76
	.10	.3996	.4039	.0043	1.06
	.3	.5636	.5810	.0174	3.10
	.5	.6650	.6668	.0218	3.18
	.7	.7317	.7535	.0218	2.89
	.9	.7783	.7986	.0203	2.55
	1.0	.7964	.8159	.0185	2.39
	2.	.8911	.9033	.0122	1.35
	3.	.9268	.9351	.0083	0.89
15.5a	.04	.3152	.3168	.0016	0.50
	.06	.3400	.3434	.0034	0.99
	.08	.3635	.3691	.0056	1.52
	.10	.3857	.3937	.0080	2.04
	.3	.5535	.5798	.0263	4.54
	.5	.6571	.6866	.0295	4.30
	.7	.7255	.7535	.0280	3.72
	.9	.7730	.7985	.0255	3.20
	1.0	.7916	.8159	.0243	2.97
	2.	.8885	.9032	.0147	1.63
	3.	.9251	.9352	.0101	1.08

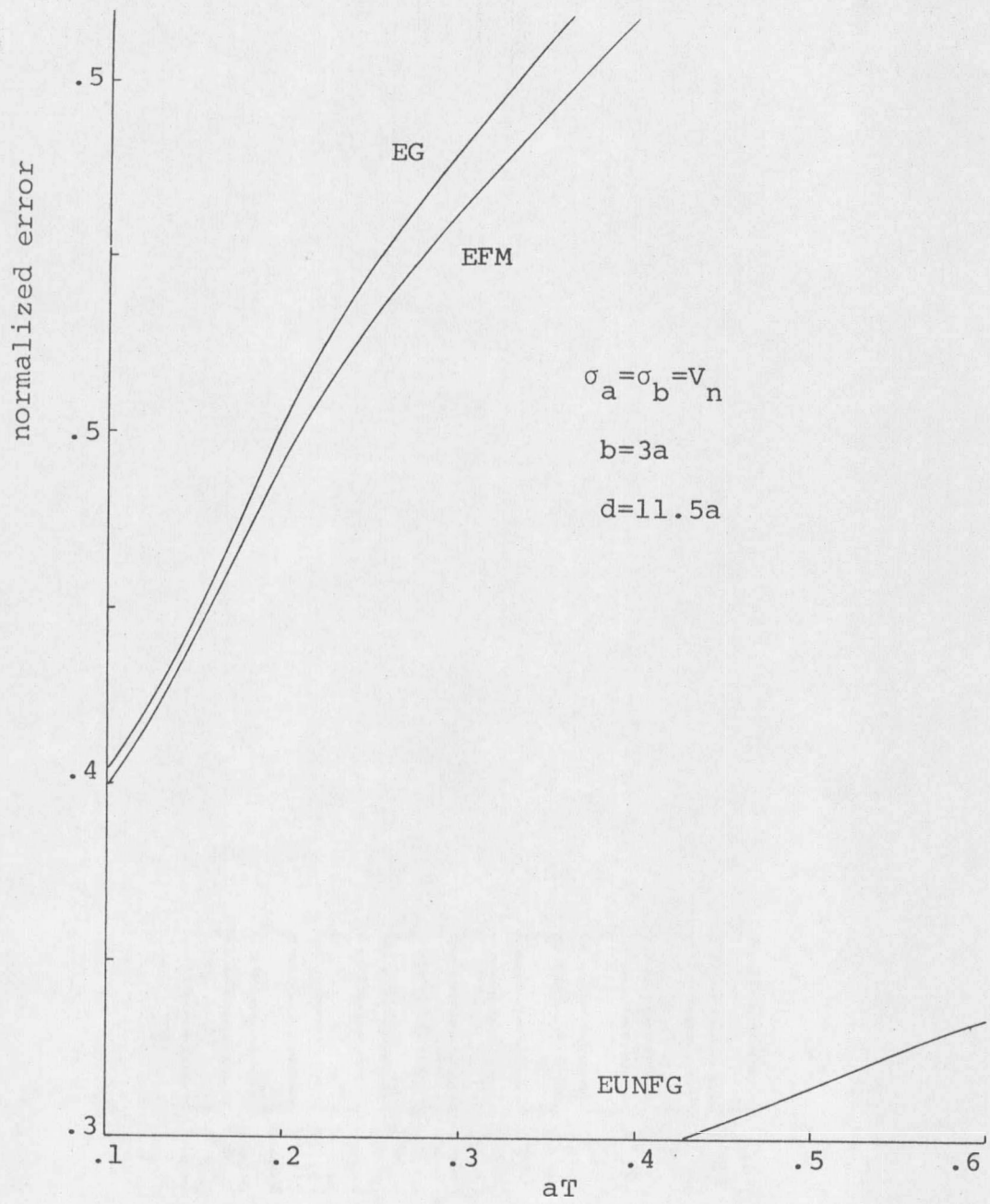


Figure 5.15 EUNFG, EFM, and EG versus aT for Case 4.

Table 5.6 Performances of various sytems for Case 4, $\sigma_a = \sigma_b$,
 $V_n/\sigma = 0.5$, $b = 3a$.

d	aT	EUNFG	EFM	EG	EFZ	EMZ	EZ
11.5a	0.04	0.2444	0.3306	0.3314	0.3309	0.3839	0.5775
	.06	.2466	.3550	.3567	.3556	.4073	.6142
	.08	.2487	.3779	.3808	.3791	.4295	.6499
	.10	.2510	.3996	.4039	.4014	.4506	.6844
	.2	.2633	.4920	.5038	.4978	.5416	.8416
	.3	.2780	.5636	.5810	.5740	.6132	.9766
	.4	.2951	.6198	.6404	.6350	.6702	1.0394
	.5	.3140	.6650	.6668	.6844	.7162	
	.6	.3340	.7015	.7237	.7248	.7536	
	.7	.3543	.7317	.7535	.7583	.7845	
	.8	.3746	.7569	.7781	.8863	.8101	
	.9	.3846	.7783	.7986	.8099	.8317	
	1.0	.4139	.7964	.8159	.8299	.8499	
	2.	.5662	.8911	.9033	.9311	.9403	
	3.	.6625	.9268	.9351	.9646	.9695	
	4.	.7267	.9451	.9514	.9820	.9820	

to be equal to $[(n_a \sigma_a + n_b \sigma_b) / (\sigma_a + \sigma_b + V_n)]$. The expression of EFZ, EMZ, and EZ are given by the following equations:

$$EFZ = \sigma_a + \sigma_b - [n_a, n_b] R_{x_s}(0) [n_a, n_b]',$$

$$EMZ = \sigma_a + \sigma_b - [(n_a \sigma_a + n_b \sigma_b)^2 / (\sigma_a + \sigma_b + V_n)],$$

and

$$EZ = 2[\sigma_a(1-n_a) + \sigma_b(1-n_b)] + V_n.$$

Plots of EFM, EFZ, EMZ, and EZ versus aT for $b=3a$, $d=11a$, and $\sigma_a = \sigma_b = V_n$ are shown in Fig. 5.16, with numerical data given in Table 5.7. It is seen from Fig. 5.16 that EFZ is very close to EFG in this case.

The error-reduction efficiency EDZ/EMZ is plotted in Fig. 5.17, with part of the numerical data given in Table 5.7. A look at Figs. 5.17 and 5.7 in comparison to Figs. 5.14 and 5.5 indicates that by adding a prefilter, greater advantage can be obtained when the postfilter is suboptimum than in the case when the postfilter is optimum. Note that the scales used in drawing these figures are different from one another.

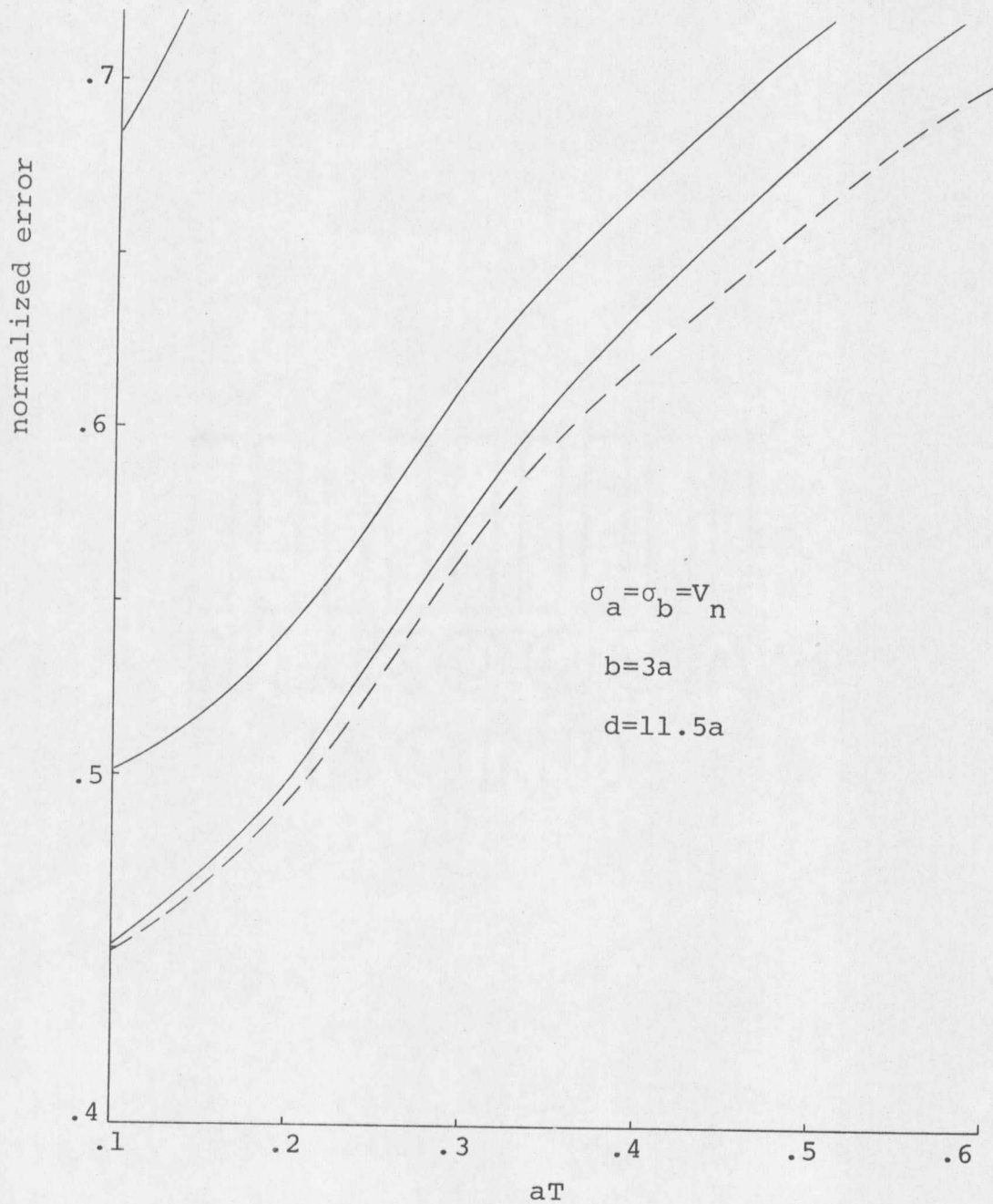


Figure 5.16 EFM, EFZ, EMZ, and EZ versus aT for Case 4.

Table 5.7 Normalized errors of the suboptimum systems with ZOH as postfilter, and with, and without an optimum prefilter for Case 4, $\sigma_a = \sigma_b$, $b=3a$, and $V_n/\sigma=0.5$.

d	aT	EFZ	EMZ	EDZ	EDZ/EMZ %	EZ
11.5a	0.04	0.3309	0.3839	0.0530	13.80	0.5774
	.06	.3556	.4073	.0517	12.68	.6142
	.08	.3791	.4295	.0504	11.74	.6499
	.10	.4014	.4506	.0492	10.92	.6844
	.3	.5740	.6132	.0392	6.39	.9766
	.5	.6844	.7162	.0318	4.42	1.1951
	.7	.7583	.7845	.0262	3.34	
	.9	.8099	.8317	.0218	2.62	
	1.0	.8299	.8499	.0200	2.35	
	2.	.9311	.9403	.0092	0.97	
	3.	.9646	.9695	.0049	0.51	
15.5a	.04	.3156	.3839	.0683	17.80	.5774
	.06	.3408	.4073	.0665	16.32	.6142
	.08	.3647	.4295	.0647	15.07	.6499
	.10	.3675	.4506	.0631	13.98	.6844
	.3	.5641	.6132	.0491	8.01	.9766
	.5	.6770	.7162	.0392	5.47	1.1951
	.7	.7526	.7845	.0319	4.06	
	.9	.8054	.8317	.0263	3.16	
	1.0	.8259	.8499	.0250	2.82	
	2.	.9299	.9403	.0104	1.15	
	3.	.9637	.9695	.0058	0.59	

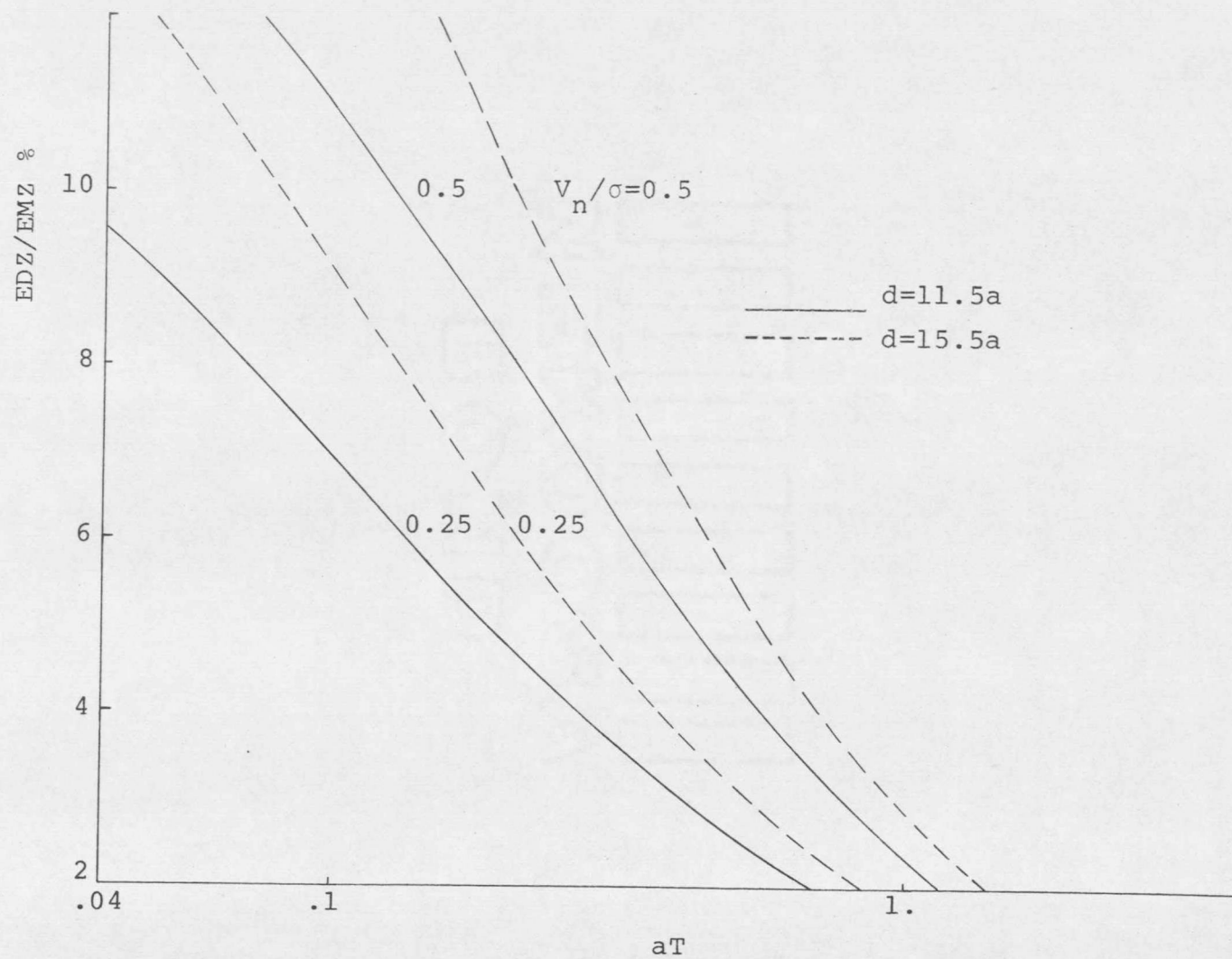
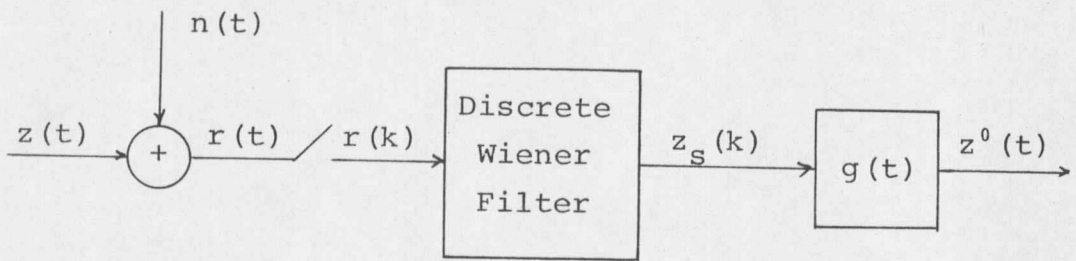


Figure 5.17 The error reduction efficiency EDZ/EMZ versus aT for Case 4.

5.3.6 Conclusion

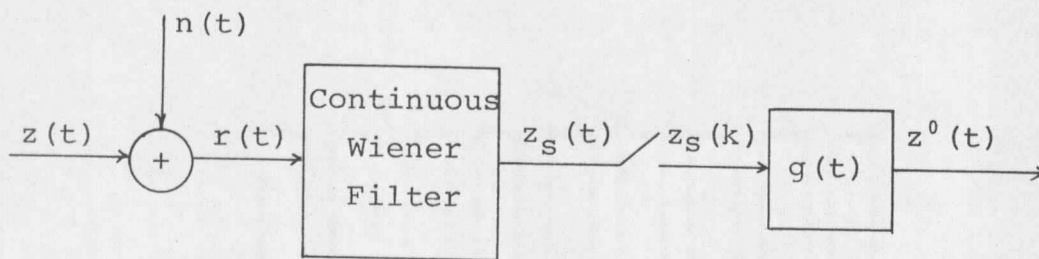
It has been shown that in the case in which $n_2=0$, the primary function of the realizable prefilter is to try to filter out the effect of the primary observation noise n_1 before it is sampled; i.e., to try to reduce the aliasing of n_1 . On the one hand, it is rather disappointing to find out that the error reduction achieved by introducing this optimum prefilter is not as much as we expected. On the other hand, although introducing a prefilter in the system complicates the synthesis process, it does not necessarily complicate the implementation of the overall system. An example is used to illustrate this point. Assume that one tries to reconstruct a first-order Gauss-Markov process $d(t)$ with autocorrelation equal to $\exp(-at)$. $d(t)$ is contaminated by noise $n_1(t)$ at the sending end. The optimum system with postfilter only is shown in Fig. 5.18_a where the postfilter is the cascade of a discrete Wiener filter followed by a filter characterized by its impulse response

$$g(t) = \begin{cases} \exp(-at) & \text{if } 0 \leq t < T \\ 0 & \text{otherwise.} \end{cases} \quad (5.70)$$



(a)

$$g(t) = \begin{cases} \exp(-at) & \text{if } 0 \leq t \leq T, \\ 0 & \text{otherwise.} \end{cases}$$



(b)

Figure 5.18 The optimum systems: (a) without a prefilter and (b) with a prefilter, where $z(t)$ is a first-order G.M. process.

The optimum system with pre- and postfilters is shown in Fig. 5.18_b, where we have a continuous Wiener filter as the optimum prefilter and $g(t)$ specified by (5.70) as the impulse response of the optimum postfilter. Thus, we have seen that in the case with prefilter, all one needs to do is to shift the discrete Wiener filter to the other side and convert it to become a continuous Wiener filter. Similar results hold for higher-order signals; i.e., the addition of an optimum prefilter simplifies the structure of the optimum postfilter. Whether this complicates or simplifies the implementation of the overall system depends on the practical situations.

Thus, when the noise n_1 is severe, or when high accuracy is required, it is probably worthwhile to include a prefilter in the system.

We have also seen that the efficiency of reducing the error by the addition of a prefilter is more prominent in the case with a suboptimum postfilter than in the case with the optimum postfilter. Since suboptimum postfilters are assumed to be fixed, the addition of a prefilter complicates the implementation in this case. However, the expressions of the optimal prefilter (either the impulse

response in the time domain or the transfer function in the frequency domain) can be obtained easily in this case; i.e., the synthesis process is straightforward.

From the results we obtained in Chapter 2, and those of Tretter and Steiglitz (57), for most practical situations, the performance of an unrealizable postfilter can be approximately achieved by allowing a finite delay before an estimation is made. Another interesting fact seen from Figure 5.9 by comparing the curves of EUNG with EFG is that in general, to allow a delay in the estimation process is often more worthwhile than to insert a prefilter in the system if the delay does not cause any serious problems.

The prefilter structure of the cascade of a continuous KSVF followed by a constant coder provides a sub-optimum system when quantization noise or/and channel additive noise are present; i.e., when $n_2 \neq 0$. We will look at this problem briefly in the following example.

Assume that the signal and the primary observation noise n_1 are the same as those given in Section 5.3.2. Now, noise n_2 has an autocorrelation function

$$R_{n_2}(t-\tau) = V_{n_2} \delta(t-\tau)$$

With $V_{n_2}/\sigma=0.02$, $V_{n_1}/\sigma=1$ ($\sigma=\sigma_a+\sigma_b$), $b=15a$, plots of EFGQ and EG versus aT are shown in Fig. 5.19. Since EFGQ represents the steady-state error resulting from using this suboptimum system with a prefilter whereas EG represents the error resulting from using the optimum system without prefilter, it is not surprising to see that in some regions of the sampling periods EFGQ is larger than EG.

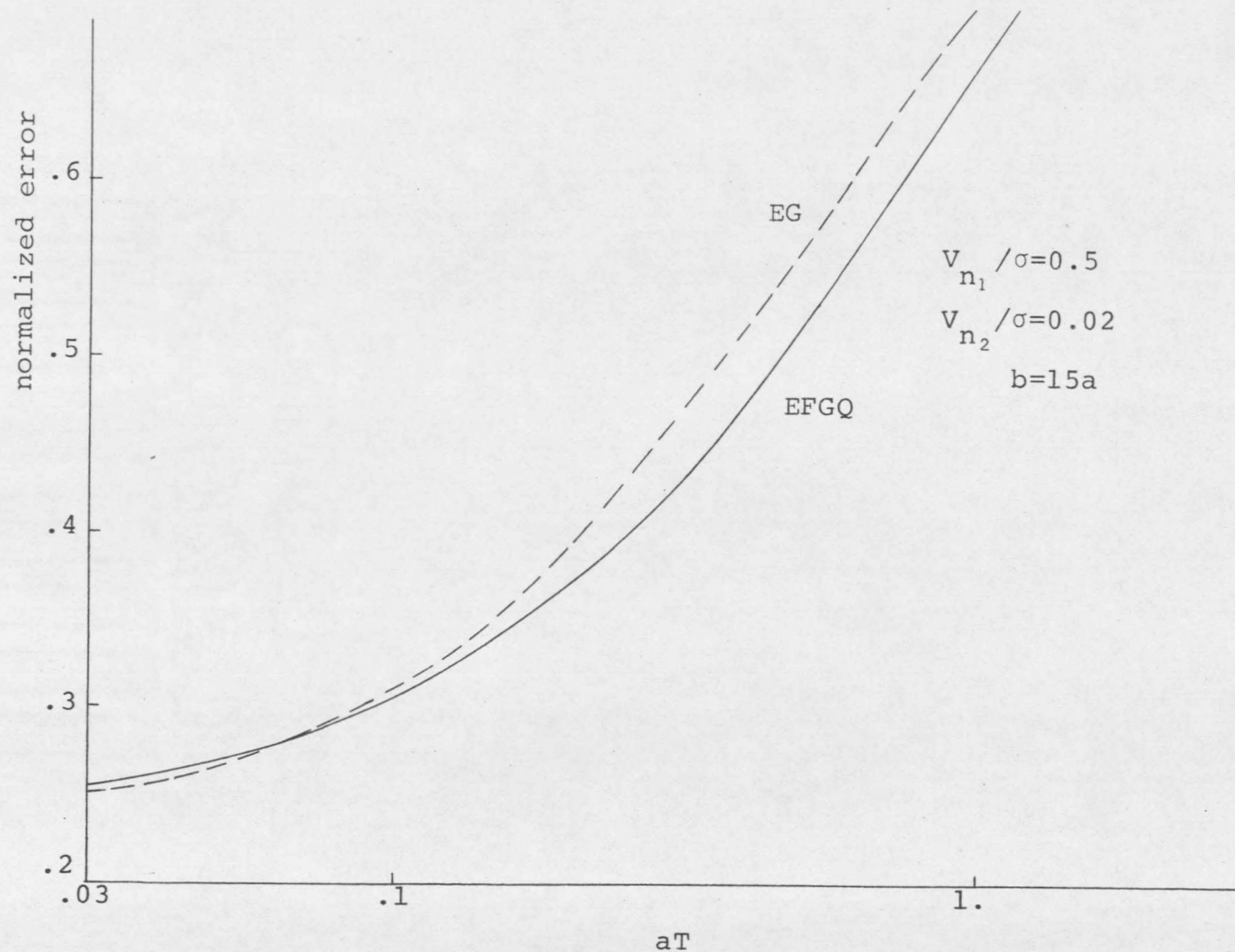


Figure 5.19 EFGQ and EG versus aT , where the signal and n_1 are the same as those given in Case 1, and n_2 represents the quantization noise.

Chapter 6

CONCLUSION, AND SUGGESTIONS FOR FURTHER WORK

The design of the optimum pre- and postfilters for the system shown in Fig. 1.1 has been solved under various assumptions. The theoretical results are presented in Chapters 2 through 4, and numerical results are presented in Chapters 2 and 5.

This work results from the fusion of three threads: the work of Robbins (49), Chang (9), Chan and Donaldson (7), Kellog (32) and others on the design of optimum pre- and postfilters in a sampled-data system; the work of Kalman (28), Kalman and Bucy (29), Petersen (41) and others on the state-space approach to the linear estimation problem; and the work of Bode and Shannon (4), Kailath (23) and others on the innovations approach to the filtering problem.

We bring together results in both the continuous and the discrete linear estimation of random functions obtained using the state-space and the innovations approach. The heart of the state-space approach is the model, or state equations, which characterize the input signal and the noises. In continuous estimation, there

are continuous state equations, and in discrete estimation there are discrete state equations. On the other hand, we have a system which is partly continuous and partly discrete. Thus, in order to use the state-space approach, one must have both continuous and discrete state equations which characterize the input signal and the discretized version of the same signal, respectively. Since the continuous state equation is assumed given in our work, discretization of the continuous state equation, which is present in Appendix C (22), is the first step in the derivation of the optimum postfilter. For the remainder of the derivation, we have repeatedly used the orthogonality principle and the innovations concept. The result shows a close relationship between the optimum postfilter we derived and both the continuous and the discrete KSVF (Kalman State Vector Filter).

We have obtained zero-lag, as well as finite-lag, optimum postfilters in Chapter 2, where the amount of delay is constrained to be an integer multiple of the sampling period T . The error reduction factor $ED(t,i)$ and the delay-efficiency factor $EN(t,i)$ have been defined as functions of the current time t , and the amount of delay

iT . ED is closely related to the "reconstruction efficiency density" defined by Tretter and Steiglitz (57), but ED has broader applicability (to nonstationary processes, time-variant systems, and nonsteady-state conditions, etc.), and ED is expressed explicitly in terms of the delay time iT .

In the steady state, numerical results obtained in Chapter 2 show that

1. By using the optimum infinite-lag postfilter in contrast to the optimum zero-lag postfilter, one can obtain from moderate to substantial error reduction depending on the inputs. For the examples worked out in Chapter 2, the range of error reduction varies from approximately 20% to 50% of the error of the system with the optimum zero-lag postfilter.

2. In general, the more severe the noise is, the greater the delay required for the performance of the optimum finite-lag postfilter to approximate that of the optimum infinite-lag postfilter. That is, error reduction converges more slowly with stronger noise.

3. The error reduction due to the first sample of delay contributes a major portion of the overall error reduction in general.

4. For the case of a first-order Gauss-Markov process with discrete white noise, it is found that the overall error reduction is essentially achieved within five samples of delay. A semi-log plot of the steady-state delay-efficiency factor $ENS(i)$ versus the delay time iT is a straight line in this case

In Chapter 3 and Appendix E, the design of the optimum prefilter with a given postfilter is studied. Unlike the design of the postfilter, where one can use the orthogonality principle to simplify the derivation, here we have to start from the integral equation obtained by the calculus of variations. But the innovations concept is still applicable.

It is not hard to imagine that there must exist some relationship between the optimum realizable prefilter and the continuous KSVF (Vector Wiener filter). We first notice that if there is no constraint on the number of the sampled-data links that are available, and quantization noise and channel additive noise are absent, then the optimum zero-lag prefilter is the continuous KSVF. In this case, the whole system acts like fixed-interval prediction system (with interval equal to the sampling

period) in the absence of the sampling process. In the case in which there is only one sampled-data link available, the continuous KSVF is still part of the optimum realizable zero-lag prefilter as shown by Theorem 3.2.

In Chapter 4, the joint optimization problem is studied. We mentioned in Chapter 4 that if one insists on using a prefilter whose impulse response is continuous in time, then the condition that $\xi_j = 0$ for $j \neq 0$ must be satisfied by the postfilter in order for the prefilter to be optimum (ξ_j is given by (3.55)). Together with the necessary conditions that must be satisfied by the pre- and postfilters (Eqs. (3.53) and (4.39)) in order for them to be optimum, this in effect yielded three equations that the pre- and postfilters must satisfy in order for them to be optimum (with the impulse response of the prefilter being continuous in time). It is well known that when there are more equations than unknowns, these equations must be consistent with one another. It has been shown by Theorem 4.3 that in the case in which $n_2 = 0$ (refer to Fig. 1.1_C), the condition that $\xi_j = 0$ for $j \neq 0$ is consistent with the necessary conditions for the pre- and postfilters to be optimum. When $n_2 \neq 0$, the condition that

$\xi_j = 0$ for $j \neq 0$ will not be satisfied by the postfilter in general. In this case, it has been shown in Theorem 4.2 and Fig. 4.1 that as far as filtering (zero-lag) is concerned, a more general approach to this problem is to sample $x_s(t)$ (the output of the continuous KSVF) first, and then combine the $x_s(k)$'s linearly using the coder N (see Eq. (4.13)) before quantization and transmission. Note that this is different from having multiple sampled-data links since we still have only one channel. Also notice that N is not a constant matrix in general with $n_2 \neq 0$, but instead, N can be considered as a discrete filter with memory (again, see Eq. (4.13)). The system shown in Fig. 4.1 has three filters. The first one is the continuous KSVF which is a well known filter. The second one (the coder) and the third one (the postfilter) try to yield a linear least square estimate of $x_s(t)$ subject to the conditions shown in Fig. 4.1. The degenerated system shown in Fig. 4.1_b is similar to the system considered by Hayes (20, also see Fig. 1.2_h), but not quite the same since we now have multiple inputs to the coder.

If the coder can be represented by a dynamic system; i.e., if the output of the coder can be characterized by

state equations, then the design of the postfilter can be solved by the method developed in Chapter 2. In the case in which $n_2=0$, Theorem 4.3 shows that the optimum coder becomes a constant matrix. The steady-state coder can then be found by nonlinear programming. The functions appearing in the nonlinear programming formulation are polynomials.

Finally, with some help from the frequency-domain approach, we have the following conclusions about the performance of various systems in the steady state as drawn from the numerical examples worked out in Chapter 5:

1. The primary function of the optimum zero-lag realizable prefilter in the absence of quantization noise or/and channel additive noises is to combat the aliasing of the primary observation noise n_1 .

2. On the one hand, when $n_2=0$, the error reduction achieved by the addition of the optimum zero-lag prefilter in contrast to the system with optimum postfilter only is rather small. But on the other hand, the addition of a prefilter does not necessarily complicate the implementation of the overall system since the addition of the optimum prefilter simplifies the structure of the optimum postfilter.

Thus, we conclude that when the primary observation noise n_1 is rather severe, or when high accuracy is required, it is probably worthwhile to include the zero-lag optimum prefilter in the system.

3. The efficiency of reducing the error by the addition of a prefilter is more prominent in the case with a suboptimum postfilter than in the case with an optimum postfilter. The addition of an optimum prefilter with a given suboptimum postfilter complicates the implementation; but the synthesis process is straightforward.

4. To allow a delay in the estimation process is often more worthwhile than to insert a prefilter in the system.

5. For the examples in Chapter 5, the error reduction achieved by using multiple sampled-data links in contrast to a single sampled-data link is small.

Comparisons of system performance are shown in Figs. 5.9, 5.11, 5.12, 5.15, and 5.16.

On the whole, the numerical part of this work should be expanded. For example, it is well known that one of the major advantages of the state-space approach is that it reduces the storage requirements. Although we have the

theoretical results, we do not have numerical examples to illustrate this point. Another major advantage of the state-space approach is in its applicability to nonstationary processes, etc. In Chapter 2, the design of the postfilter has included the situation with nonstationary processes, time-variant systems, etc., but again, no example has been worked out. Also, the state-space approach enables one to look into the transient response of the system. We have not probed into this problem. Suboptimum systems have not been given a careful look either.

There are several important extensions that merit further work in this area:

1. As mentioned above, one of the main reasons that the state-space approach is adopted is because of its applicability to nonstationary processes, etc. In deriving the optimum prefilter, and in solving the joint optimization problem, we have excluded nonstationary processes and time-variant systems. Inclusion of those cases is certainly one of the important extensions that merits further work. This extension can be started from Eq. (E.18).

2. We have pointed out in Chapter 3, the essential functions of the prefilter are to combat the error due to sampling and to combat the error due to quantization noise or/and channel additive noises. This work deals primarily with its first function. One can start from the degenerated system shown in Fig. 4.1 to investigate both the first and the second functions of the prefilter. This involves the interesting problem of the trade-off between the number of quantization levels and the sampling rate. It also involves the problem of taking the channel disturbance into consideration.

3. We have obtained the optimum fixed-lag post-filter with no consideration given to the prefilter in Chapter 2. It is interesting to notice that the smoothing problem with a prefilter is a two-dimensional problem, namely, delay at the sending end and delay at receiving end. In view of the substantial error reduction achieved by using unrealizable filters, further work in this respect is surely worthwhile.

4. Numerical methods have been mentioned in the Introductory Chapter. With the understanding of the general structure of the optimum pre- and postfilters

obtained from the state-space approach, it is likely that numerical methods can be used to greater advantage. That is, the state-space method used in conjunction with numerical methods might provide a simpler solution to the filtering problem. For example, (3.72) might be solved by numerical methods with some simplified assumptions on the characteristics of the filters.

1. A steady-state solution corresponds to a stable system with an infinite past data available. It is desirable to remove the assumption about the availability of an infinite past data. That is, it is of interest to study the transient response of the systems.

There are other important extensions such as the comparison of the present work with the rate distortion function; the modeling of the state equations; the existence problem of the innovation processes, and the extension to correlated noises.

APPENDIX

Appendix A

CONTINUOUS KALMAN-BUCY STATE-VECTOR FILTER AND ITS EXTENSIONS

Given the state and observation models

$$\dot{x}(t) = A(t) x(t) + \Gamma(t) w(t) \quad (A.1)$$

$$z(t) = H(t) x(t) \quad (A.2)$$

$$r(t) = z(t) + n(t) \quad (A.3)$$

with prior statistics

$$E[w(t)] = E[n(t)] = E[x(0)] = 0 \quad (A.4)$$

$$R_w(t, \tau) = V_w(t) \delta(t - \tau) \quad (A.5)$$

$$R_{xw}(0, \tau) = 0, \quad \text{for } \tau > 0 \quad (A.6)$$

$$R_{wn}(t, \tau) = V_{wn}(t) \delta(t - \tau) \quad (A.7)$$

and

$$R_x(0, 0) = V_x(0). \quad (A.8)$$

We seek the optimal estimate of $x(t)$, $\hat{x}(t/T_i, T_f, r(\tau))$,
(from here on, $\hat{x}(t/T_i, T_f, r(\tau))$ will be denoted by $x(t/T_f)$),
based on $r(\tau)$ for $T_i \leq \tau \leq T_f$ using a minimum-mean-square
error criterion.

Assume first that we are given the innovation process $u(t)$ or $r(t)$. Then from the argument made in Section 1.4.3 of this thesis,

$$\hat{x}(t/T_f) = \int_{T_i}^{T_f} g(t, v) u(v) dv, \quad (A.9)$$

where $g(t, v)$ is to be determined from the orthogonality principle, i.e., from the equation

$$E\{[x(t) - \hat{x}(t/T_f)] u'(\tau)\} = 0 \quad \text{for } T_i \leq \tau \leq T_f \quad (A.10)$$

Substitute (A.9) into (A.10). Since

$$E[u(t) u'(\tau)] = V_u(t) \delta(t - \tau),$$

$$\begin{aligned} E[x(t) u'(\tau)] &= \int_{T_i}^{T_f} g(t, v) E[u(v) u'(\tau)] dv \\ &= g(t, \tau) V_u(\tau); \end{aligned} \quad (A.11)$$

i.e.,

$$g(t, \tau) = E[x(t) u(\tau)] V_u^{-1}(\tau). \quad (A.12)$$

For realizable filters, $T_f = t$ and (A.9) becomes

$$\hat{x}(t) = \int_{T_i}^t g(t, v) u(v) dv. \quad (A.13)$$

Differentiating both sides of (A.13) with respect to t yields

$$\dot{\hat{x}}(t) = \int_{T_i}^t \frac{\partial g(t, \tau)}{\partial t} u(v) dv + g(t, t) u(t). \quad (A.14)$$

From (A.12) and (A.1)

$$\begin{aligned}
 \frac{\partial g(t, v)}{\partial t} &= E\left[\frac{dx(t)}{dt} u(v)\right] V_u^{-1}(v) \\
 &= E\{[A(t) x(t) + \Gamma(t) w(t)] u(v)\} V_u^{-1}(v) \\
 &= A(t) g(t, v) + \Gamma(t) R_{wu}(t, v) V_u^{-1}(v) \quad (A.15)
 \end{aligned}$$

Substituting (A.15) into (A.14) yields

$$\begin{aligned}
 \dot{\hat{x}}(t) &= A(t) \hat{x}(t) + g(t, t) u(t) + \\
 &\quad \int_{T_i}^t \Gamma(t) R_{wu}(t, v) V_u^{-1}(v) u(v) dv. \quad (A.16)
 \end{aligned}$$

It can be shown that with the given prior statistics,

$$R_{wu}(t, \tau) = V_{wn}(t) \delta(t - \tau) \quad (A.17)$$

Substituting (A.17) into (A.16) yields

$$\dot{\hat{x}}(t) = A(t) \hat{x}(t) + [g(t, t) + (1/2) V_{wn}(t)] u(t). \quad (A.18)$$

Thus, the key steps in this innovations approach to the Kalman-Bucy filtering problem and its extensions are first, to find the innovation process using an invertible transformation; second, to find $V_u^{-1}(t)$ and $E[x(t) u'(\tau)]$;

and finally, to specify the optimum filter as the solution to a differential equation.

We will not go into the detail of the derivation. The interested reader is directed to Kailath (23) or Sage and Melsa (51). Instead, we will discuss the interpretations of the innovation process. Let us look at two different situations:

1. When $n(t)$ in (A.3) is a white noise, i.e.,

$$R_n(t, \tau) = V_n(t) \delta(t - \tau), \quad (\text{A.19})$$

the innovation process is given by

$$u(t) = r(t) - \hat{z}(t) \quad (\text{A.20})$$

First, let us prove that $u(t)$ is a white noise. From (A.2) and (A.3),

$$\begin{aligned} u(t) &= H(t) x(t) + n(t) - H(t) \hat{x}(t) \\ &= H(t) \tilde{x}(t) + n(t). \end{aligned} \quad (\text{A.21})$$

Thus,

$$\begin{aligned} E[u(t) u'(\tau)] &= E\{[H(t) \tilde{x}(t) + n(t)][r(\tau) - \hat{z}(\tau)]'\} \\ &= H(t) E\{\tilde{x}(t)[r(\tau) - \hat{z}(\tau)]'\} \\ &\quad + E[n(t) r'(\tau)] - E[n(t) \hat{z}'(\tau)]. \end{aligned} \quad (\text{A.22})$$

For $\tau < t$, from the orthogonality principle,

$$E[\tilde{x}(t) [r(\tau) - \hat{z}(\tau)]'] = 0 \quad (\text{A.23})$$

From (A.7) and (A.19),

$$E[n(t) r'(\tau)] = 0 \quad \text{for } \tau < t \quad (\text{A.24})$$

Also, since $\hat{z}(\tau)$ depends on $r(\rho)$ and $n(\rho)$ for $\rho \leq \tau < t$ only,

$$E[n(t) \hat{z}'(\tau)] = 0 \quad (\text{A.25})$$

From (A.22) through (A.25),

$$E[u(t) u'(\tau)] = 0 \quad \text{for } \tau < t \quad (\text{A.26})$$

A similar argument applies for $\tau > t$; hence, for $\tau \neq t$,

$$E[u(t) u'(\tau)] = 0 \quad (\text{A.27})$$

For $\tau = t$,

$$\begin{aligned} E[u(t) u'(t)] &= E\{[H(t) \tilde{x}(t) + n(t)] \cdot [H(t) \tilde{x}(t) + n(t)]'\} \\ &= H(t) V_{\tilde{x}}(t) H'(t) + H(t) \cdot E[\tilde{x}(t) n'(t)] + \\ &\quad E[n(t) \tilde{x}'(t)] H'(t) + E[n(t) n'(t)]. \quad (\text{A.28}) \end{aligned}$$

Because of (A.3) and the orthogonality principle,

$$\begin{aligned} E[\tilde{x}(t) n'(t)] &= E[\tilde{x}(t) r'(t)] - E[\tilde{x}(t) x'(t)] H'(t) \\ &= -E[\tilde{x}(t) x'(t)] H'(t) \\ &= -V_{\tilde{x}}(t) H'(t), \quad (\text{A.29}) \end{aligned}$$

i.e.,

$$E[\tilde{x}(t) n'(t)] = \{E[n(t) x'(t)]\}' = -V_x^v(t) H'(t) \quad (A.30)$$

Since $V_x^v(t) \leq V_x(t)$, a finite quantity, and $V_{wn}(t)$ is also given a finite value, all the quantities in (A.28) are of finite value except the term $E[n(t) n'(t)]$ which is infinite. Hence,

$$E[u(t) u'(\tau)] = E[n(t) n'(\tau)]. \quad (A.31)$$

From (A.31) and (A.27),

$$E[u(t) u'(\tau)] = V_n(t) \delta(t-\tau). \quad (A.32)$$

The optimal zero-lag realizable filter for estimating $x(t)$ from $r(t)$ with $n(t)$ being a white noise is specified by the following equations: (51)

$$\dot{\hat{x}}(t) = A(t) \hat{x}(t) + K_c(t) u(t) \quad (A.33)$$

$$u(t) = r(t) - H(t) \hat{x}(t) \quad (A.34)$$

$$K_c(t) = [V_x^v(t) H'(t) + \Gamma(t) V_{wn}(t)] V_u^{-1}(t) \quad (A.35)$$

$$\begin{aligned} \dot{V}_x^v(t) = & A(t) V_x^v(t) + V_x^v(t) A'(t) + \Gamma(t) V_w(t) \Gamma'(t) \\ & - K_c(t) V_n(t) K_c'(t) \end{aligned} \quad (A.36)$$

2. $n(t)=0$; in this case, the right-hand side of (A.20) is zero, and it no longer represents the innovation process of $r(t)$. Instead, the innovation process is given by

$$u(t) = \dot{r}(t) - H^*(t) x(t) \quad (\text{A.37})$$

where

$$H^*(t) = \dot{H}(t) + H(t) A(t). \quad (\text{A.38})$$

and

$$\begin{aligned} \dot{r}(t) &= \frac{d[H(t) \cdot x(t)]}{dt} \\ &= \dot{H}(t) x(t) + H(t) \dot{x}(t) \\ &= \dot{H}(t) x(t) + H(t) [A(t) x(t) + \Gamma(t) w(t)] \\ &= H^*(t) x(t) + H(t) \Gamma(t) w(t). \end{aligned} \quad (\text{A.39})$$

Substituting (A.39) into (A.37) yields

$$u(t) = H^*(t) \tilde{x}(t) + H(t) \Gamma(t) w(t). \quad (\text{A.40})$$

(A.40) is in the same form as (A.21) with $H^*(t)$ replacing $H(t)$, and $w(t)$ playing the role of $n(t)$. The proof that $u(t)$ given by (A.40) is a white noise is similar to the previous proof; the key equations are, for $\tau < t$,

$$E[\tilde{x}(t) \dot{r}'(\tau)] = \frac{\partial E[\tilde{x}(t) r(\tau)]}{\partial \tau} = 0 \quad (\text{A.41})$$

$$E[\tilde{x}(t) \hat{x}'(\tau)] = 0 \quad (\text{A.42})$$

$$E[w(t) \dot{r}'(\tau)] = \frac{\partial E[w(t) r(\tau)]}{\partial \tau} = 0 \quad (\text{A.43})$$

$$E[w(t) \hat{x}(\tau)] = 0 \quad (\text{A.44})$$

The first two equations follow from the orthogonality principle, and the last two equations follow from the fact that $w(t)$ is uncorrelated with $x(\tau)$ and $n(\tau)$ for $\tau < t$.

Thus, for $\tau < t$,

$$\begin{aligned} E[u(t) u'(\tau)] &= E\{[H^*(t) \tilde{x}'(t) + H(t) \Gamma(t) w(t)] \\ &\quad [\dot{r}(\tau) - H^*(\tau) x(\tau)]'\} \\ &= 0 \end{aligned} \quad (\text{A.45})$$

A similar argument applies for $\tau > t$. For $\tau = t$,

$$E[u(t) u'(t)] = E\{[H^*(t) \tilde{x}(t) + H(t) \Gamma(t) w(t)]^2\} \quad (\text{A.46})$$

Similarly, $H(t) \cdot \Gamma(t) \cdot E[w(t) \cdot w'(t)] \cdot \Gamma'(t) \cdot H'(t)$ is the only term in (A.46) that has an infinite value, thus,

$$E[u(t) u'(\tau)] = V_u(t) \delta(t-\tau), \quad (\text{A.47})$$

where

$$V_u(t) = H(t) \Gamma(t) V_w(t) \Gamma'(t) H'(t). \quad (A.48)$$

The results for the optimal zero-lag realizable filter in this case are given by the following equations: (pp. 338, reference 51)

$$\dot{\hat{x}}(t) = A(t) \hat{x}(t) + K_C(t) u(t). \quad (A.49)$$

$$u(t) = \dot{r}(t) - H^*(t) \hat{x}(t) \quad (A.50)$$

$$K_C(t) = [V_x(t) H^{*'}(t) + \Gamma(t) V_w(t) \Gamma'(t) H'(t)] \times [H(t) \Gamma(t) V_w(t) \Gamma'(t) H'(t)]^{-1} \quad (A.51)$$

$$\begin{aligned} \dot{V}_x(t) = & A(t) V_x(t) + V_x(t) A'(t) + \Gamma(t) V_w(t) \Gamma'(t) \\ & - K_C(t) [H(t) \Gamma(t) V_w(t) \Gamma'(t) H'(t)] K_C'(t), \end{aligned} \quad (A.52)$$

where $H^*(t)$ is given by (A.38).

For a discussion about how to specify the initial conditions, refer to pp. 544 of Van Trees (58).

When $n(t)$ in (A.3) is a colored noise, then depending on whether it contains a white noise component or not, the problem can be reduced to one of the above two cases.

Let

$$n(t) = n_w(t) + n_r(t), \quad (A.53)$$

where $n_w(t)$ is a white noise, and $n_r(t)$ is specified by the state equations

$$\dot{\zeta}(t) = A(t) \zeta(t) + \Gamma_r(t) w_r(t) \quad (\text{A.54})$$

$$n_r(t) = I_c(t) w_r(t), \quad (\text{A.55})$$

where

$$E[w_r(t) w_r(\tau)] = V_{w_r} \delta(t-\tau). \quad (\text{A.56})$$

Augment the original state equation to include $\delta(t)$, and define

$$x_a(t) = [x(t), \zeta(t)]', \quad (\text{A.57})$$

$$w_a(t) = [w(t), w_r(t)] \quad (\text{A.58})$$

Then

$$\dot{x}_a(t) = \begin{bmatrix} A(t) & 0 \\ 0 & B(t) \end{bmatrix} x_a(t) + \begin{bmatrix} \Gamma(t) & 0 \\ 0 & \Gamma_r(t) \end{bmatrix} w_a(t), \quad (\text{A.59})$$

and

$$r(t) = [H(t), I_c(t)] x_a(t) + n_w(t) \quad (\text{A.60})$$

where

$$R_{w_a}(t, \tau) = E[w_a^2(t)] \cdot \delta(t-\tau)$$

$$= \begin{bmatrix} V_w(t) & V_{wwr}(t) \\ V_{wrw}(t) & V_{wr}(t) \end{bmatrix} \delta(t-\tau). \quad (A.61)$$

If $n_w(t) \neq 0$, the above equations are the model of the case with additive white noise. If $n_w(t) = 0$, Eq. (A.57) through (A.60) is the model of the noise-free case.

To give a heuristic look at the invertibility of the transformation which yields the innovation process, note that from (A.33) or (A.49),

$$\dot{\hat{x}}(t) = A(t) \hat{x}(t) + K_C(t) u(t).$$

For the case with additive white noise, we have from (A.34)

$$r(t) = u(t) + H(t) \hat{x}(t). \quad (A.62)$$

For the noise-free case,

$$\dot{r}(t) = u(t) + H^*(t) \hat{x}(t). \quad (A.63)$$

Thus, if $u(t)$ is given, $r(t)$ can be obtained either by (A.33) and (A.62), or by (A.49) and (A.63) depending on the characteristic of the noise.

Substituting (A.34) into (A.33) yields

$$\dot{\hat{x}}(t) = [A(t) - K_C(t) H(t)] \hat{x}(t) + K_C(t) r(t) \quad (A.64)$$

(A.63) represents a feedback system. A similar situation can be found in the noise-free case. Due to this feedback, there is no need to store $r(\tau)$ for $\tau < t$. This feedback loop accounts for the recursiveness of the solution.

Appendix B

DISCRETE KALMAN-BUCY STATE-VECTOR FILTER AND ITS EXTENSIONS

The discrete state and observation equations are

$$x(k) = \Phi(k, k-1) x(k-1) + w_d(k-1), \quad (B.1)$$

$$z(k) = H(k) x(k), \quad (B.2)$$

and

$$r(k) = z(k) + n(k). \quad (B.3)$$

with prior statistics

$$E[w_d(k)] = E[n(k)] = E[x(0)] = 0 \quad (B.4)$$

$$R_{w_d}(j, k) = V_{w_d}(j) \delta_{jk}, \quad (B.5)$$

$$R_n(j, k) = V_n(j) \delta_{jk}, \quad (B.6)$$

$$R_{nw_d}(j, k) = V_{nw_d}(j) \delta_{jk}, \quad (B.7)$$

and

$$R_x(0, 0) = V_x(0). \quad (B.8)$$

We seek for the linear least-squares estimate of $x(k)$, $\hat{x}(k/\ell_i, \rho)$ (to be denoted by $\hat{x}(k/\rho)$ from here on), based on the observation of $r(j)$ for $\ell_i \leq j \leq \rho$.

The general procedure used in deriving the optimal discrete filter is similar to that used for the continuous optimal filter described in Appendix A, and is simpler. Thus, the details will not be given here. For the details, refer to Sage and Melsa (51; Chapter 7). The solutions are

$$\hat{x}[(j+1)/j] = \Phi(j+1, j) \hat{x}(j/j) + K_0(j) [r(j) - H(j) \hat{x}(j/j)] \quad (B.9)$$

$$\hat{x}(j/j) = \hat{x}(j/(j-1)) + K_d(j) u_d(j), \quad (B.10)$$

$$u_d(j) = r(j) - H(j) \hat{x}(j/(j-1)), \quad (B.11)$$

$$K_0(j) = V_{wn}(j) V_n^{-1}(j), \quad (B.12)$$

$$K_d(j) = V_x(j/(j-1)) H'(j) V_{u_d}^{-1}(j), \quad (B.13)$$

$$V_{u_d}(j) = H(j) V_x(j/(j-1)) H'(j) + V(j), \quad (B.14)$$

$$\begin{aligned} V_x(j/(j-1)) &= [\Phi(j, j-1) - K_0(j-1) H(j-1)] V_x(j-1) [\Phi(j, j-1) \\ &\quad - K_0(j-1) H(j-1)]' + V_w(j-1) \\ &\quad - K_0(j-1) V_w(j-1) K_0'(j-1), \end{aligned} \quad (B.15)$$

$$V_x(j) = [I - K_d(j) H(j)] V_x(j/(j-1)), \quad (B.16)$$

where $u_d(j)$ is the discrete innovation process of $r(j)$.

The above equations can be applied to the case when

$n(k) \neq 0$ as well as to the case when $n(k)=0$. When $n(k) \neq 0$, (B.13) can be written in the alternative form

$$K_d(j) = V_x(j) H'(j) V_n^{-1}(j). \quad (B.17)$$

Thus, in discrete estimation, it is no longer necessary to distinguish between the case in which $n(k)$ is present and that in which $n(k)$ is absent. The second case is simply a special case of the former one. To prove that $u_d(j)$ is a white noise, write $u_d(j)$ in the form

$$\begin{aligned} u_d(j) &= r(j) - H(j) \hat{x}(j/(j-1)) \\ &= H(j) \tilde{x}(j/(j-1)) + n(j) \end{aligned} \quad (B.18)$$

Then

$$\begin{aligned} E[u_d(j) u_d'(k)] &= E\{[H(j) \tilde{x}(j/(j-1)) + n(j)] \\ &\quad [r(k) - H(k) \hat{x}(k/(k-1))]\} \end{aligned} \quad (B.19)$$

If $k < j$, it follows from the orthogonality principle that

$$E[\tilde{x}(j/(j-1)) r'(k)] = 0 \quad (B.20)$$

and

$$E[\tilde{x}(j/(j-1)) \hat{x}'(k/(k-1))] = 0. \quad (B.21)$$

From (B.6) and (B.7),

$$E[n(j) r'(k)] = 0. \quad (B.22)$$

Also, since $\hat{x}(k/(k-1))$ depends on $x(\ell)$ and $n(\ell)$ for $\ell \leq k-1$ only,

$$E[n(j) \hat{x}(k/(k-1))] = 0. \quad (B.23)$$

From (B.19) through (B.23), for $k < j$,

$$E[u_d(j) u_d'(k)] = 0. \quad (B.24)$$

Similarly, for $k > j$,

$$E[u_d(j) u_d'(k)] = 0 \quad (B.25)$$

For $j=k$,

$$\begin{aligned} E[u_d(k) u_d'(k)] &= E\{[H(k) \hat{\tilde{x}}(k/(k-1)) + n(k)] \\ &\quad [H(k) \hat{\tilde{x}}(k/(k-1)) + n(k)]'\} \\ &= H(k) V_{\tilde{x}}(k/(k-1)) H'(k) + V_n(k) \\ &\quad + E[n(k) \hat{\tilde{x}}'(k/(k-1))] H'(k) \\ &\quad + H(k) E[\hat{\tilde{x}}(k/(k-1)) n'(k)]. \end{aligned} \quad (B.26)$$

Since $\hat{\tilde{x}}(k/(k-1))$ is a linear combination of $r(\ell)$ for $\ell \leq k-1$, which are uncorrelated with $n(k)$, and $x(k)$ is also uncorrelated with $n(k)$ (from (B.1) and (B.7)),

$$\begin{aligned} E[\hat{\tilde{x}}(k/(k-1)) n'(k)] &= \{E[n(k) \hat{\tilde{x}}'(k/(k-1))]\}' \\ &= 0 \end{aligned} \quad (B.27)$$

Hence,

$$V_{u_d}(k) = E[u_d(k) u_d'(k)] = H(k) \cdot V_x^u(k/(k-1)) \cdot H'(k) + V_n(k). \quad (B.28)$$

From (B.21), (B.22), and (B.25),

$$E[u_d(k) u_d'(j)] = V_{u_d}(k) \delta_{kj} \quad (B.29)$$

with $V_{u_d}(k)$ given by (B.28). If $n(k)=0$, (B.29) reduces to

$$E[u_d(k) u_d'(j)] = [H(k) V_x^u(k/(k-1)) H'(k)] \delta_{kj}. \quad (B.30)$$

Colored noise problems can again be handled by augmenting the state representation. Assume that

$$n(k) = n_w(k) + n_c(k) \quad (B.31)$$

where $n_w(k)$ is a discrete white noise, and $n_c(k)$ is specified by the state equations

$$\xi(k) = \phi_n(k, k-1) \xi(k-1) + w_{dn}(k-1) \quad (B.32)$$

$$n_c(k) = I_d(k) \xi(k) \quad (B.33)$$

where

$$R_{w_{dn}}(k, j) = V_{w_{dn}} \xi(k) \delta_{kj} \quad (B.34)$$

and

$$R_{w_{dn} w_d}(k, j) = V_{nd}(k) \delta_{kj}. \quad (B.35)$$

Let

$$x_a(k) = [x(k), \xi(k)]'. \quad (B.36)$$

and

$$w_a(k) = [w_d(k), w_{dn}(k)]'. \quad (B.37)$$

Then

$$x_a(k) = \begin{bmatrix} \Phi(k, k-1) & 0 \\ 0 & \Phi_n(k, k-1) \end{bmatrix} x_a(k-1) + w_a(k-1) \quad (B.38)$$

and

$$r(k) = [H(k), I_d(k)] x_a(k) + n_w(k) \quad (B.39)$$

with

$$E[w_a(k) w_a'(j)] = \begin{bmatrix} V_{w_d}(k) & V_{dn}(k) \\ V_{nd}(k) & V_{w_{dn}}(k) \end{bmatrix} \delta_{kj} \quad (B.40)$$

Thus, the problem reduces to the white noise, or to the noise-free case.

Appendix C

DISCRETIZATION OF CONTINUOUS STATE EQUATION

The problem is to discretize the continuous state equation given by (2.1) with prior statistics given by (2.5) through (2.8). From (2.1) it follows that

$$x(t) = \Phi(t, \tau) x(\tau) + \int_{\tau}^t \Phi(t, v) \Gamma(v) w(v) dv \quad (C.1)$$

where $\Phi(t, v)$ is the transition matrix of $A(t)$ with the following properties (51):

Property 1.

$$\partial \Phi(t, \tau) / \partial t = A(t) \Phi(t, \tau). \quad (C.2)$$

Property 2.

$$\Phi(t, \tau) \cdot \Phi(\tau, \rho) = \Phi(t, \rho). \quad (C.3)$$

Property 3.

$$\Phi(t, t) = I \triangleq \text{the identity matrix.} \quad (C.4)$$

Substituting $t=kT$ and $\tau=(k-1)T$ into (C.1) yields

$$x(k) = \Phi(k, k-1) x(k-1) + \int_{(k-1)T}^{kT} \Phi(kT, v) \Gamma(v) w(v) dv. \quad (C.5)$$

Let (22)

$$w_d(k-1) = \int_{(k-1)T}^{kT} \Phi(kT, \tau) \Gamma(\tau) w(\tau) d\tau. \quad (C.6)$$

Then,

$$x(k) = \Phi(k, k-1) x(k-1) + w_d(k-1) \quad (C.7)$$

Some of the properties of $w_d(k)$ are the following:

Property 1.

$$E[w_d(k) w_d'(j)] = V_{w_d}(k) \delta_{kj} \quad (C.8)$$

where

$$V_{w_d}(k) = \int_{kT}^{(k+1)T} \Phi((k+1)T, \tau) \Gamma(\tau) V_w(\tau) \Gamma'(\tau) \Phi'((k+1)T, \tau) d\tau. \quad (C.9)$$

This property can be established as follows: From (2.28),

$$\begin{aligned} E[w_d(k) w_d'(j)] &= \int_{kT}^{(k+1)T} \int_{jT}^{(j+1)T} \Phi((k+1)T, v) \Gamma(v) \\ &\quad \times E[w(v) \cdot w'(\tau)] \Gamma'(\tau) \Phi'((j+1)T, \tau) dv d\tau \\ &\quad (C.10) \end{aligned}$$

$$\begin{aligned} &= \int_{kT}^{(k+1)T} \int_{jT}^{(j+1)T} \Phi((k+1)T, v) \Gamma(v) V_w(v) \times \\ &\quad \delta(v-\tau) \Gamma'(\tau) \Phi'((j+1)T, \tau) dv d\tau; \quad (C.11) \end{aligned}$$

i.e.,

$$\begin{aligned} E[w_d(k) \cdot w_d'(j)] &= \int_{kT}^{(k+1)T} \Phi((k+1)T, v) \Gamma(v) V_w(v) \Gamma'(v) \\ &\quad \times \Phi'((k+1)T, v) dv \quad (C.12_a) \end{aligned}$$

if $j=k$, and

$$E[w_d(k) w_d'(j)] = 0 \quad \text{otherwise.} \quad (C.12_b)$$

Property 2.

$$E[x(t) w_d'(kT)] = 0 \quad \text{for } t < kT. \quad (C.13)$$

This property follows directly from (2.7), (2.6), and the definition of $w_d(k)$ given by (C.6).

Property 3.

$$E[n(k) w_d'(j)] = 0 \quad (C.14)$$

This follows directly from (2.8) and (C.6). Thus, the discrete state equation given by (C.7) fits the general form of the discrete state equation as given in Appendix B with $V_{wn}(k)=0$.

There is one more property of interest.

Property 4.

$$E[w(t) w_d'(k-1)] = 0 \quad \text{for } t < (k-1)T, \text{ or } t > kT. \quad (C.15)$$

To establish this property, we write

$$E[w(t) w_d'(k-1)] = \int_{(k-1)T}^{kT} E[w(t) w'(\nu)] \Gamma'(\nu) \Phi'(kT, \nu) d\nu. \quad (C.16)$$

From (2.7), it follows that

$$E[w(t) w_d'(k-1)] = \int_{(k-1)T}^{kT} V_w(t) \delta(t-v) \Gamma'(v) \Phi'(kT, v) dv. \quad (C.17)$$

Thus,

$$E[w(t) w_d'(k-1)] = \begin{cases} V_w(t) \Gamma'(t) \Phi'(kT, t) & \text{for } (k-1)T < t < kT, \\ (1/2) V_w(t) \Gamma'(t) \Phi'(kT, t) & \text{for } t = kT, (k-1)T, \\ 0 & \text{otherwise.} \end{cases} \quad (C.18)$$

In this work, the observation process is assumed given in discrete form, where $z(k)$ in (2.3) is given by

$$z(k) = H(k) x(k). \quad (C.19)$$

If a continuous noise $n(t)$ is added before sampling, discretization of $n(t)$ is needed. This does not present any problem if $n(t)$ does not contain a white noise component. The corresponding discretized noise is simply $n(k)$ with autocorrelation function equal to $R_n(t, \tau)$, where $t = kT$ and $\tau = jT$ for any integers k, j . If $n(t)$ does contain a white noise component, the discretization of $n(t)$ is not so straightforward. Since in a continuous time system, the autocorrelation function of a white noise is an impulse function at $t=0$; whereas it is senseless to have an

impulse function appear in the discrete model. The point is that the white noise representation is only an approximated model of some random process. To find out the discretized model of a noise which is approximated by a continuous white noise, one has to go back to the original correlation function that characterize the noise. For example, a noise $n_0(t)$ with an autocorrelation function shown in Fig. C.1 can be approximated by a continuous white noise, i.e.,

$$R_{n_0}(\tau) = V_{n_0} \delta(\tau) \quad (C.20)$$

where $\delta(\tau) \triangleq$ the unit impulse function, and

$$V_{n_0} = \int_{-\infty}^{\infty} R_{n_0}(\tau) d\tau. \quad (C.21)$$

If this noise is sampled with sampling period T , then the only values of $R_{n_0}(t)$ that needed to be investigated are at $t = \dots, -T, 0, T, \dots$. If $R_{n_0}(jT)/R_{n_0}(0)$ for $j \neq 0$ is small enough, one can approximate $n_0(k)$ by a discrete white noise, i.e.,

$$R_{n_0}(k-j) = V_{n_0} \delta_{kj} \quad (C.22)$$

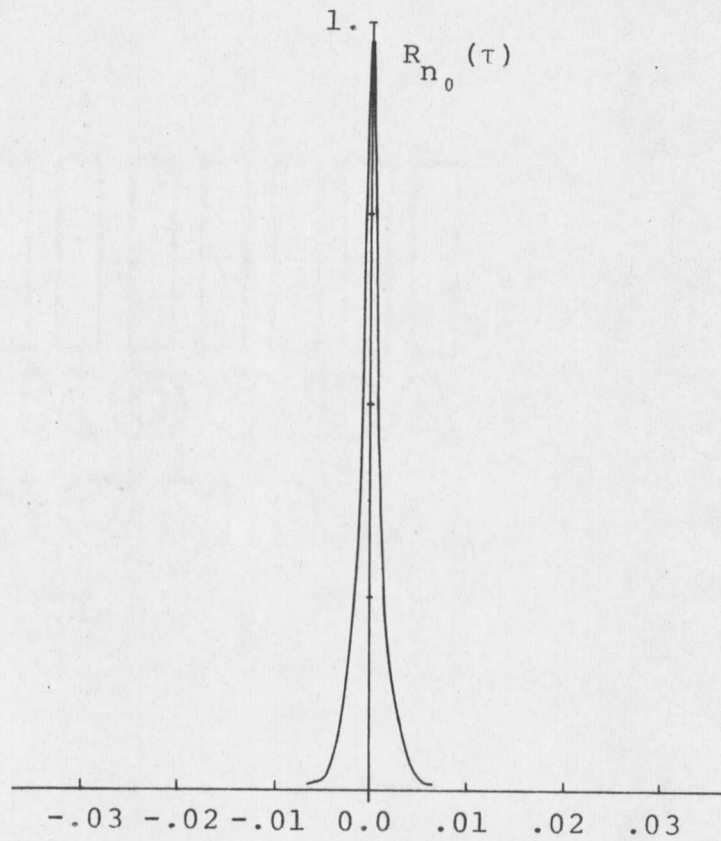


Figure C.1 The autocorrelation function of random process which can be approximated by a white noise.

where

$$V_{n_0 d} = R_{n_0}(0). \quad (C.23)$$

Thus, one can approximate both $n_0(t)$ by a continuous white noise, and $n_0(k)$ by a discrete white noise but with different variances, i.e., $V_{n_0} \neq V_{n_0 d}$ in general.

Appendix D.

ERROR COVARIANCE OF LINEAR ESTIMATION

A general expression of the error covariance of the filtering, prediction, and smoothing problems is developed first. Let $V_{\xi}(t) = E[\xi^2(t)]$. Then from (2.31), it follows that

$$V_{\hat{x}}(t/l_f) = \sum_{l=l_i}^{l_f} \sum_{k=l_i}^{l_f} g(t, lT) E[u_d(l) u_d(k)] g'(t, kT), \quad (D.1)$$

where $\hat{x}(t/l_f) = \hat{x}(t/l_i, l_f, r(k)) \triangleq$ linear least-square estimate of $x(t)$ given the $r(k)$'s for $l_i \leq k \leq l_f$. Substituting (B.29) into (D.1) yields

$$V_{\hat{x}}(t/l_f) = \sum_{l=l_i}^{l_f} g(t, lT) V_{u_d}(l) g'(t, lT), \quad (D.2)$$

where $V_{u_d}(l)$ is given by (B.28). From (2.12), it follows that

$$\begin{aligned} V_{\hat{x}}(t/l_f) &= E\{[x(t) - \hat{x}(t/l_f)]^2\} \\ &= V_x(t) - E[x(t) \hat{x}'(t/l_f)] - E[\hat{x}(t/l_f) x'(t)] \\ &\quad + V_{\hat{x}}(t/l_f). \end{aligned} \quad (D.3)$$

From (2.12) and the orthogonality principle

$$E[x(t) \hat{x}'(t/\ell_f)] = E[\hat{x}(t/\ell_f) x'(t)] = V_{\hat{x}}(t/\ell_f) \quad (D.4)$$

Thus, (D.3) becomes

$$\begin{aligned} V_{\hat{x}}(t/\ell_f) &= V_x(t) - V_{\hat{x}}(t/\ell_f) \\ &= V_x(t) - \sum_{\ell=\ell_i}^{\ell_f} g(t, \ell T) V_{u_d}(\ell) g'(t, \ell T), \end{aligned} \quad (D.5)$$

where the second equality results from the use of (D.2).

Since $d(t) = C(t) \cdot x(t)$,

$$V_d(t/\ell_f) = V_d(t) - \sum_{\ell=\ell_i}^{\ell_f} C(t) g(t, \ell T) V_{u_d}(\ell) g'(t, \ell T) C'(t). \quad (D.6)$$

Depending on whether $\ell_f < m$, $\ell_f = m$, or $\ell_f > m$, where $m = \langle t/T \rangle$; we are dealing with the prediction, filtering, or smoothing problems, respectively. Let

$$g_d(t, \ell T) = C(t) g(t, \ell T) \sqrt{V_{u_d}(\ell)} \quad (D.7)$$

Then (D.6) becomes

$$V_d(t/\ell_f) = V_d(t) - \sum_{\ell=\ell_i}^{\ell_f} g_d^2(t, \ell T). \quad (D.8)$$

For prediction and filtering problem, $\ell_f \leq m$. Substituting (2.41) into (D.7) to obtain, for $\ell \leq m$.

$$\begin{aligned} g_d(t, \ell T) &= C(t) \Phi(t, \ell T) V_x^\sim(\ell/\ell-1) H'(\ell) \sqrt{V_{u_d}^{-1}(\ell)} \\ &= C(t) \Phi(t, \ell T) K_d(\ell) \sqrt{V_{u_d}(\ell)}, \end{aligned} \quad (D.9)$$

where $K_d(\ell)$ is given by (B.13). When $\ell_f = m$ (zero-lag), Eq. (D.8) becomes

$$V_d^\sim(t/m) = V_d^\sim(t) = V_d(t) - \sum_{\ell=\ell_i}^m g_d^2(t, \ell T) \quad (D.10)$$

with $g(t, \ell T)$ for $\ell \leq m$ given by (D.9). For smoothing problem $\ell_f > m$, (D.7) becomes

$$\begin{aligned} V_d^\sim(t/\ell_f) &= V_d^\sim(t, n_d) = V_d^\sim(t) - \sum_{\ell=\ell_i}^{m+n_d} g_d^2(t, \ell T) \\ &= V_d^\sim(t) - \sum_{k=1}^{n_d} g_d^2(t, (m+k)T), \end{aligned} \quad (D.11)$$

where $n_d = \ell_f - m > 0$, $V_d^\sim(t)$ is given by (D.10), and $g_d(t, (m+k)T)$ for $k > 0$ can be obtained by substituting (2.90) into (D.7), i.e.,

$$g_d(t, (m+k)T) = C(t) V_x(t) \Phi'((m+1)T, t) \Lambda_m(k) H'((m+k)T) \sqrt{V_{u_d}^{-1}(m+k)}$$

for $k > 0$. (D.12)

For a stationary input signal, $\Phi(t, \ell T) = \Phi(t - \ell T)$, and $C(t)$ and $H(k)$ become constants. At the steady state $(\ell_i \rightarrow -m)$, $K_d(\ell)$ and $V_{u_d}(\ell)$ become constants. Under these conditions $g_d(t, \ell T) = g_d(t - \ell T)$ for all ℓ . Substituting this expression into (D.8) and integrating both sides of (D.5) from $t=0$ to $t=T$ to obtain

$$\begin{aligned} (1/T) \int_0^T V_d^2(t/m+n_d) dt &= V_d - (1/T) \int_0^T \sum_{\ell_i=-\infty}^{m+n_d} g_d^2(t-\ell T) dt \\ &= V_d - (1/T) \int_{-n_d}^{\infty} g_d^2(v) dv \end{aligned} \quad (D.13)$$

where the second equality comes from letting $v=t-\ell T$ and interchanging the order of integration and summation; $g_d(v)$ is given by

$$g_d(v) = C \Phi(v) K_d \sqrt{V_{u_d}^{-1}} \quad (D.14)$$

for $v > 0$, and

$$g_d(v) = C V_x^2(v+kT) \Phi(-v-(k-1)T) \{[I-H' K_d'] \Phi'(T)\}^k H \cdot V_{u_d}^{-1} \quad (D.15)$$

for $-kT \leq v < -(k-1)T$, where $k > 0$. The expression

$(1/T) \int_0^T \tilde{V}_d^v(t) \cdot dt$ represents the error with a penalty function

$$p(t) = \begin{cases} (1/T) & \text{if } 0 \leq t < T \\ 0 & \text{otherwise.} \end{cases} \quad (D.16)$$

The $g_d(v)$ corresponds to the 'reconstruction efficiency density' defined by Tretter and Steiglitz (57). More generally, g_d is a function of two variables as shown in Eq. (D.7).

To obtain a differential equation for the zero-lag error covariance $V_x^v(t)$, we first write

$$\frac{d}{dt} E[\tilde{x}(t) \tilde{x}'(t)] = E[(d\tilde{x}(t)/dt) \tilde{x}'(t)] + E[\tilde{x}(t) (d\tilde{x}'(t)/dt)]. \quad (D.17)$$

Subtracting (2.43) from (2.1) to obtain for $t \neq kT$, $k=0, 1, \dots$,

$$d\tilde{x}(t)/dt = A(t) \tilde{x}(t) + \Gamma(t) w(t) \quad (D.18)$$

Using the result of (D.18), the first term on the right-hand side of (D.17) can be written as

$$\begin{aligned}
 E[(dx(t)/dt) \tilde{x}'(t)] &= E\{[A(t) \tilde{x}(t) + \Gamma(t) w(t)] \tilde{x}'(t)\} \\
 &= A(t) V_{\tilde{x}}(t) + \Gamma(t) E[w(t) \tilde{x}'(t)].
 \end{aligned}$$

(D.19)

Again from (D.18), it follows that for $kT \leq \rho \leq (k+1)T$,

$k=0,1,\dots$,

$$\tilde{x}(t) = \Phi(t, \rho) \tilde{x}(\rho) + \int_{\rho}^t \Phi(t, v) \Gamma(v) w(v) dv \quad (D.20)$$

Thus,

$$\begin{aligned}
 E[w(t) \tilde{x}'(t)] &= E[w(t) \tilde{x}'(\rho)] \Phi'(t, \rho) + \int_{\rho}^t E[w(t) w(v)] \\
 &\quad \Phi'(t, v) \Gamma'(v) dv.
 \end{aligned} \quad (D.21)$$

Since $\tilde{x}(\rho)$ depends on $x(\tau)$ and $n(\tau)$ for $\tau \leq \rho$ only,

$$E[w(t) \tilde{x}'(\rho)] = 0 \quad \text{for } \rho < t. \quad (D.22)$$

From (2.6), it follows that

$$\begin{aligned}
 \int_{\rho}^t E[w(t) w(v)] \Phi'(t, v) \Gamma'(v) dv &= \int_{\rho}^t V_w(t) \delta(t-v) \Phi'(t, v) \Gamma'(v) dv \\
 &= (1/2) V_w(t) \Gamma'(t).
 \end{aligned} \quad (D.23)$$

Substituting (D.23) into (D.19) to obtain

$$E[(dx(t)/dt) \tilde{x}'(t)] = A(t) V_{\tilde{x}}(t) + (1/2) \Gamma(t) V_w(t) \Gamma'(t). \quad (D.24)$$

Similarly,

$$E[\hat{\mathbf{x}}(t) (d\hat{\mathbf{x}}'(t)/dt)] = \mathbf{V}_{\mathbf{x}}^{\gamma}(t) \mathbf{A}'(t) + (1/2) \Gamma(t) \mathbf{V}_{\mathbf{w}}(t) \Gamma'(t). \quad (\text{D.25})$$

Substituting (D.24) and (D.25) into (D.17) yields, for $mT < t < (m+1)T$,

$$\dot{\mathbf{V}}_{\mathbf{x}}^{\gamma}(t) = \mathbf{A}(t) \mathbf{V}_{\mathbf{x}}^{\gamma}(t) + \mathbf{V}_{\mathbf{x}}^{\gamma}(t) \mathbf{A}'(t) + \Gamma(t) \mathbf{V}_{\mathbf{w}}(t) \Gamma'(t). \quad (\text{D.26})$$

At the sampling instants, from (2.12), (2.44), and (2.45), it follows that

$$\hat{\mathbf{x}}(mT^-) = \mathbf{x}(mT) - \sum_{k=\ell_i}^{m-1} \Phi(m, k) \mathbf{K}_d(k) \mathbf{u}_d(k) \quad (\text{D.27})$$

and

$$\begin{aligned} \hat{\mathbf{x}}(mT^+) = \mathbf{x}(mT) - \left[\sum_{k=\ell_i}^{m-1} \Phi(m, k) \mathbf{K}_d(k) \mathbf{u}_d(k) \right. \\ \left. + \mathbf{K}_d(m) \mathbf{u}_d(m) \right] \end{aligned} \quad (\text{D.28})$$

Thus,

$$\hat{\mathbf{x}}(mT^+) = \hat{\mathbf{x}}(mT^-) - \mathbf{K}_d(m) \mathbf{u}_d(m) \quad (\text{D.29})$$

and

$$\begin{aligned}
V_{\tilde{x}}(mT^+) &= V_{\tilde{x}}(mT^-) - E[\tilde{x}(mT^-) u_d'(mT)] K_d'(m) - \\
&\quad K_d(m) E[u_d(mT^-) \tilde{x}'(mT^-)] \\
&\quad + K_d(m) V_{u_d}(m) K_d'(m). \quad (D.30)
\end{aligned}$$

From (D.27), it follows that

$$\begin{aligned}
E[\tilde{x}(mT^-) u_d'(mT)] &= E\{[x(mT) - \sum_{i=\ell_i}^{m-1} \Phi(m,i) K_d(i) u_d(i)] u_d'(m)\} \\
&= E[x(mT) u_d'(mT)] \\
&\quad - \sum_{i=\ell_i}^{m-1} \Phi(m,i) \cdot K_d(i) E[u_d(i) u_d'(m)]. \quad (D.31)
\end{aligned}$$

The second term on the right-hand side of (D.31) is zero since $u_d(k)$ is a discrete white noise. Thus,

$$\begin{aligned}
E[\tilde{x}(mT^-) u_d'(mT)] &= E[x(m) u_d'(m)] \\
&= V_{\tilde{x}}(m/m-1) H'(m) \\
&= K_d(m) V_{u_d}(m) \quad (D.32)
\end{aligned}$$

The second equality follows from (2.40), and the third equality follows from (B.13). Through the use (D.32), we obtain

$$V_x^{\sim}(mT^+) - V_x^{\sim}(mT^-) = -K_d(m) V_{u_d}^{\sim}(m) K_d'(m) \quad (D.33)$$

Eq. (D.33) can also be obtained from (D.5). Combining (D.33) and (D.26) yields, for all $t > l_i T \triangleq$ the initial time,

$$\begin{aligned} V_x^{\sim}(t) = & A(t) V_x^{\sim}(t) + V_x^{\sim}(t) A'(t) + \Gamma(t) V_w(t) \Gamma'(t) \\ & - \sum_{l=l_i}^{\infty} K_d(l) V_{u_d}^{\sim}(l) K_d'(l) \delta(t-lT). \end{aligned} \quad (D.34)$$

Another form of the zero-lag error expression can be obtained as follows. Substituting $\rho = mT$ into (D.20) to yield for $mT \leq t < (m+1)T$,

$$\tilde{x}(t) = \phi(t, mT) \tilde{x}(m) + \int_{mT}^t \phi(t, v) \Gamma(v) w(v) dv. \quad (D.35)$$

Thus,

$$\begin{aligned} V_x^{\sim}(t) = & E \{ [\phi(t, mT) \tilde{x}(m) + \int_{mT}^t \phi(t, v) \Gamma(v) w(v) dv] \times \\ & [\phi(t, mT) \tilde{x}(m) + \int_{mT}^t \phi(t, \tau) \Gamma(\tau) w(\tau) d\tau] \}. \end{aligned} \quad (D.36)$$

Since $\tilde{x}(m)$ is independent of $w(v)$ for $mT < v$, (D.36) becomes

$$V_x^u(t) = \Phi(t, mT) V_x^u(m) \Phi'(t, mT) + \int_{mT}^t \Phi(t, v) \Gamma(v) V_w(v) \Gamma'(v) \\ \times \Phi'(t, v) dv. \quad (D.37)$$

Appendix E

CONDITIONS FOR THE OPTIMAL PREFILTER

From Fig. 1.1, ignoring realizability and assuming $n_2(t_k)=0$ at first, one can write

$$y(t) = \int_{T_i}^{\infty} F(t,v) r(v) dv \quad (E.1)$$

and

$$d^O(t) = \sum_{k=l_i}^{\infty} G(t,t_k) y(t_k), \quad (E.2)$$

where T_i and l_i denote the initial conditions. Let $d(t)=[d_1(t), d_2(t), \dots, d_p(t)]$; and $d^O(t)=[d_1^O(t), d_2^O(t), \dots, d_p^O(t)]$. Then given $R_d(t,\tau)$, $R_{dr}(t,\tau)$, $R_r(t,\tau)$, and $G(t,t_k)$, the problem is to minimize

$$T_i = \int_{\infty}^{\infty} q_i(t) E \{ [d_i(t) - d_i^O(t)]^2 \} dt \quad (E.3)$$

for $i=1,2,\dots,p$, where $q_i(t)$ is a nonnegative weighting function. Let

ξ_{dia} = A column matrix consisting of the diagonal terms of the square matrix ξ ,

and

$Q(t)$ = A diagonal matrix with its i th diagonal element equal to q_i .

Thus, the problem is to minimize every element of J_{dia} , where

$$\begin{aligned} J_{dia} &= \left[\int_{-\infty}^{\infty} Q(t) E \{ [d(t) - d^0(t)] [d(t) - d^0(t)]' \} dt \right]_{dia} \\ &= \int_{-\infty}^{\infty} Q(t) [R_d(t, t) - R_{dd^0}(t, t) - R_{d^0d}(t, t) + R_{d^0d^0}(t, t)]_{dia} dt \end{aligned} \quad (E.4)$$

with R_d , R_{dr} , R_r , and G given. From (E.2), it follows that

$$R_{d^0d^0}(t, t) = \sum_{k=\ell_i}^{\infty} \sum_{j=\ell_i}^{\infty} G(t, t_k) R_Y(t_k, t_j) G'(t, t_j). \quad (E.5)$$

Through the use of (E.1), $R_Y(t_k, t_j)$ is given by

$$R_Y(t_k, t_j) = \int_{T_i}^{\infty} \int_{T_i}^{\infty} F(t_k, \alpha) R_r(\alpha, \beta) F'(t_j, \beta) d\alpha d\beta. \quad (E.6)$$

Again, from (E.1) and (E.2), it follows that

$$R_{dd^0}(t, t) = \sum_{k=\ell_i}^{\infty} R_{dy}(t, t_k) G'(t, t_k) \quad (E.7)$$

and

$$R_{dy}(t, t_k) = \int_{T_i}^{\infty} R_{dr}(t, \beta) F'(t_k, \beta) d\beta. \quad (E.8)$$

Furthermore, since $[d(t) \cdot (d^0(t))'] = [d^0(t) d'(t)]'$,

$$[E\{d(t) (d^0(t))'\}]_{dia} = [E\{d^0(t) d'(t)\}]_{dia} \quad (E.9)$$

By the use of Eqs. (E.5) through (E.9), (E.4) can be written as

$$\begin{aligned} J_{dia} = & \int_{-\infty}^{\infty} Q(t) [R_d(t, t) - 2 \sum_{k=\ell_i}^{\infty} \int_{T_i}^{\infty} R_{dr}(t, \beta) F'(t_k, \beta) G'(t, t_k) d\beta \\ & + \sum_{j=\ell_i}^{\infty} \sum_{k=\ell_i}^{\infty} G(t, t_k) \int_{T_i}^{\infty} \int_{T_i}^{\infty} F(t_k, \alpha) R_r(\alpha, \beta) F'(t_j, \beta) d\alpha d\beta \\ & \times G'(t, t_j)]_{dia} dt. \end{aligned} \quad (E.10)$$

In order to use the method of calculus of variations, let

$$F(t, \tau) = F_0(t, \tau) + \varepsilon \delta F(t, \tau), \quad (E.11)$$

where F_0 denotes the optimum prefilter. $\delta F(t, \tau)$ denotes the variation, and ε is an arbitrary quantity. Substituting (E.11) into (E.10), we have (after some manipulation),

$$J_{dia} = \int_{-\infty}^{\infty} Q(t) [P_0(t) + \varepsilon P_1(t) + \varepsilon^2 P_2(t)]_{dia} dt, \quad (E.12)$$

where

$$\begin{aligned}
 P_0(t) = & R_d(t, t) - \sum_{j=\ell_i}^{\infty} \int_{T_i}^{\infty} R_{dr}(t, \beta) F'_0(t_j, \beta) G'(t, t_j) d\beta \\
 & + \sum_{k=\ell_i}^{\infty} \sum_{j=\ell_i}^{\infty} G(t, t_k) \left[\int_{T_i}^{\infty} \int_{T_i}^{\infty} F_0(t_k, \alpha) R_r(\alpha, \beta) F'_0(t_j, \beta) \right. \\
 & \left. \times d\alpha d\beta \right] \cdot G'(t, t_j), \quad (E.13)
 \end{aligned}$$

$$\begin{aligned}
 P_1(t) = & \sum_{k=\ell_i}^{\infty} \sum_{j=\ell_i}^{\infty} G(t, t_k) \left[\int_{T_i}^{\infty} \int_{T_i}^{\infty} F_0(t_k, \alpha) R_r(\alpha, \beta) F'(t_j, \beta) \right. \\
 & \left. \times d\alpha d\beta \right] G'(t, t_j) + \sum_{k=\ell_i}^{\infty} \sum_{j=\ell_i}^{\infty} G(t, t_k) \int_{T_i}^{\infty} \int_{T_i}^{\infty} F(t_k, \alpha) \\
 & \times R_r(\alpha, \beta) F'_0(t_j, \beta) d\alpha d\beta G'(t, t_j) - 2 \sum_{k=\ell_i}^{\infty} \int_{T_i}^{\infty} R_{dr}(t, \beta) \\
 & \times \delta F'(t_k, \beta) G'(t, t_k) d\beta, \quad (E.14)
 \end{aligned}$$

and

$$\begin{aligned}
 P_2(t) = & \sum_{k=\ell_i}^{\infty} \sum_{j=\ell_i}^{\infty} G(t, t_k) \int_{T_i}^{\infty} \int_{T_i}^{\infty} \delta F(t_k, \alpha) R_r(\alpha, \beta) \delta F'(t_j, \beta) \times \\
 & d\alpha d\beta G'(t, t_j). \quad (E.15)
 \end{aligned}$$

A necessary condition for F_0 to be the optimum prefilter is

$$\frac{\partial J_{\text{dia}}}{\partial \varepsilon} = \int_{-\infty}^{\infty} Q(t) [P_1(t)]_{\text{dia}} dt = 0 \quad (\text{E.16})$$

Since $Q(t)$ is nonnegative and $[P_2(t)]_{\text{dia}}$ represents the variance functions of the outputs of the postfilter if $\delta F(t, \tau)$ is the prefilter,

$$\frac{\partial^2 J_{\text{dia}}}{\partial \varepsilon^2} = 2 \int_{-\infty}^{\infty} Q(t) [P_2(t)]_{\text{dia}} dt > 0. \quad (\text{E.17})$$

Thus, (E.16) is both necessary and sufficient to specify the optimum prefilter. Substitute (E.14) into (E.16). Since the diagonal elements of the first term on the right-hand side of (E.14) are equal to those of the second term, we have

$$\begin{aligned} \sum_{j=\ell_i}^{\infty} \int_{T_i}^{\infty} Q(t) \{ \int_{T_i}^{\infty} [\sum_{k=\ell_i}^{\infty} G'(t, t_k) \int_{T_i}^{\infty} F_0(t_k, \alpha) R_r(\alpha, \beta) d\alpha - R_{dr}(t, \beta)] \\ \times \delta F'(t_j, \beta) G'(t, t_j) d\beta \}_{\text{dia}} dt = 0 \end{aligned} \quad (\text{E.18})$$

for an arbitrary $\delta F(t, \beta)$. If all the quantities in (E.18) are scalars, rearranging the terms in (E.18) to obtain

$$\begin{aligned} \sum_{j=\ell_i}^{\infty} \int_{T_i}^{\infty} \delta f'(t_j, \beta) \{ \int_{-\infty}^{\infty} q(t) g(t, t_j) [R_{dr}(t, \beta) - \sum_{k=\ell_i}^{\infty} g(t, t_k) \cdot \times \\ \int_{T_i}^{\infty} f_0(t_k, \alpha) R_r(\alpha, \beta) d\alpha] dt \} d\beta = 0, \end{aligned} \quad (\text{E.19})$$

where we have used lower case f , g , and q to denote F , G , and Q , respectively, in order to emphasize that they are scalars. Eq. (E.19) is equivalent to Eq. (B.8) obtained by Kellog (32).

Let us investigate the case in which the signal and noises are stationary, and the filters are constrained to be time invariant, i.e.,

$$R.(t, \tau) = R.(t - \tau) \quad (E.20)$$

$$F(t, \tau) = F(t - \tau) \quad (E.21)$$

and

$$G(t, \tau) = G(t - \tau) \quad (E.22)$$

Also, assume that the sampling process is periodic; i.e., $t_k = kT$, and let

$$Q(t) = \begin{cases} (1/T) \cdot I & \text{if } 0 \leq t < T \\ 0 & \text{otherwise,} \end{cases} \quad (E.23)$$

where I is the identity matrix, and T is the sampling period. Furthermore, assume $T_i, \ell_i \rightarrow -\infty$. From these assumptions, a solution could be obtained from (E.18) directly, but in order to simplify later development, let us start from (E.4). By use of (E.23) and (E.5), we have

$$\int_{-\infty}^{\infty} Q(t) R_d(t, t) dt = (1/T) \int_0^T \sum_{k=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} G(t - kT) R_y(k - j) \times \\ G'(t - jT) dt. \quad (E.24)$$

Let $v=t-jT$ and $\ell=k-j$; after interchanging the order of integration and summation, we have

$$\begin{aligned} \int_0^T \sum_{k=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} G(k-kT) R_Y(k-j) G'(t-jT) dt \\ = \sum_{\ell=-\infty}^{\infty} \left[\sum_{j=-\infty}^{\infty} \int_{-jT}^{-(j-1)T} G(v-\ell T) R_Y(\ell) G'(v) dv \right]. \end{aligned} \quad (\text{E.25})$$

The inner summation is equal to $\int_{-\infty}^{\infty} G(v-\ell T) R_Y(\ell) G'(v) dv$; thus,

$$\int_{-\infty}^{\infty} Q(t) R_{dd}^o(t, t) dt = (1/T) \sum_{\ell=-\infty}^{\infty} \int_{-\infty}^{\infty} G(v-\ell T) R_Y(\ell) G'(v) dv, \quad (\text{E.26})$$

where

$$R_Y(\ell) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(\alpha) R_r(\ell T + \alpha - \beta) F'(\beta) d\alpha d\beta \quad (\text{E.27})$$

from (E.6). Similarly,

$$\int_{-\infty}^{\infty} Q(t) R_{dd}^o(t, t) dt = (1/T) \int_{-\infty}^{\infty} R_{dy}(v) G'(v) dv \quad (\text{E.28})$$

where

$$R_{dy}(v) = \int_{-\infty}^{\infty} R_{dr}(\beta + v) F'(\beta) d\beta. \quad (\text{E.29})$$

From Eqs. (E.26) through (E.29), and the fact that

$[R_{dd}^o(t,t)]_{dia} = [R_d^o(t,t)]_{dia}$, Eq. (E.4) becomes

$$J_{dia} = [R_d(0) - (2/T) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} R_{dr}(\beta+v) F'(\beta) G'(v) dv d\beta +$$

$$(1/T) \sum_{\ell=-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} G(v-\ell T) F(\alpha) R_r(\ell T + \beta - \alpha) F'(\beta) \times$$

$$G'(v) dv d\alpha d\beta]_{dia} \quad (E.30)$$

Following the same procedure as for the time-variant systems, the optimum time-invariant prefilter is found to be specified by the equation

$$[\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} R_{dr}(\beta+v) \delta F'(\beta) G'(v) dv d\beta -$$

$$\sum_{\ell=-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} G(v-\ell T) F_o(\alpha) R_r(\ell T + \alpha - \beta) F'(\beta) G'(v) \times$$

$$dv d\alpha d\beta]_{dia}$$

$$= 0 \quad (E.31)$$

With a single sampled-data link, and a single input and output, all the quantities become scalars, and (E.31) reduces to

$$\int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} R_{dr}(\beta+v) g(v) dv - \int_{-\infty}^{\infty} \sum_{l=-\infty}^{\infty} \left[\int_{-\infty}^{\infty} g(v-lT) g(v) dv \right] \times \right. \\ \left. R_r(lT+\alpha-\beta) f_0(\alpha) d\alpha \right] \delta f(\beta) d\beta = 0. \quad (E.32)$$

Since $\delta f(\beta)$ is an arbitrary function in β , it is both sufficient and necessary that

$$\int_{-\infty}^{\infty} R_{dr}(\beta+v) g(v) dv - \int_{-\infty}^{\infty} \sum_{l=-\infty}^{\infty} \left[\int_{-\infty}^{\infty} g(v-lT) g(v) dv \right] R_r(lT+\alpha-\beta) \\ \times f_0(\alpha) d\alpha = 0. \quad (E.33)$$

When n_2 is absent (see Fig. 1.1), there is no need to impose a power constraint on the prefilter since one can add a gain factor K to the prefilter and the inverse $1/K$ to the postfilter without influencing the performance of the system. When $n_2 \neq 0$, (E.2) becomes

$$d^0(t) = \sum_{k=-\infty}^{\infty} G(t, t_k) [Y(t_k) + n_2(t_k)]. \quad (E.34)$$

Then

$$R_d^0(t, t) = \sum_{k=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} G(t, t_k) [R_Y(t_k, t_j) + R_{Yn_2}(t_k, t_j) \\ + R_{n_2 Y}(t_k, t_j) + R_{n_2}(t_k, t_j)] G'(t, t_j) \quad (E.35)$$

and

$$R_{dd}^o(t, t) = \sum_{k=-\infty}^{\infty} [R_{dy}(t, t_k) + R_{dn_2}(t, t_k)] G'(t, t_k). \quad (E.36)$$

Substituting (E.35) and (E.36) into (E.4), and using (E.9) to obtain

$$\begin{aligned} J_{dia} = & \int_{-\infty}^{\infty} Q(t) \{ R_d(t, t) + \sum_{k=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} G(t, t_k) [R_Y(t_k, t_j) \\ & + R_{yn_2}(t_k, t_j) + R_{n_2Y}(t_k, t_j) + R_{n_2}(t_k, t_j)] G'(t, t_j) \\ & - 2 \sum_{j=-\infty}^{\infty} [R_{dy}(t, t_j) + R_{dn_2}(t, t_j)] G'(t, t_j) \}_{dia} dt. \end{aligned} \quad (E.37)$$

From here on, the problem under consideration will be confined to the special case of stationary inputs, time-invariant systems, periodic sampling, steady-state conditions, and $Q(t)$ given by (E.23). In this case, (E.37) reduces to

$$\begin{aligned} J_{dia} = & \{ R_d(0) - (2/T) \int_{-\infty}^{\infty} [R_{dy}(v) + R_{dn_2}(v)] G'(v) dv + \\ & \sum_{\ell=-\infty}^{\infty} (1/T) \int_{-\infty}^{\infty} G(v-\ell T) [R_Y(\ell) + R_{yn_2}(\ell) + R_{n_2Y}(\ell) + \\ & R_{n_2}(\ell)] G'(v) dv \}_{dia}. \end{aligned} \quad (E.38)$$

Let us look at two different situations:

1. $n_2(kT)$ is an independent noise; then

$$R_{dn_2}(v) = R_{yn_2}(\ell T) = R_{n_2Y}(\ell T) = 0. \quad (E.39)$$

It is clear that if there is no power constraint on the prefilter in this case, one can simply add a gain factor of $K \rightarrow \infty$ to the prefilter, and add $1/K$ to the postfilter. Then the effect of n_2 on the signal disappears because of an infinite signal-to-noise ratio, whereas the effect of other factors remain unchanged. Thus, it is necessary to have a power constraint on the prefilter in this case.

This constraint is given by

$$P_Y = [R_Y(0)]_{\text{dia}} = \left[\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(\alpha) R_r(\beta - \alpha) F'(\beta) d\alpha d\beta \right]_{\text{dia}}. \quad (E.40)$$

Then the function to be minimized is

$$J_{\text{dia}}^0 = J_{\text{dia}} + \left[\lambda \cdot \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(\alpha) R_r(\beta - \alpha) F'(\beta) d\alpha d\beta \right]_{\text{dia}}, \quad (E.41)$$

where J_{dia} is given by (E.38), and λ is the Lagrange multiplier. Substituting (E.27) and (E.29) into (E.38), then substituting the result into (E.41) to obtain

$$\begin{aligned}
J_{\text{dia}}^0 = & [R_d(0) - (2/T) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} R_{dr}(\beta+v) F'(\beta) G'(v) dv d\beta + \\
& (1/T) \sum_{\ell=-\infty}^{\infty} \int_{-\infty}^{\infty} G(v-\ell T) \{ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(\alpha) R_r(\ell T+\beta-\alpha) F'(\beta) \times \\
& d\alpha d\beta + R_{n_2}(\ell) G'(v) dv + \lambda \cdot \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(\alpha) R_r(\beta-\alpha) \times \\
& F'(\beta) d\alpha d\beta \}_{\text{dia}}.
\end{aligned} \tag{E.42}$$

Using the method we used before for the case when $n_2=0$, we find that the optimal prefilter is specified by the equation

$$\begin{aligned}
& \left[\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} R_{dr}(\beta+v) \delta F'(\beta) G'(v) dv d\beta - \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \sum_{\ell=-\infty}^{\infty} \int_{-\infty}^{\infty} G(v-\ell T) \times \right. \\
& F_0(\alpha) R_r(\ell T+\beta-\alpha) \delta F'(\beta) G'(v) dv - T \lambda F_0(\alpha) R_r(\beta-\alpha) \times \\
& \left. \delta F'(\beta) \right] d\alpha d\beta \}_{\text{dia}} = 0,
\end{aligned} \tag{E.43}$$

where λ is to be chosen to satisfy (E.40).

2. $n_2(kT)$ represents quantization noise specified by (see Fig. 1.1)

$$s(k) = a_1 y(k) + n_2(k) \tag{E.44}$$

where y is the input and s is the output of the quantizer, and a_1 is given by Eq. (1.7) and is assumed to be absorbed in the prefilter. Then the quantization noise is uncorrelated with (but not independent of) the input signal and the observation noise; i.e., (refer to Fig. 1.1)

$$R_{dn_2}(v) = R_{n_2 Y}(\ell T) = R_{n_1 n_2} = 0. \quad (E.45)$$

Assume that the quantization noise can be approximately characterized by a discrete white noise (refer to Section 1.5 of this thesis); i.e.,

$$R_{n_2}(k-j) = R_{n_2}(0) \delta_{kj} \quad (E.46)$$

Let y_n and s_n denote the normalized input and output of the quantizer, and let y_n^i and s_n^i denote the normalized i th end point of the input ranges and the normalized output level for $y_n^{i-1} \leq y_n < y_n^i$, respectively. Then

$$\begin{aligned} R_{n_2}(k-j) &\stackrel{\sim}{=} R_{n_2}(0) \delta_{kj} \\ &= R_Y(0) \int_{y_n^{i-1}}^{y_n^i} (s_n^i - y_n)^2 P_r(y_n) dy_n \end{aligned} \quad (E.47)$$

since both the actual quantization steps and the output levels are linear proportional to $[R_Y(0)]^{1/2}$

(8; Eq. (11_a)). The $P_r(y_n)$ in (E.47) represents the probability density function of y_n . From (E.47), it follows that

$$R_{n_2}(0)/R_y(0) = \int_{y_n^{i-1}}^{y_n^i} (s_n^i - y_n)^2 P_r(y_n) dy_n = c_2. \quad (\text{E.48})$$

Thus, it is seen that with the simplified assumption on the quantization noise, the noise-to-signal ratio depends only on the quantizer and not on the prefilter. The power of the prefilter in this case is to be optimized instead of being constrained to be a certain value. Substituting (E.47) and (E.45) into (E.38), we have

$$\begin{aligned} J_{\text{dia}} = & [R_d(0) - (2/T) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} R_{dr}(\beta+v) F'(\beta) G'(v) dv d\beta + \\ & (1/T) \sum_{\ell=-\infty}^{\infty} \int_{-\infty}^{\infty} G(v-\ell T) \{ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(\alpha) R_r(\ell T + \beta - \alpha) F'(\beta) \times \\ & d\alpha d\beta G'(v) dv + c_2 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} G(v) F(\alpha) R_r(\beta - \alpha) \times \\ & F'(\beta) G'(v) dv d\alpha d\beta \}_{\text{dia}}, \end{aligned} \quad (\text{E.49})$$

where c_2 is given by (E.48). (E.49) is similar to (E.42). Thus, the solution can be written down directly in the form

$$\begin{aligned}
& \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} R_{dr}(\beta+v) \delta F'(\beta) G'(v) dv d\beta - \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left[\sum_{l=-\infty}^{\infty} \int_{-\infty}^{\infty} G(v-lT) \times \right. \\
& \left. F_O(\alpha) R_r(lT+\beta-\alpha) + c_2 T \int_{-\infty}^{\infty} G(v) F_O(\alpha) R_r(\beta-\alpha) \right] \times \\
& \left. \delta F'(\beta) G'(v) dv d\alpha d\beta \right\}_{dia} = 0. \quad (E.50)
\end{aligned}$$

For the case when both quantization noise and an additive independent channel noise are present, one will need to impose a power constraint on the prefilter in general. In this case, the quantization noise can be approximated by an independent white noise since $R_y(0)$ in (E.47) is preassigned a certain value.

REFERENCES

REFERENCES

1. Aasnaes, H. B., "Estimation and Identification in Discrete-Time System Using the Innovations Approach," Technical Report No. 7050-20, Sel-73-001, Information System Laboratory, Stanford Electronics Laboratories, Stanford University, Stanford, California, May 1971.
2. Berger, T., Rate Distortion Theory, A Mathematical Basis for Data Compression, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1971.
3. Berger, T., and Tufts, D. W., "Optimum Pulse Amplitude Modulation, Part I: Transmitter-Receiver Design and Bounds from Information Theory," IEEE Transactions on Information Theory, Vol. IT-13, April 1967, pp. 196-208.
4. Bode, H. W., and Shannon, C. E., "A Simplified Derivation of Linear Least Square Smoothing and Prediction Theory," Proceedings IRE, Vol. 38, April 1950, pp. 417-425.
5. Brown, W. M., "Optimum Prefiltering of Sampled Data," IRE Transactions on Information Theory, (Correspondence), Vol. IT-7, October 1961, pp. 269-270.
6. Busgang, J. J., "Cross Correlation Functions of Amplitude-Distorted Gaussian Signals," Technical Report 216, Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Mass., March 26, 1952.
7. Chan, D., and Donaldson, R. W., "Optimum Pre- and Postfiltering of Sampled Signals with Application to Pulse Modulation and Data Compression Systems," IEEE Transactions on Communication Technology, Vol. Com-19, April 1971, pp. 141-157.
8. _____, and _____, "Correlation Functions and Reconstruction Error for Quantized Gaussian Signals Transmitted over Discrete Memoryless Channels," IEEE Transactions on Information Theory, Vol. IT-18, No. 4, July 1972, pp. 519-523.

9. Chang, S. S. L., Synthesis of Optimum Control Systems, McGraw-Hill Book Company, New York, 1961.
10. Costas, J. P., "Coding With Linear Systems," Proceedings IRE, Vol. 40, September 1952, pp. 1101-1103.
11. Curry, R. E., Estimation and Control with Quantized Measurement, Cambridge, Mass., MIT Press, 1970.
12. Derusso, P. M., "Optimum Linear Filtering of Signals Prior to Sampling," AIEE Transactions, Vol. 79, Part II, Applications and Industry, January 1961, pp. 549-555.
13. Fraser, D. C., "On the Application of Optimum Linear Smoothing Techniques to Linear and Nonlinear Dynamic Systems," Ph.D. Thesis, Massachusetts Institute of Technology, January 1967.
14. Frost, P. A., and Kailath, T., "An Innovations Approach to Least-Squares Estimation--Part III: Nonlinear Estimation on White Gaussian Noise," IEEE Transactions on Automatic Control, Vol. AC-16, No. 3, June 1971, pp. 217-226.
15. Gallagher, R. G., Information Theory and Reliable Communication, John Wiley & Sons, Inc., New York, 1969.
16. Gevers, M. R., "Structural Properties of Realizations of Discrete Time Markovian Processes," Technical Report No. 7050-19, SU-SEL-72-016, Information System Laboratory, Stanford Electronics Laboratories, Stanford University, Stanford, California, May 1972.
17. Goblick, T. J., Jr., and Holsinger, J. L., "Analog Source Digitization: A Comparison of Theory and Practice," IEEE Transactions on Information Theory (correspondence), Vol. IT-13, April 1967, pp. 323-326.
18. Goodman, L. M., "Optimum Sampling and Quantization Rate," Proceedings IEEE (letter), Vol. 54, January 1966, pp. 90-92.

19. Goodman, L. M., and Drouilhet, P. R., Jr., "Asymptotically Optimum Pre-emphasis and De-emphasis Networks for Sampling and Quantizing," Proceedings IEEE (letter) Vol. 54, May 1966, pp. 795-796.
20. Hayes, J. F., "Optimum Prefiltering and Postfiltering of Sampled Waveforms," IEEE Transactions on Communication Technology, Vol. COM-19, No. 4, August 1971, pp. 532-535.
21. Huang, J. J. Y., and Schultheiss, P. M., "Block Quantization of Correlated Gaussian Random Variables," IEEE Transactions on Communication Systems, Vol. CS-11, September 1963, pp. 289-296.
22. Jazwinski, A. H., Stochastic Processes and Filtering Theory, N.Y., Academic Press, 1970.
23. Kailath, T., "An Innovations Approach to Least-Squares Estimation--Part I: Linear Filtering in Additive White Noise," IEEE Transactions on Automatic Control, Vol. AC-13, No. 6, December 1968, pp. 646-655.
24. _____, "The Innovations Approach to Detection and Estimation Theory," Proceedings IEEE, Vol. 58, No. 5, May 1970, pp. 680-695.
25. _____, and Frost, P., "An Innovations Approach to Least-Squares Estimation Part II: Linear Smoothing in Additive White Noise," IEEE Transactions on Automatic Control, Vol. AC-13, No. 6, December 1968, pp. 655-660.
26. _____, and Geesey, R. A., "An Innovations Approach to Least-Squares Estimation--Part IV: Recursive Estimation Given Lumped Covariance Functions," IEEE Transactions on Automatic Control, Vol. AC-16, No. 6, December 1971, pp. 720-726.
27. _____, and _____, "Covariance Factorization--An Explication Via Examples," Proceedings 2nd Asilomar Conference on Circuits and Systems, Monterey, California, November 1968.

28. Kalman, R. E., "A New Approach to Linear Filtering and Prediction Problems," Journal of Basic Engineering, Transaction ASME, Vol. 82D, March 1960_a, pp. 34-45.
29. _____, and Bucy, R. S., "New Results in Linear Filtering and Prediction Theory," Journal of Basic Engineering, Transaction ASME, Vol. 83D, March 1961, pp. 95-108.
30. Katzenelson, J., "On Errors Introduced by Combined Sampling and Quantization," IRE Transactions on Automatic Control, Vol. AC-7, April 1962, pp. 58-68.
31. Kellog, W. C., "Information Rates in Sampling and Quantization," IEEE Transactions on Information Theory, Vol. IT-13, No. 3, July 1967, pp. 506-511.
32. _____, "Numerical Operations on Random Functions," Ph.D Dissertation, Division of Engineering and Applied Physics, Harvard University, Cambridge, Massachusetts, February 1966.
33. _____, and Chang, S. S. L., "Jointly Optimized Realizable Linear Pre- and Postfilters for Systems with Samplers," Proceedings IEEE (correspondence), Vol. 53, June 1965, pp. 623-624.
34. Kolmogorov, A. N., "Interpolation and Extrapolation of Stationary Random Sequences," Bull. Academic Sci. USSR, Math. Ser., Vol. 5, 1941, a translation has been published by the RAND Corp., Santa Monica, Calif., as Memo. RM-3090=PR.
35. Kuo, B. C., Discrete Data Control System, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1970.
36. Max, J., "Quantizing for Minimum Distortion," IRE Transactions on Information Theory, Vol. IT-6, March 1960, pp. 7-12.
37. Meditch, J. S., Stochastic Optimal Linear Estimation and Control, McGraw Hill Book Company, New York, 1969.

38. Mehra, R. K., "On Optimal and Suboptimal Linear Smoothing," Proceedings 1968 National Electronics Conference, pp. 119-124.
39. Moore, J. B., and Anderson, B. D. O., "Spectral Factorization of Time Varying Covariance Functions: The Singular Case," Math. Systems Theory, Vol. 4, No. 1, 1970, pp. 10-23.
40. _____, and Hetrakul, P., "Optimal Demodulation of PAM Signals," IEEE Transactions on Information Theory, Vol. IT-19, No. 2, March 1973, pp. 188-196.
41. Papoulis, A., Probability, Random Variables, and Stochastic Processes, McGraw-Hill Book Company, New York, 1965.
42. Petersen, D. P., "Reexamination of the Prefiltering Problem in a State Variable Formulation," IEEE Transactions on Information Theory, Vol. IT-16, No. 4, July 1970, pp. 480-483.
43. _____, and Middleton, D., "Linear Interpolation, Extrapolation, and Prediction of Random Space-Time Fields with a Limited Domain of Measurement," IEEE Transaction on Information Theory, Vol. IT-11, No. 1, January 1965, pp. 18-30.
44. _____, and Theodore, N. T., "Spectral Analysis and Data Processing for Multivariable, Multidimensional Fields," Final Report EE-178 (70) DC-114, prepared for the United States Department of Commerce, August 14, 1970, Chapter 6.
45. Pierre, D. A., "Multiplier Algorithm for Nonlinear Programming," Presented at the VIII International Symposium on Mathematical Programming, August 27-31, 1973.
46. _____, Optimization Theory with Applications, John Wiley & Sons, Inc., New York, 1969.
47. Ragazzini, J. R., and Franklin, G. F., Sampled Data Control Systems, McGraw-Hill Book Company, New York, 1958.

48. Rauch, H. E., "Solutions to the Linear Smoothing Problem," IEEE Transactions on Automatic Control, Vol. AC-8, No. 4, October 1963, pp. 371-372.
49. Robbins, H. M., "An Extension of Wiener Filter Theory to Partly Sampled Systems," IRE Transactions on Circuit Theory, Vol. CT-6, December 1959, pp. 362-370.
50. Ruchkin, D. S., "Linear Reconstruction of Quantized and Sampled Random Signals," IRE Transactions on Communication Systems, Vol. CS-9, December 1961, pp. 350-355.
51. Sage, A. P., and Melsa, J. L., Estimation Theory with Applications to Communications and Control, McGraw-Hill Book Company, New York, 1968.
52. Saucedo, R., and Schiring, E. E., Introduction to Continuous and Digital Control Systems, The Macmillan Company, New York, 1968.
53. Spilker, J. J., Jr., "Theoretical Bounds on the Performance of Sample Data Communications Systems," IRE Transactions on Circuit Theory, Vol. CT-7, September 1960, pp. 335-341.
54. Steiglitz, K., "Transmission of An Analog Signal Over a Fixed Bit Rate Channel," IEEE Transactions on Information Theory, Vol. IT-12, October 1966, pp. 469-474.
55. Taub, H., and Schilling, D. L., Principle of Communication Systems, McGraw-Hill Book Company, New York, 1971.
56. Thomas, J. B., An Introduction to Statistical Communication Theory, John Wiley & Sons, Inc., New York, 1969.
57. Tretter, S. A., and Steiglitz, K., "Some Properties of Minimum Mean-Squares Error Filters for Sampled-Data Reconstruction," IEEE International Convention Record, Part 7, New York, March 1965, pp. 134-146.

58. Van Trees, H. L., Detection, Estimation, and Modulation Theory, Part I, John Wiley & Sons, Inc., New York, 1968.
59. Whittle, P., Prediction and Regulation, Princeton, N. J., Van Nostrand, 1963.
60. Widrow, B., "Statistical Analysis of Amplitude Quantized and Sampled Data Systems," AIEE Transactions, Vol. 79, Part 2, Applications and Industry, January 1961, pp. 555-568.
61. Wiener, N., The Extrapolation, Interpolation, and Smoothing of Stationary Time Series with Engineering Applications, John Wiley & Sons, Inc., New York, 1949.

MONTANA STATE UNIVERSITY LIBRARIES
3 1762 10011695 1

D378
Y35 Yeh, Yern
cop.2 A state-space approach
to the design of pre-
and postfilters in a
sampled-data...

NAME AND ADDRESS

D378
Y35
cop.2