NONGAUSSIAN ESTIMATION USING A MODIFIED GAUSSIAN SUM ADAPTIVE FILTER

by

Mauro J. Caputi

Dissertation submitted to the Faculty of the

Virginia Polytechnic Institute and State University

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Electrical Engineering

APPROVED:

Richard L. Moose

Richard L. Moose, Chairman

William T. Baumann

William T. Baumann

Hugh F. VanLandingham

Charles E. Nunnally

Charles E. Nunnally

Werner E. Kohler

Werner E. Kohler

April, 1991

Blacksburg, Virginia

NONGAUSSIAN ESTIMATION USING A MODIFIED GAUSSIAN SUM ADAPTIVE FILTER

by

Mauro J. Caputi

Richard L. Moose, Chairman

Electrical Engineering

(ABSTRACT)

This investigation is concerned with effective state estimation of a system driven by an unknown nonGaussian input with additive white Gaussian noise, and observed by measurements containing feedthrough of the same nonGaussian input and corrupted by additional white Gaussian noise. A Gaussian sum (GS) approach has previously been developed [6-8] which can cope with the nonGaussian nature of the input signal. Due to a serious growing memory problem in this approach, a modified Gaussian sum (MGS) estimation technique is developed that avoids the growing memory problem while providing effective state estimation. Several differences between the MGS and GS algorithms are examined. An MGS adaptive filter is derived for a general system and a modal system, with simulation examples performed using a nonGaussian input signal. The modal system simulation results are compared to those produced from an augmented Kalman filter based on an augmented modal system model assuming a narrowband Gaussian input signal. A necessary condition for effective MGS estimation is derived. Alternate estimation procedures are developed to compensate for situations when this condition is not met. Several configurations are simulated and their performance results are analyzed and compared. Two methods of monitoring and updating key parameters of the MGS filter are developed. Simulation results are analyzed to investigate the performance of these methods.

ACKNOWLEDGEMENTS

I wish to extend my sincere appreciation and thanks to my advisory committee: Dr. Richard L. Moose, Chairman, Dr. William T. Baumann, Dr. Hugh F. VanLandingham, Dr. Charles E. Nunnally, and Dr. Werner E. Kohler. Their insightful guidance, technical assistance, and moral support have made this endeavor a very pleasurable and worthwhile experience.

I am grateful to my mother and father for the support and encouragement they have lovingly provided me with. I will always remember the valuable lessons they taught me.

My wife, Joan, has made all the difference. As we traveled together on this long journey she has been a wonderful companion, a terrific cheerleader, a wise counselor, and my best friend. Throughout all the ups and downs we faced, she remained steadfast and strong. We laughed and cried together, and now we can rejoice together for an accomplishment well done. Thank you for loving me and believing in me. Thank you for being you.

Psalm 46:10

DEDICATION

This dissertation is dedicated in loving memory of our son, Jonathan David, who would have been one year old on June 8, 1991.

For Thou didst form my inward parts;
Thou didst weave me in my mother's womb.
I will give thanks to Thee, for I am fearfully and wonderfully made;
wonderful are Thy works, and my soul knows it very well.
My frame was not hidden from Thee, when I was made in secret,
and skillfully wrought in the depths of the earth.
Thine eyes have seen my unformed substance;
and in Thy book they were all written,
the days that were ordained for me,
when as yet there was not one of them.

Psalm 139:13-16

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

1.0 INTRODUCTION

This investigation is concerned with effective state estimation of a system driven by an unknown nonGaussian input with additive white Gaussian noise, and observed by measurements containing feedthrough of the same nonGaussian input and corrupted by additional white Gaussian noise. The motivation behind this work is the problem of robust modal control of large, multiple input multiple output (MIMO), heavily damped structures. The scope of the problem involves system identification of the system parameters, optimal estimation of the system states, and robust control of the heavily damped modes of the system. This investigation will deal specifically with the state estimation portion of the problem.

In an earlier study [1], Kalman filter estimation techniques [2], [3] are developed for a lightly damped, simply supported plate. The input forcing function is assumed to be narrowband Gaussian, which is adequately modeled by passing white Gaussian noise through a linear, time-invariant filter. Thus, knowledge of the power spectral density of the narrowband process is sufficient for the design of an adequate estimator. Since the plant and measurement equations for this system are linear, and the inputs and plant and measurement disturbances are Gaussian, the Kalman filter will provide the optimum minimum mean-square error estimate of the system states [4].

Although the Gaussian assumption for modeling many types of inputs and noise processes is valid in a wide range of applications, in practice it may not be a good assumption for some signal models. For instance, also considered in [1] are large, heavily damped structures modeled by a plant with complex modes. The input to the plant is unknown, but its frequency characteristics are known and the input is always present. The signal model chosen to represent this input is a stochastic FM signal, generated by frequency modulating a sinusoid with a Gaussian process. This stochastic FM signal has a highly nonGaussian probability density function. The power spectral density of a signal of this type appears similar to that of a Gaussian narrowband signal. An estimator design based solely on the power spectral density of a nonGaussian signal assumed to be a Gaussian narrowband signal may provide very poor estimates.

Estimation techniques need to be developed that can cope effectively with the nonGaussian nature of certain signals. One such approach is the Gaussian sum technique, developed by Alspach [6-8]. The density function of each nonGaussian process of the system is approximated by a weighted sum of Gaussian density functions. The conditional density of the state given the available measurement sequence, necessary in the Kalman filter development, is updated using the Gaussian sum approximations and Bayes' rule.

A serious limitation in Alspach's approach is that the number of Gaussian terms used to approximate the density functions increases at each time iteration. A modified approach is required to alleviate this limitation. A modified estimation structure is developed based on an adaptive Kalman filter scheme first presented by Magill [16], and extended by Moose [23]. Essentially, a parameter vector is used to uniquely describe each Gaussian term in the estimator. The parameter vector is restricted to be randomly chosen from the same finite set of known values at each iteration. By using a nonGaussian signal model in conjunction with the modified formulation of the Gaussian sum estimator, the number of Gaussian terms at each iteration of the estimator will be fixed, thereby avoiding the growing memory problem.

The modified estimation algorithm is termed the *modified Gaussian sum adaptive filter* and forms the basis for the nonGaussian state estimation in this investigation. An algorithm similar to this one that does not make reference to the Gaussian sum characteristic is known in the literature as the parallel processing algorithm [4], or the multiple model algorithm [5]. Both the Gaussian sum and modified Gaussian sum algorithms have been implemented in a variety of important engineering applications such as phase and frequency estimation [36], geophysical field navigation [37], maneuvering target tracking [24], and specification of route widths for air traffic controllers [20].

Several contributions to the field of applied estimation theory are made from this investigation. These include:

1. The development of a modified Gaussian sum algorithm with nongrowing memory based on a nonGaussian signal model with a Gaussian sum probability density function. Parameters from this model are used directly in the modified Gaussian sum adaptive filter structure.

2. A comparison between the Gaussian sum (GS) filter of [6] and the modified Gaussian sum (MGS) adaptive filter. The two are similar, but the comparison shows the MGS adaptive filter to be a good improvement to the GS filter.

3. An examination of a necessary condition for effective MGS estimation. This condition provides a simple test to determine if the MGS adaptive filter will work properly for a given system.

4. An alternate configuration of the MGS adaptive filter when the necessary condition of 3 above is not met. This configuration is applied in several ways, and each is evaluated on a performance basis.

5. Two methods of monitoring and updating key parameters of the MGS adaptive filter. These allow the estimator to react to changes in the input signal level which cause the signal to be nonstationary over long periods of time.

The dissertation begins with a literature review in Chapter 2. This chapter covers previous work and applications in the areas of Gaussian sum estimation, adaptive Kalman filtering, and nonGaussian estimation. Chapter 3 begins with the development of the Gaussian sum estimator. Then, a detailed description of the nonGaussian signal model and the alternate formulation of the Gaussian sum density approximation is given. A curve fitting procedure used to find the initial modified Gaussian sum parameter vector is outlined. The MGS adaptive filter is developed for a general system and a simulation example is given. Several differences between the MGS and GS algorithms are examined. Chapter 4 describes the modal system, the nonGaussian input model, and the MGS adaptive filter based on this system. A simulation example using the MGS adaptive filter is given and the results are compared to those produced from an augmented Kalman filter based on an augmented system model assuming a narrowband Gaussian input signal. A necessary condition for effective MGS estimation is derived. Alternate estimation procedures are developed to compensate for situations when this condition is not met. Several configurations are simulated and their performance results are analyzed and compared. Chapter 5 discusses two methods of monitoring and updating key parameters of the MGS adaptive filter. Simulation results are analyzed to investigate the performance of these methods. Finally, Chapter 6 gives the conclusions and outlines suggested directions for future investigations.

2.0 PREVIOUS RESEARCH AND RELATED WORK

This chapter presents a review of the literature relevant to this investigation. It begins with a discussion of the Gaussian sum estimation technique. Many examples of its use in a wide range of applications are given. Next, an adaptive Kalman filtering method is outlined along with several modifications. Several examples of the various applications of this method are presented. Finally, an overview of alternate nonGaussian filtering and its applications is discussed.

2.1 Gaussian Sum Estimation

An estimation technique, applicable to both linear systems with nonGaussian inputs and nonlinear systems with Gaussian inputs, has been developed by Alspach [6-8]. In the case of linear systems, the noise processes associated with the plant and measurement are assumed known and nonGaussian. The probability density function of each noise process is approximated by a Gaussian sum; that is, a weighted sum of Gaussian density functions. The Gaussian sum approximation is written as

$$p_{GS}(x) = \sum_{i=1}^{M} \alpha_i \mathrm{N}[\mu_i, \sigma_i^2] \tag{2.1.1}$$

where

$$\sum_{i=1}^{M} \alpha_i = 1; \quad \alpha_i \geq 0 \ \text{ for } i = 1, 2, \ldots, M \tag{2.1.2}$$

and

$$\mathrm{N}[\mu_i, \sigma_i^2] = \frac{1}{\sqrt{2\pi}\sigma_i} \, \mathrm{e}^{-\frac{1}{2}\left(\frac{x - \mu_i}{\sigma_i}\right)^2} \tag{2.1.3}$$

For sufficiently large M, any density function can be closely approximated by a Gaussian sum. As long as condition (2.1.2) holds, the Gaussian sum is always a valid density function.

The basic filtering problem is to estimate the state x_k of the system from the current measurement sequence $Z_k = \{z_1, z_2, \ldots, z_{k-1}, z_k\}$. The "best" estimate $\hat{x}_k$ is found by minimizing or maximizing a particular performance criterion, such as minimizing the mean-square error between x_k and $\hat{x}_k$. In the Bayesian approach to estimation, this requires the use of the *a posteriori*, or conditional, density function $p(x_k \mid Z_k)$.

The Gaussian sum estimator [6-8] is briefly described here and is more fully developed in chapter 3. The density functions of the initial state, the plant noise process, and measurement noise process are approximated by Gaussian sum densities of the form of (2.1.1). These densities are then directly used to form the conditional density function $p(x_k \mid Z_k)$ as a sum of Gaussian densities. An unfortunate consequence of using these Gaussian sum densities is that the number of Gaussian terms that forms $p(x_k \mid Z_k)$ increases at each stage of the estimator.

This *growing memory* problem is a serious limitation in Alspach's development. In chapter 3, a nonGaussian signal model is developed and used in conjunction with a modified Gaussian sum adaptive filter that avoids the growing memory problem.

Alspach has applied the Gaussian sum technique to a wide variety of applications. In conjunction with Sorenson, he applies the technique to linear systems with nonGaussian noise [7], and to nonlinear systems with Gaussian noise [8]. In [9], [10], and with Scharf and Abiri in [11], he addresses the problem of linear systems with Gaussian noise processes having unknown covariances. Joint identification, tracking, and prediction in a multi-target, multi-sensor environment is considered in [13], and with LaGrotta in [12]. Alspach and Sorenson [14] use the Gaussian sum technique in conjunction with proving the validity of the separation theorem for linear, nonGaussian, optimal control problems. Alspach extends this work in [15] to nonlinear systems. Here he does not use the separation theorem, but certainty equivalence control instead, producing a suboptimal control algorithm useful for off-line investigations.

Other researchers have used the Gaussian sum approach in their work. Tam and Moore [36] developed a Gaussian sum estimator using extended Kalman filters in the problem of angle demodulation. Dmitriev and Shimelevich [37] applied the technique in determining the coordinates of a moving vehicle from geophysical field measurements. Sirisena and Brown [38] applied the algorithm to probabilistic and stochastic load flow problems. Namera and Stubberud [39] used the Gaussian sum approach in solving nonlinear fixed-point prediction problems. Gauvrit [40] developed a Gaussian sum filter to track targets in clutter with unknown noise variances. Tanaka and Katayama [41] devised a robust Kalman filter that consists of the Kalman filter plus MAP estimates of the selection parameters that specify noises from the Gaussian sum densities. Their filter is actually a smoother applied to a linear system with Gaussian sum noises. Kitagawa [42] modeled sudden changes of trend or seasonal components due to the structural changes of the economic system by the presence of outliers by a nonGaussian model. He used Gaussian sum approximations of the nonGaussian densities.

2.2 Adaptive Kalman Filtering

In a typical application, the various parameters that describe the system are assumed to be known. Even if they are time-varying, the variation is assumed to be known. The parameters are then used in the Kalman filter design as the true model of the system. For many physical problems, the parameters may not be known exactly or may change at unknown times. In such cases, it is highly desirable to design the filter to be self-learning, so that it can adapt itself to the particular situation at hand.

One solution to this problem was formulated first by Magill [16]. He considered the problem of estimation of a Gaussian random process when some parameters of the process are initially unknown and remain constant with time. The parameters are assumed to come from a finite set of known values. The optimal adaptive estimate is a weighted sum of conditional estimates, which are formed by a bank of Kalman filters. Each Kalman filter is based on a particular parameter set. The weighting coefficients are determined by a nonlinear function of the measurement residuals of the filters. The measurement residual of the filter possessing the parameter set that matches the actual set will have the smallest expected value (typically, zero mean). The residuals of all the mismatched filters will be biased. Under the Gaussian assumption, the probability of the matched filter will be the largest among all the filters.

Along similar lines, Ackerson and Fu [17] developed an adaptive state estimator for a linear system operating in a switching environment. The noise affecting the system comes from a group of Gaussian densities acting one at a time. The transitions from one noise source to the next is determined by a Markov transition probability matrix. This effort differed from that of Magill in that the parameter set describing the noise source does not remain constant with time, but is allowed to switch from one set to another at random intervals. The adaptive estimator has the same growing memory problem as that of Alspach, and consequently a suboptimal finite memory estimator was proposed.

Following Ackerson and Fu, Jaffer and Gupta [18] developed an adaptive estimator with a fixed number of filters for the problem of signal estimation under conditions of intermittent failure in the observations. Fujita and Fukao [19] established the validity of the separation theorem for the problem of determining an overall optimal control policy for a linear system with interrupted observations. Bruckner, Scott, and Rea [20] extended the adaptive algorithm of [17] to the switching of plant parameters in an air traffic control system.

A slightly different version of the adaptive filter was introduced by Moose [23]. He modeled the variations in the switching plant or switching environment by a semi-Markov process. Briefly stated, a semi-Markov process is a probabilistic system that makes its state transitions according to the transition probability matrix of a conventional Markov process. However, the amount of time spent in state i before the next transition to state j is a random variable [21], [22]. By incorporating the semi-Markov statistics into the learning portion of the adaptive filter, the problem of growing memory is completely eliminated.

This technique was successfully used by Moose and Wang [24] in the switching plant problem. Other applications include: modeling large scale depth variations of submarines, Moose [25]; modeling accelerations of a maneuvering target in three-dimensional space using spherical observations of noisy radar data, Gholson and Moose [26]; incorporating a correlated acceleration model in an airborne target tracking system, Moose, VanLandingham, and McCabe [27]; passive underwater tracking using polar coordinates, McCabe and Moose [28]; using the correlated acceleration model in the passive underwater range tracking problem, Moose and Dailey [29]; extending the algorithm to passive underwater depth tracking, Moose and Godiwala [30]; modeling unknown biases in measurement devices, Moose, Sistanizadeh, and Skagfjord [31]; and incorporating a nonlinear system block in an underwater tracking system which decouples the bearing and range estimators, Moose [32].

Many others have contributed to the field of adaptive state estimation and multiple model filtering. Tugnait and Haddad [33], [34] studied state estimation in linear systems with random Markovian noise statistics. The noise comes from a group of Gaussian distributions with different means and covariances. The transitions are determined by an unknown Markov transition probability matrix. Akashi and Kumamoto [35] devised an adaptive filter in which the overall estimate is calculated using a relatively small number of individual conditional estimates sampled at random from a larger set of individual conditional estimates.

Maybeck and Suizu [46] and Tobin and Maybeck [47] used multiple model filtering in the problem of accurately tracking the azimuth and elevation of a highly maneuverable airborne target, using outputs from a forward-looking infrared sensor as measurements. Blom and Bar-Shalom [48] described a method of timing the merging of various individual filters in an adaptive estimator to overcome the growing memory problem. Bar-Shalom, Chang, and Blom [49] used the multiple model filter in the problem of input estimation, providing estimates of the magnitude and onset time of the input. Emre and Seo [50] presented a global modeling approach to data association of multiple targets and maneuver detection/estimation of single targets. These two problems are solved simultaneously using system identification techniques, which leads to a multiple model estimator with a finite number of terms.

2.3 Alternate NonGaussian Filtering

Several alternatives to the Gaussian sum and adaptive Kalman filter algorithms have been proposed for systems containing nonGaussian inputs. Masreliez [43] introduced an approximate nonGaussian filtering method for linear systems. Instead of approximating the densities using Gaussian sums, he used a nonlinear score function of the measurement residuals to produce the Kalman filter equations. The technique works best however when only some of the random processes are nonGaussian, with the rest being Gaussian. Masreliez and Martin [44] extend this approximate nonGaussian filtering method by applying an influence function of min-max robustness theory to replace the previously used score function. Tsai and Kurz [45] devised a robust Kalman filter based on an m-interval polynomial approximation method for unknown nonGaussian noise. As the score function of [43] is partitioned into m segments, each is better approximated with a low-order polynomial function which is easier to implement and adapt to the density function of the residual process.

Other robust estimation techniques are used by Kirlin and Moghaddamjoo for systems with unknown inputs and nonGaussian measurement errors. In [51] the input forcing function and measurement bias are estimated using a moving data window median calculation. The plant and measurement noise covariances are found using biweights and rank correlation. In [52] they discussed adaptive estimation of unknown inputs and measurement noise covariance using a running window curve-fitting algorithm. Estimation of the plant noise covariance is accomplished using an independent technique, based on the residuals and a stochastic approximation method.

3.0 NONGAUSSIAN ESTIMATION

This chapter discusses the mathematical details of the Gaussian sum approach to nonGaussian signal estimation. It begins with an outline of Alspach's development of the Gaussian sum estimation technique [6-8]. A major problem with this technique is indicated which necessitates a modified approach. Next, a detailed description of the nonGaussian signal model is given which leads to an alternate formulation of the Gaussian sum density approximation. The initial parameter vector describing the density approximation is found using a curve fitting procedure. A modified Gaussian sum adaptive filter is developed for a general discrete-time system and a simulation example is given. Finally, key differences between the Gaussian sum and modified Gaussian sum algorithms are examined.

3.1 Gaussian Sum Estimation Technique

As previously stated in chapter 2, an estimation technique, applicable to both linear systems with nonGaussian inputs and nonlinear systems with Gaussian inputs, has been developed by Alspach [6-8]. The basic problem is to estimate the state x_k of a discrete-time system from the current measurement sequence $Z_k = \{z_1, z_2, \ldots, z_{k-1}, z_k\}$. The "best" or optimal estimate $\hat{x}_k$ of the system state x_k is chosen as that estimate which minimizes the mean-square error between x_k and $\hat{x}_k$. This results in the conditional mean estimate

$$\hat{x}_k = E[x_k \mid Z_k] = \int_{-\infty}^{\infty} x_k \, p(x_k \mid Z_k) \, dx_k \tag{3.1.1}$$

where $p(x_k \mid Z_k)$ is the conditional density function of the state given the current measurement sequence.

The Gaussian sum estimator is developed by Alspach using the following system model

$$x_k = \Phi_k x_{k-1} + w_{k-1} \tag{3.1.2}$$

$$z_k = H_k x_k + v_k \tag{3.1.3}$$

where the initial state has a Gaussian sum density of the form

$$p(x_0) = \sum_{i=1}^{\xi_0'} \alpha_{0_i}' \mathrm{N}[\mu_{0_i}', P_{0_i}'] \tag{3.1.4}$$

Assume that the plant and measurement noise processes, w_k and v_k, are statistically independent, nonGaussian, white noise sequences with Gaussian sum densities

$$p(w_k) = \sum_{n=1}^{q_k} \beta_{k_n} \mathrm{N}[\omega_{k_n}, Q_{k_n}] \tag{3.1.5}$$

and

$$p(v_k) = \sum_{m=1}^{r_k} \gamma_{k_m} \mathrm{N}[\nu_{k_m}, R_{k_m}] \tag{3.1.6}$$

The key idea of Gaussian sum estimation is to use (3.1.4) - (3.1.6) in the approximation of $p(x_k \mid Z_k)$ as a sum of Gaussian densities. Using Bayes' rule, this conditional density function can be determined recursively [7] from

$$p(x_k \mid Z_k) = \frac{p(x_k \mid Z_{k-1}) p(z_k \mid x_k)}{p(z_k \mid Z_{k-1})} \tag{3.1.7}$$

$$p(x_k \mid Z_{k-1}) = \int_{-\infty}^{\infty} p(x_{k-1} \mid Z_{k-1}) p(x_k \mid x_{k-1}) dx_{k-1} \tag{3.1.8}$$

where the normalizing constant $p(z_k \mid Z_{k-1})$ in (3.1.7) is given by

$$p(z_k \mid Z_{k-1}) = \int_{-\infty}^{\infty} p(x_k \mid Z_{k-1}) p(z_k \mid x_k) dx_k \quad . \tag{3.1.9}$$

The density $p(z_k \mid x_k)$ in (3.1.7) is determined by the measurement noise density $p(v_k)$ of (3.1.6) and the measurement equation (3.1.3). Similarly, the density $p(x_k \mid x_{k-1})$ in (3.1.8) is determined by the plant noise density $p(w_k)$ of (3.1.5) and the plant equation (3.1.2). Knowledge of $p(z_k \mid x_k)$, $p(x_k \mid x_{k-1})$, and the initial state density $p(x_0)$ of (3.1.4) determines $p(x_k \mid Z_k)$ for all k. Alspach [6], [7] derives two additional theorems which are briefly stated here. The measurement update theorem provides passage of $p(x_k \mid Z_{k-1})$ to $p(x_k \mid Z_k)$, while the time update theorem provides passage of $p(x_k \mid Z_k)$ to $p(x_{k+1} \mid Z_k)$.

Measurement Update Theorem

Suppose that the k^{th} stage *prediction* density function is

$$p(x_k \mid Z_{k-1}) = \sum_{i=1}^{\xi'_k} \alpha'_{k_i} \, \mathrm{N}[\mu'_{k_i}, P'_{k_i}] \tag{3.1.10}$$

Then the k^{th} stage *filtering* density is given by

$$p(x_k \mid Z_k) = \sum_{j=1}^{\xi_k} \alpha_{k_j} \mathrm{N}[\mu_{k_j}, P_{k_j}] \tag{3.1.11}$$

where the mean and covariance terms are found from the Kalman filter measurement update equations

$$\mu_{k_j} = \mu'_{k_i} + K_{k_j}(z_k - H_k \mu'_{k_i} - \nu_{k_m}) \tag{3.1.12}$$

$$P_{k_j} = [I - K_{k_j} H_k] P'_{k_i} \tag{3.1.13}$$

$$K_{k_j} = P'_{k_i} H_k^{\mathrm{T}} (H_k P'_{k_i} H_k^{\mathrm{T}} + R_{k_m})^{-1} \tag{3.1.14}$$

The filtering density (3.1.11) is the result of a double sum formed by the product of the prediction density (3.1.10) and the measurement noise density (3.1.6). The double sum is rewritten as a single sum with upper summation limit

$$\xi_k = \xi'_k r_k \tag{3.1.15}$$

produced by the product of upper summation limits ξ'_k of (3.1.10) and r_k of (3.1.6). As a result, the number of terms of the filtering Gaussian sum density grows geometrically at each measurement update with a ratio equal to the number of Gaussian sum terms of the measurement noise.

The α_{k_j} term of (3.1.11) is found from

$$\alpha_{k_j} = \frac{\alpha'_{k_i} \gamma_{k_m} W_{k_j}}{\sum_{i=1}^{\xi'_k} \alpha'_{k_i} \gamma_{k_m} W_{k_j}} \tag{3.1.16}$$

where the index of summation j is defined as

$$j = i + (m-1)\xi'_k \qquad \begin{matrix} i = 1, \ldots, \xi'_k \\ m = 1, \ldots, r_k \end{matrix}$$

The W_{k_j} term of (3.1.16) is the Gaussian probability density function

$$W_{k_j} = \mathrm{N}[H_k \mu'_{k_i} + \nu_{k_m}, H_k P'_{k_i} H_k^{\mathrm{T}} + R_{k_m}] \tag{3.1.17}$$

with terms that are readily found from measurement update equations (3.1.12) and (3.1.14). The α_{k_j} term (3.1.16) is a nonlinear function of the current measurement data, with its denominator acting as a scale factor so that the α_{k_j} terms remain bounded between 0 and 1 at each iteration.

Time Update Theorem

Given that the k^{th} stage *filtering* density function is

$$p(x_k \mid Z_k) = \sum_{j=1}^{\xi_k} \alpha_{k_j} \mathrm{N}[\mu_{k_j}, P_{k_j}] \tag{3.1.18}$$

Then the k^{th} stage *prediction* density is given by

$$p(x_{k+1} \mid Z_k) = \sum_{i=1}^{\xi'_{k+1}} \alpha'_{k+1_i} \mathrm{N}[\mu'_{k+1_i}, P'_{k+1_i}] \tag{3.1.19}$$

where the mean and covariance terms are found from the Kalman filter time update equations

$$\mu'_{k+1_i} = \Phi_{k+1} \mu_{k_j} + \omega_{k_n} \tag{3.1.20}$$

$$P'_{k+1_i} = \Phi_{k+1} P_{k_j} \Phi^{\mathrm{T}}_{k+1} + Q_{k_n} \tag{3.1.21}$$

The prediction density (3.1.19) is the result of a double sum formed by the product of the filtering density (3.1.18) and the plant noise density (3.1.5). The double sum is rewritten as a single sum with upper summation limit

$$\xi'_{k+1} = \xi_k q_k \tag{3.1.22}$$

produced by the product of upper summation limits ξ_k of (3.1.18) and q_k of (3.1.5). As a result, the number of terms of the prediction Gaussian sum density grows geometrically at each time update with a ratio equal to the number of Gaussian sum terms of the plant noise.

The α'_{k+1_i} term of (3.1.19) is found from

$$\alpha'_{k+1_i} = \alpha_{k_j} \beta_{k_n} \tag{3.1.23}$$

where the index of summation i is defined as

$$i = j + (n-1)\xi_k \qquad \begin{array}{l} j = 1, \ldots, \xi_k \\ n = 1, \ldots, q_k \end{array}$$

Alspach's two theorems show that for the system defined by (3.1.2) - (3.1.6) the Gaussian sum form repeats itself from one stage to the next so that (3.1.10) and (3.1.11) are the general forms for an arbitrary stage. The measurement update equations (3.1.12) - (3.1.14) and time update equations (3.1.20) and (3.1.21) comprise a bank of Kalman filters operating in parallel. Using the Gaussian sum approximation of the conditional density function (3.1.11), the overall state estimate $\hat{x}_k$ is found as a weighted sum of the individual conditional mean values from each filter

$$\hat{x}_k = E[x_k \mid Z_k] = \sum_{j=1}^{\xi_k} \alpha_{k_j} \mu_{k_j} \tag{3.1.24}$$

The corresponding conditional covariance is

$$P_k = E[(x_k - \hat{x}_k)(x_k - \hat{x}_k)^{\mathrm{T}} \mid Z_k] = \sum_{j=1}^{\xi_k} \alpha_{k_j} [P_{k_j} + (\hat{x}_k - \mu_{k_j})(\hat{x}_k - \mu_{k_j})^{\mathrm{T}}] \tag{3.1.25}$$

It is evident from (3.1.15) and (3.1.22) that the number of Gaussian terms increases at each stage of the estimator. This *growing memory* problem is a serious limitation in Alspach's development. A modified approach is required to alleviate this limitation.

3.2 NonGaussian Signal Model Development

Let u be a random noise process or random input signal with a nonGaussian density function. It can be modeled as the sum of two statistically independent random processes

$$u = b + n \tag{3.2.1}$$

The first term, b, is a semi-Markov process with state transitions governed by the transition probability matrix of a conventional Markov process. Markov processes have the property that a transition is made at every time instant. The transition may return the process to the state it previously occupied, but a transition occurs nevertheless. However, in the semi-Markov case, the amount of time between transitions is a random variable [22]. The value of b is randomly selected from a fixed set of discrete values, characterized by a delta probability density function

$$p(b) = \sum_{i=1}^{M} P_i \delta(b - b_i) \tag{3.2.2}$$

with

$$\sum_{i=1}^{M} P_i = 1; \quad P_i \geq 0 \;\text{ for } i = 1, 2, \ldots, M \tag{3.2.3}$$

This process can be thought of as a randomly-switching bias, each bias value b_i having probability P_i.

The second term, n, is a zero mean white Gaussian process with variance σ_n^2. With both densities known, the density function of u can be found using the convolution relationship between u, n, and b [53]

$$p(u) = \int_{-\infty}^{\infty} p_n(u - b) p_b(b) db \tag{3.2.4}$$

where $p_n(u-b)$ is the Gaussian density with $n = u - b$

$$p_n(u - b) = \frac{1}{\sqrt{2\pi}\sigma_n} e^{-\frac{1}{2}\left(\frac{u-b}{\sigma_n}\right)^2} \tag{3.2.5}$$

Substituting (3.2.2) and (3.2.5) into the convolution integral (3.2.4) gives

$$p(u) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma_n} e^{-\frac{1}{2}\left(\frac{u-b}{\sigma_n}\right)^2} \sum_{i=1}^{M} P_i \delta(b - b_i) db \tag{3.2.6}$$

Interchanging integration and summation

$$p(u) = \sum_{i=1}^{M} P_i \left[\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma_n} \, e^{-\frac{1}{2}\left(\frac{u-b}{\sigma_n}\right)^2} \delta(b - b_i) db \right] \tag{3.2.7}$$

Using the sifting property of the delta function [54]

$$f(a) = \int_{-\infty}^{\infty} f(x)\delta(x-a)dx \tag{3.2.8}$$

the integral of (3.2.7) is evaluated as

$$p(u) = \sum_{i=1}^{M} P_i \left[\frac{1}{\sqrt{2\pi}\sigma_n} \, e^{-\frac{1}{2}\left(\frac{u-b_i}{\sigma_n}\right)^2} \right] \tag{3.2.9}$$

or

$$p(u) = \sum_{i=1}^{M} P_i \mathrm{N}[b_i, \sigma_n^2] \tag{3.2.10}$$

Thus, the nonGaussian density function of u can be modeled as a Gaussian sum. The weight P_i of each Gaussian term is the probability of the i^{th} bias term. The bias term b_i is restricted to be randomly selected from the same fixed set of bias values at each iteration. Using this model in conjunction with the modified Gaussian sum adaptive filter developed in the next section avoids the growing memory problem of Alspach's development.

Selecting the parameters P_i, b_i, and σ_n in (3.2.10) to obtain the "best" approximation p_{GS} to some actual nonGaussian density function p_A is accomplished by means of minimizing the L^k norm

$$\| p_A - p_{GS} \|^k = \int_{-\infty}^{\infty} \left| p_A(u) - \sum_{i=1}^{M} P_i \mathrm{N}[b_i, \sigma_n^2] \right|^k du \tag{3.2.11}$$

This curve fitting exercise can be done off-line using several values of M until a suitable trade-off between minimum norm and minimum M is obtained. Alspach [6], [7] performed this curve fitting procedure using L^1 and L^2 norms for a uniform density and a Gamma density. It was found that minimizing the L^2 norm resulted in many fewer terms in the Gaussian sum and a considerably better looking approximation for both densities compared to minimizing the L^1 norm.

Figure 1 compares a Gamma density with a four-term Gaussian sum density approximation minimizing the L^2 norm. The Gaussian sum curve is shown to fit the Gamma curve reasonably well. The Gamma density used is

$$p(u) = \begin{cases} \dfrac{u^3 e^{-u}}{6} & u \geq 0 \\ 0 & u < 0 \end{cases} \tag{3.2.12}$$

Each term of the Gaussian sum has a fixed value of $\sigma_n = 1$. Table 1 lists the values of P_i and b_i used in the Gaussian sum.

Table 1.

Gaussian sum P_i, b_i values, $\sigma_n = 1$.

i	P_i	b_i
1	0.081	2.537
2	0.432	2.553
3	0.356	4.555
4	0.131	6.933

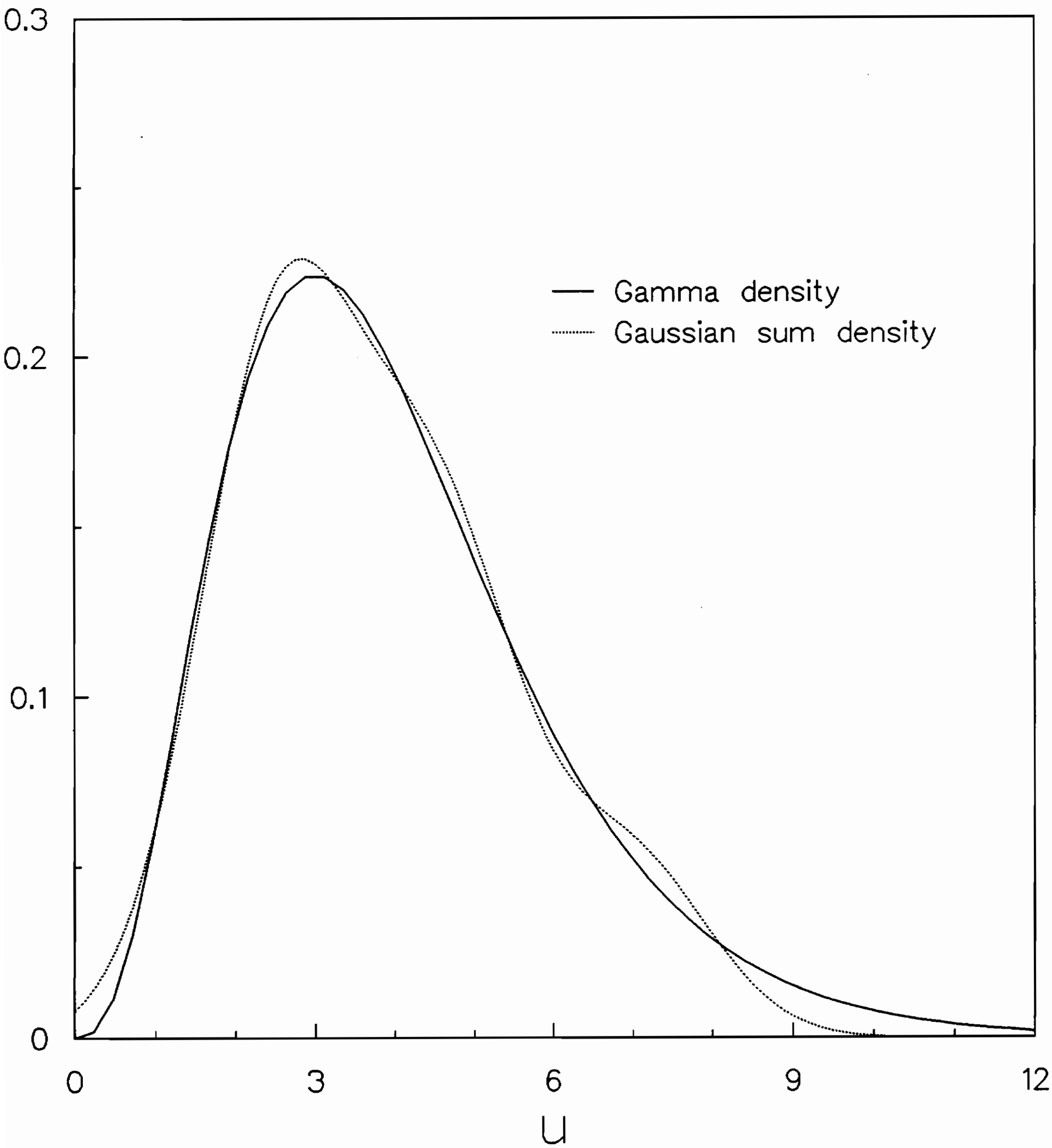


Figure 1. Gamma and Gaussian sum probability density functions.

3.3 Modified Gaussian Sum Adaptive Filtering

A modified Gaussian sum adaptive filter is now developed for a linear system with a deterministic input signal and nonGaussian plant and measurement noise. The system is modeled in standard discrete-time state-space form as

$$x_{k+1} = \Phi x_k + \Gamma u_k + \Psi w_k \tag{3.3.1}$$

$$w_k = a_k + m_k \tag{3.3.2}$$

$$z_k = H x_k + v_k \tag{3.3.3}$$

$$v_k = b_k + n_k \tag{3.3.4}$$

where

x_{k+1}	is the state vector
u_k	is a known deterministic input
w_k	is the vector Gaussian sum signal model of the actual nonGaussian plant noise process, comprised of semi-Markov bias vector a_k, and zero mean white Gaussian noise m_k with covariance Q
z_k	is the measurement vector
v_k	is the vector Gaussian sum signal model of the actual nonGaussian measurement noise process, comprised of semi-Markov bias vector b_k, and zero mean white Gaussian noise n_k with covariance R
Φ, Γ, Ψ, H	are the respective constant transition matrices

and the random quantities x_k, a_k, m_k, b_k, n_k are assumed to be mutually statistically independent.

The optimal estimate $\hat{x}_{k+1}$ of the state vector is found by minimizing the mean-square error between x_{k+1} and $\hat{x}_{k+1}$. This results in the conditional mean estimate

$$\hat{x}_{k+1} = E[x_{k+1} \mid Z_{k+1}] = \int_{-\infty}^{\infty} x_{k+1} p(x_{k+1} \mid Z_{k+1})\, dx_{k+1} \tag{3.3.5}$$

where Z_{k+1} is the current measurement sequence $\{z_1, z_2, \ldots, z_{k+1}\}$. The conditional density function of (3.3.5) can be written as the ratio of the corresponding joint and marginal densities

$$p(x_{k+1} \mid Z_{k+1}) = \frac{p(x_{k+1}, Z_{k+1})}{p(Z_{k+1})} \tag{3.3.6}$$

The two bias vectors a and b are explicitly brought into (3.3.6) by considering the joint density $p(x_{k+1}, Z_{k+1})$ to be a marginal density found from $p(x_{k+1}, Z_{k+1}, a = a_i, b = b_j)$ by summing over the a and b terms

$$p(x_{k+1} \mid Z_{k+1}) = \frac{\sum_{i=1}^{N} \sum_{j=1}^{M} p(x_{k+1}, Z_{k+1}, a = a_i, b = b_j)}{p(Z_{k+1})} \tag{3.3.7}$$

Expanding (3.3.7) and using $p(x_{k+1}, Z_{k+1}, a_i, b_j)$ as shorthand for $p(x_{k+1}, Z_{k+1}, a = a_i, b = b_j)$ gives

$$p(x_{k+1} \mid Z_{k+1}) = \frac{\sum_{i=1}^{N} \sum_{j=1}^{M} p(x_{k+1} \mid Z_{k+1}, a_i, b_j) p(Z_{k+1}, a_i, b_j)}{p(Z_{k+1})}$$

$$p(x_{k+1} \mid Z_{k+1}) = \frac{\sum_{i=1}^{N} \sum_{j=1}^{M} p(x_{k+1} \mid Z_{k+1}, a_i, b_j) p(a_i, b_j \mid Z_{k+1}) p(Z_{k+1})}{p(Z_{k+1})}$$

and finally

$$p(x_{k+1} \mid Z_{k+1}) = \sum_{i=1}^{N} \sum_{j=1}^{M} p(x_{k+1} \mid Z_{k+1}, a_i, b_j) p(a_i, b_j \mid Z_{k+1}) \tag{3.3.8}$$

Substituting (3.3.8) into the conditional mean (3.3.5) results in

$$\hat{x}_{k+1} = \int_{-\infty}^{\infty} x_{k+1} \left[\sum_{i=1}^{N} \sum_{j=1}^{M} p(x_{k+1} \mid Z_{k+1}, a_i, b_j) p(a_i, b_j \mid Z_{k+1}) \right] dx_{k+1} \tag{3.3.9}$$

Interchanging integration and summation gives

$$\hat{x}_{k+1} = \sum_{i=1}^{N} \sum_{j=1}^{M} \left[\int_{-\infty}^{\infty} x_{k+1} p(x_{k+1} \mid Z_{k+1}, a_i, b_j)\, dx_{k+1} \right] p(a_i, b_j \mid Z_{k+1}) \tag{3.3.10}$$

The bracketed integral in (3.3.10) is the conditional mean estimate of x_{k+1} given that $a = a_i$ and $b = b_j$, denoted by

$$\hat{x}_{k+1}^{ij} = \int_{-\infty}^{\infty} x_{k+1} p(x_{k+1} \mid Z_{k+1}, a_i, b_j)\, dx_{k+1} \tag{3.3.11}$$

In effect, $\hat{x}_{k+1}^{ij}$ represents the estimate for the ij^{th} density combination from the two Gaussian sums. A set of $N \times M$ (NM) estimators is needed to provide all the individual $\hat{x}_{k+1}^{ij}$ estimates. The overall estimate from (3.3.10) and (3.3.11),

$$\hat{x}_{k+1} = \sum_{i=1}^{N} \sum_{j=1}^{M} \hat{x}_{k+1}^{ij} p(a_i, b_j \mid Z_{k+1}) \tag{3.3.12}$$

is a weighted sum of the NM individual estimates. The weighting factor $p(a_i, b_j \mid Z_{k+1})$ is the probability that $a = a_i$ and $b = b_j$ given the current measurement sequence.

Each estimate (3.3.11) is found from a modified linear state estimator. Rewriting the measurement sequence as $Z_{k+1} = \{Z_k, z_{k+1}\}$, the conditional density function of (3.3.11) is

$$p(x_{k+1} \mid Z_k, z_{k+1}, a_i, b_j) = \frac{p(x_{k+1}, Z_k, z_{k+1}, a_i, b_j)}{p(Z_k, z_{k+1}, a_i, b_j)} \tag{3.3.13}$$

Expanding as before gives

$$p(x_{k+1} \mid Z_k, z_{k+1}, a_i, b_j) = \frac{p(z_{k+1} \mid x_{k+1}, a_i, b_j, Z_k) p(x_{k+1} \mid a_i, b_j, Z_k)}{p(z_{k+1} \mid a_i, b_j, Z_k)} \tag{3.3.14}$$

The first term of the numerator of (3.3.14) is Gaussian since x_{k+1}, $a = a_i$, and $b = b_j$ are given. The conditional mean value is found by first combining (3.3.3) and (3.3.4) at time iteration $k+1$

$$z_{k+1} = Hx_{k+1} + b_{k+1} + n_{k+1} \tag{3.3.15}$$

and then taking the expected value

$$m_{k+1}^{ij} = E[z_{k+1} \mid x_{k+1}, a_i, b_j, Z_k] = Hx_{k+1} + b_j + E[n_{k+1} \mid x_{k+1}, a_i, b_j, Z_k]$$

Since n_{k+1} has a mean value of zero, the expected value is

$$m_{k+1}^{ij} = Hx_{k+1} + b_j \tag{3.3.16}$$

The conditional covariance is found next as

$$\begin{aligned}C^{ij}_{k+1} &= E[(z_{k+1} - m^{ij}_{k+1})(z_{k+1} - m^{ij}_{k+1})^{\mathrm{T}} \mid x_{k+1}, a_i, b_j, Z_k] \\ &= E[(Hx_{k+1} + b_{k+1} + n_{k+1} - Hx_{k+1} - b_j)(z_{k+1} - m^{ij}_{k+1})^{\mathrm{T}} \mid x_{k+1}, a_i, b_j, Z_k]\end{aligned}$$

Knowing that $b = b_j$

$$C^{ij}_{k+1} = E[n_{k+1} n^{\mathrm{T}}_{k+1} \mid x_{k+1}, a_i, b_j, Z_k] = R \tag{3.3.17}$$

The conditional density is therefore

$$p(z_{k+1} \mid x_{k+1}, a_i, b_j, Z_k) = \mathrm{N}[m^{ij}_{k+1}, R] = \mathrm{N}[Hx_{k+1} + b_j, R] \tag{3.3.18}$$

The second term of the numerator of (3.3.14) is also Gaussian since $a = a_i$ and $b = b_j$ are given. The conditional mean value is found by first combining (3.3.1) and (3.3.2)

$$x_{k+1} = \Phi x_k + \Gamma u_k + \Psi a_k + \Psi m_k \tag{3.3.19}$$

and then taking the expected value

$$\tilde{x}^{ij}_{k+1} = E[x_{k+1} \mid a_i, b_j, Z_k] = \Phi E[x_k \mid a_i, b_j, Z_k] + \Gamma u_k + \Psi a_i + \Psi E[m_k \mid a_i, b_j, Z_k]$$

Since m_k has a mean value of zero, the expected value is

$$\tilde{x}^{ij}_{k+1} = \Phi \hat{x}^{ij}_k + \Gamma u_k + \Psi a_i \tag{3.3.20}$$

where $\tilde{x}_{k+1}^{ij}$ is the *prediction* at time iteration $k+1$ given prior measurements up to only iteration k, while $\hat{x}_k^{ij}$ is the previous *estimate* at iteration k given measurements up to iteration k.

The conditional covariance is found next as

$$\begin{aligned}M_{k+1}^{ij} &= E[(x_{k+1} - \tilde{x}_{k+1}^{ij})(x_{k+1} - \tilde{x}_{k+1}^{ij})^{\mathrm{T}} \mid a_i, b_j, Z_k] \\ &= E[(\Phi x_k + \Gamma u_k + \Psi a_k + \Psi m_k - \Phi \hat{x}_k^{ij} - \Gamma u_k - \Psi a_i)(x_{k+1} - \tilde{x}_{k+1}^{ij})^{\mathrm{T}} \mid a_i, b_j, Z_k]\end{aligned}$$

Since $a = a_i$ and recalling that x_k and m_k are statistically independent

$$M_{k+1}^{ij} = \Phi E[(x_k - \hat{x}_k^{ij})(x_k - \hat{x}_k^{ij})^{\mathrm{T}} \mid a_i, b_j, Z_k]\Phi^{\mathrm{T}} + \Psi E[m_k m_k^{\mathrm{T}} \mid a_i, b_j, Z_k]\Psi^{\mathrm{T}}$$

Letting $P_k^{ij} = E[(x_k - \hat{x}_k^{ij})(x_k - \hat{x}_k^{ij})^{\mathrm{T}} \mid a_i, b_j, Z_k]$ be the conditional covariance of the error term $x_k - \hat{x}_k^{ij}$,

$$M_{k+1}^{ij} = \Phi P_k^{ij} \Phi^{\mathrm{T}} + \Psi Q \Psi^{\mathrm{T}} \tag{3.3.21}$$

Since the covariances Q and R are the same for each ij^{th} estimator, the respective conditional covariances simplify to $P_k^{ij} = P_k$ and $M_{k+1}^{ij} = M_{k+1}$, and the conditional density is therefore

$$p(x_{k+1} \mid a_i, b_j, Z_k) = \mathrm{N}[\tilde{x}_{k+1}^{ij}, M_{k+1}] = \mathrm{N}[\Phi \hat{x}_k^{ij} + \Gamma u_k + \Psi a_i, \Phi P_k \Phi^{\mathrm{T}} + \Psi Q \Psi^{\mathrm{T}}] \tag{3.3.22}$$

The least-mean-square (LMS) estimator is found by taking the derivative of the natural logarithm of (3.3.14) with respect to x_{k+1}, setting the result equal to zero, and determining the value of $x_{k+1} = \hat{x}_{k+1}^{ij}$ that produces this result. Since the denominator term of (3.3.14) does not depend on x_{k+1}, it will not play a role in determining $\hat{x}^{ij}_{k+1}$. Now, set

$$\frac{\partial}{\partial x_{k+1}}[\ln\ p(x_{k+1} \mid Z_k, z_{k+1}, a_i, b_j)] = 0 \tag{3.3.23}$$

Using (3.3.18) and (3.3.22) in (3.3.14), (3.3.23) becomes

$$\frac{\partial}{\partial x_{k+1}}[\ln\ \mathrm{N}[m^{ij}_{k+1}, R] + \ln\ \mathrm{N}[\tilde{x}^{ij}_{k+1}, M_{k+1}] - 0] = 0$$

$$\frac{\partial}{\partial x_{k+1}}[-\tfrac{1}{2}(z_{k+1} - m^{ij}_{k+1})^{\mathrm{T}}R^{-1}(z_{k+1} - m^{ij}_{k+1}) - \tfrac{1}{2}(x_{k+1} - \tilde{x}^{ij}_{k+1})^{\mathrm{T}}M^{-1}_{k+1}(x_{k+1} - \tilde{x}^{ij}_{k+1})] = 0 \tag{3.3.24}$$

Combining (3.3.16) and using a rule of matrix differentiation (if $C = C^{\mathrm{T}}$)

$$\frac{\partial}{\partial x}(Ax + b)^{\mathrm{T}}C(Ax + b) = 2A^{\mathrm{T}}C(Ax + b) \tag{3.3.25}$$

the first term of (3.3.24) is arranged in the form of (3.3.25)

$$\frac{\partial}{\partial x_{k+1}}[-\tfrac{1}{2}(-Hx_{k+1} + z_{k+1} - b_j)^{\mathrm{T}}R^{-1}(-Hx_{k+1} + z_{k+1} - b_j)]$$

with $A = -H$, $b = z_{k+1} - b_j$, and $C = R^{-1}$. Applying (3.3.25) results in

$$-H^{\mathrm{T}}R^{-1}(Hx_{k+1} - z_{k+1} + b_j) \tag{3.3.26}$$

Similarly, the second term of the derivative (3.2.24) is

$$-M^{-1}_{k+1}(x_{k+1} - \tilde{x}^{ij}_{k+1}) \tag{3.3.27}$$

Combining (3.3.20), (3.3.26), and (3.3.27) and setting $x_{k+1} = \hat{x}^{ij}_{k+1}$ gives

$$-H^{\mathrm{T}}R^{-1}(H\hat{x}^{ij}_{k+1} - z_{k+1} + b_j) - M^{-1}_{k+1}(\hat{x}^{ij}_{k+1} - \Phi\hat{x}^{ij}_{k} - \Gamma u_k - \Psi a_i) = 0$$

or

$$(M^{-1}_{k+1} + H^{\mathrm{T}}R^{-1}H)\hat{x}^{ij}_{k+1} = H^{\mathrm{T}}R^{-1}(z_{k+1} - b_j) + M^{-1}_{k+1}(\Phi\hat{x}^{ij}_{k} + \Gamma u_k + \Psi a_i) \tag{3.3.28}$$

Now, define

$$P^{-1}_{k+1} = (M^{-1}_{k+1} + H^{\mathrm{T}}R^{-1}H) \tag{3.3.29}$$

Using (3.3.29) and solving for $\hat{x}^{ij}_{k+1}$ in (3.3.28) gives

$$\hat{x}^{ij}_{k+1} = P_{k+1}H^{\mathrm{T}}R^{-1}(z_{k+1} - b_j) + P_{k+1}M^{-1}_{k+1}(\Phi\hat{x}^{ij}_{k} + \Gamma u_k + \Psi a_i) \tag{3.3.30}$$

To rewrite $P_{k+1}M^{-1}_{k+1}$, first premultiply (3.3.29) by P_{k+1}, giving the identity matrix

$$I = P_{k+1}(M^{-1}_{k+1} + (H^{\mathrm{T}}R^{-1}H) \tag{3.3.31}$$

then solve (3.3.31) for $P_{k+1}M^{-1}_{k+1}$ and substitute into (3.3.30)

$$\hat{x}^{ij}_{k+1} = P_{k+1}H^{\mathrm{T}}R^{-1}(z_{k+1} - b_j) + (I - P_{k+1}H^{\mathrm{T}}R^{-1}H)(\Phi\hat{x}^{ij}_{k} + \Gamma u_k + \Psi a_i) \tag{3.3.32}$$

Now, let

$$K_{k+1} = P_{k+1}H^{\mathrm{T}}R^{-1} \tag{3.3.33}$$

be the Kalman gain, and rearrange using (3.3.29) to produce

$$\begin{aligned} K_{k+1} &= (M_{k+1} + H^{-1}RH^{\mathrm{T}^{-1}})H^{\mathrm{T}}R^{-1} \\ &= M_{k+1}H^{\mathrm{T}}R^{-1} + H^{-1} \\ &= (M_{k+1}H^{\mathrm{T}}(R^{-1} + H^{\mathrm{T}^{-1}}M_{k+1}^{-1}H^{-1}) \end{aligned}$$

or

$$K_{k+1} = M_{k+1}H^{\mathrm{T}}(HM_{k+1}H^{\mathrm{T}} + R)^{-1} \tag{3.3.34}$$

Substituting (3.3.33) into (3.3.32) and rearranging produces the familiar Kalman filter equation below, with modifications to allow for the two bias terms of the Gaussian sum densities

$$\hat{x}_{k+1}^{ij} = \Phi\hat{x}_k^{ij} + \Gamma u_k + \Psi a_i + K_{k+1}[z_{k+1} - b_j - H(\Phi\hat{x}_k^{ij} + \Gamma u_k + \Psi a_i)] \tag{3.3.35}$$

As a final step, substitute (3.3.33) into (3.3.31) to produce

$$P_{k+1}M_{k+1}^{-1} = (I - K_{k+1}H)$$

or

$$P_{k+1} = (I - K_{k+1}H)M_{k+1} \tag{3.3.36}$$

With the individual ij^{th} estimate (3.3.35) determined, the next step in producing the overall estimate (3.3.12) is to find the weighting term $p(a_i, b_j \mid Z_{k+1})$. Using $Z_{k+1} = \{Z_k, z_{k+1}\}$ again, the weighting term becomes

$$p(a_i, b_j \mid Z_{k+1}) = \frac{p(z_{k+1}, a_i, b_j, Z_k)}{p(z_{k+1}, Z_k)}$$

Using Bayes' theorem produces

$$p(a_i, b_j \mid Z_{k+1}) = \frac{p(z_{k+1} \mid a_i, b_j, Z_k) p(a_i, b_j \mid Z_k)}{p(z_{k+1} \mid Z_k)} \tag{3.3.37}$$

The first term of the numerator can be modeled as a Gaussian density if the bias terms switch slowly compared to the time interval k. This assumption will be made here and has been validated by extensive simulation and analysis by Moose and Wang [24]. Combining (3.3.15) and (3.3.19), z_{k+1} is written as

$$z_{k+1} = H(\Phi x_k + \Gamma u_k + \Psi a_k + \Psi m_k) + b_{k+1} + n_{k+1} \tag{3.3.38}$$

Taking the expected value, the conditional mean is

$$\begin{aligned} \tilde{z}_{k+1}^{ij} &= E[z_{k+1} \mid a_i, b_j, Z_k] \\ &= H\Phi E[x_k \mid a_i, b_j, Z_k] + H\Gamma u_k + H\Psi a_i + b_j \end{aligned}$$

or

$$\tilde{z}_{k+1}^{ij} = H(\Phi \hat{x}_k^{ij} + \Gamma u_k + \Psi a_i) + b_j \tag{3.3.39}$$

The conditional covariance is found next as

$$R_{k+1}^{ij} = E[(z_{k+1} - \tilde{z}_{k+1}^{ij})(z_{k+1} - \tilde{z}_{k+1}^{ij})^{\mathrm{T}} \mid a_i, b_j, Z_k] \tag{3.3.40}$$

Using (3.3.38) and (3.3.39), the $z_{k+1} - \tilde{z}_{k+1}^{ij}$ term is written as

$$z_{k+1} - \tilde{z}_{k+1}^{ij} = H(\Phi x_k + \Gamma u_k + \Psi a_k + \Psi m_k) + b_{k+1} + n_{k+1} - H(\Phi \hat{x}_k^{ij} + \Gamma u_k + \Psi a_i) - b_j \tag{3.3.41}$$

Inserting (3.3.41) into (3.3.40) gives

$$R_{k+1}^{ij} = E[(H\{\Phi(x_k - \hat{x}_k^{ij}) + \Psi m_k\} + n_{k+1})(z_{k+1} - \tilde{z}_{k+1}^{ij})^{\mathrm{T}} \mid a_i, b_j, Z_k]$$

and using (3.3.17) and (3.3.21) produces

$$R_{k+1}^{ij} = HM_{k+1}^{ij}H^{\mathrm{T}} + R \tag{3.3.42}$$

Since the covariances R and $M_{k+1}^{ij} = M_{k+1}$ are the same for each ij^{th} estimator, the conditional covariance simplifies to $R_{k+1}^{ij} = R_{k+1}$, and the conditional density is therefore

$$\begin{aligned} p(z_{k+1} \mid a_i, b_j, Z_k) &= \mathrm{N}[\tilde{z}_{k+1}^{ij}, R_{k+1}] \\ &= \mathrm{N}[H(\Phi\hat{x}_k^{ij} + \Gamma u_k + \Psi a_i) + b_j, HM_{k+1}H^{\mathrm{T}} + R] \end{aligned} \tag{3.3.43}$$

The second term of the numerator of (3.3.37) can be rewritten to explicitly include the time interval for the two bias terms

$$p(a_i, b_j \mid Z_k) = p(a_{k+1} = a_i, b_{k+1} = b_j \mid Z_k) \tag{3.3.44}$$

Using Bayes' theorem and the definition of marginal densities, (3.3.44) is expanded into

$$p(a_i, b_j \mid Z_k) = \frac{\sum_{\alpha=1}^{N}\sum_{\beta=1}^{M} p(a_{k+1} = a_i, a_k = a_\alpha, b_{k+1} = b_j, b_k = b_\beta, Z_k)}{p(Z_k)}$$

$$= \sum_{\alpha=1}^{N}\sum_{\beta=1}^{M} p(a_{k+1} = a_i, b_{k+1} = b_j \mid a_k = a_\alpha, b_k = b_\beta, Z_k)p(a_k = a_\alpha, b_k = b_\beta, Z_k) \tag{3.3.45}$$

Recalling that a_{k+1} and b_{k+1} are assumed statistically independent, and given that Z_k, $a_k = a_\alpha$, and $b_k = b_\beta$ are known, (3.3.45) becomes

$$= \sum_{\alpha=1}^{N} \sum_{\beta=1}^{M} p(a_{k+1} = a_i \mid a_k = a_\alpha) p(b_{k+1} = b_j \mid b_k = b_\beta) p(a_k = a_\alpha, b_k = b_\beta \mid Z_k) \tag{3.3.46}$$

The measurement sequence Z_k is not needed in the first two terms of (3.3.46) since $a_k = a_\alpha$ and $b_k = b_\beta$ are known at time interval k and are independent of Z_k. The following notation will be used to express the three terms of the summation (3.3.46)

$$\theta_a^{i\alpha} = p(a_{k+1} = a_i \mid a_k = a_\alpha) \tag{3.3.47}$$

$$\theta_b^{j\beta} = p(b_{k+1} = b_j \mid b_k = b_\beta) \tag{3.3.48}$$

$$w_k^{\alpha\beta} = p(a_k = a_\alpha, b_k = b_\beta \mid Z_k) \tag{3.3.49}$$

The terms (3.3.47) and (3.3.48) are Markov transition probabilities [22]; that is, $\theta_a^{i\alpha}$ is the conditional probability that $a = a_i$ at time interval k+1, given that $a = a_\alpha$ at time interval k. The $\theta_b^{j\beta}$ term is similarly defined. The term (3.3.49) is of the same form as (3.3.37) and is just the previous weighting term at the previous time interval k.

The denominator term of (3.3.37) is independent of ij. Therefore it is the same for each ij^{th} estimator and becomes a scale factor.

Combining (3.3.43) and (3.3.47) - (3.3.49), the weighting term (3.3.37) is written as

$$w_{k+1}^{ij} = p(a_i, b_j \mid Z_{k+1}) = C_{k+1}^{o} \mathrm{N}[\tilde{z}_{k+1}^{ij}, R_{k+1}] \sum_{\alpha=1}^{N} \sum_{\beta=1}^{M} \theta_a^{i\alpha} \theta_b^{j\beta} w_k^{\alpha\beta} \tag{3.3.50}$$

where C^o_{k+1} is a scale factor determined at each time interval such that

$$\sum_{i=1}^{N} \sum_{j=1}^{M} w^{ij}_{k+1} = 1 \tag{3.3.51}$$

guaranteeing that the sum of all the weighting terms (3.3.50) is equal to one. The mean and covariance of the Gaussian density function are available from the Kalman filter equation (3.3.35) and Kalman gain (3.3.34).

The equations for the modified Gaussian sum adaptive filter for the system of (3.3.1) - (3.3.4) are summarized in Table 2. The structure of the overall adaptive filter is a bank of Kalman filters operating in parallel, with each individual estimate multiplied by its own corresponding weighting term. The ij^{th} estimator based on the bias terms that most closely matches the actual bias terms of the modeled system will have a corresponding weighting term that tends closer to one, while the weights of the other mismatched estimators will tend towards zero. A block diagram of the adaptive filter is shown in Figure 2.

A simulation example is presented in the following section.

Table 2.

Modified Gaussian sum adaptive filter equations.

System:

$$x_{k+1} = \Phi x_k + \Gamma u_k + \Psi w_k$$
$$w_k = a_k + m_k$$
$$z_k = H x_k + v_k$$
$$v_k = b_k + n_k$$

Overall estimate:

$$\hat{x}_{k+1} = \sum_{i=1}^{N} \sum_{j=1}^{M} \hat{x}_{k+1}^{ij} p(a_i, b_j \mid Z_{k+1})$$

Kalman filter equation:

$$\hat{x}_{k+1}^{ij} = \Phi \hat{x}_k^{ij} + \Gamma u_k + \Psi a_i + K_{k+1}[z_{k+1} - b_j - H(\Phi \hat{x}_k^{ij} + \Gamma u_k + \Psi a_i)]$$

Kalman gain equations:

$$M_{k+1} = \Phi P_k \Phi^{\mathrm{T}} + \Psi Q \Psi^{\mathrm{T}}$$
$$K_{k+1} = M_{k+1} H^{\mathrm{T}} (H M_{k+1} H^{\mathrm{T}} + R)^{-1}$$
$$P_{k+1} = (I - K_{k+1} H) M_{k+1}$$

Weighting term:

$$w_{k+1}^{ij} = p(a_i, b_j \mid Z_{k+1}) = C_{k+1}^{o} p(z_{k+1} \mid a_i, b_j, Z_k) \sum_{\alpha=1}^{N} \sum_{\beta=1}^{M} \theta_a^{i\alpha} \theta_b^{j\beta} w_k^{\alpha\beta}$$

with

$$p(z_{k+1} \mid a_i, b_j, Z_k) = \mathrm{N}[H(\Phi \hat{x}_k^{ij} + \Gamma u_k + \Psi a_i) + b_j, H M_{k+1} H^{\mathrm{T}} + R]$$
$$\theta_a^{i\alpha} = p(a_{k+1} = a_i \mid a_k = a_\alpha)$$
$$\theta_b^{j\beta} = p(b_{k+1} = b_j \mid b_k = b_\beta)$$
$$w_k^{\alpha\beta} = p(a_k = a_\alpha, b_k = b_\beta \mid Z_k)$$

and scale factor C_{k+1}^{o} such that $\sum_{i=1}^{N} \sum_{j=1}^{M} w_{k+1}^{ij} = 1$

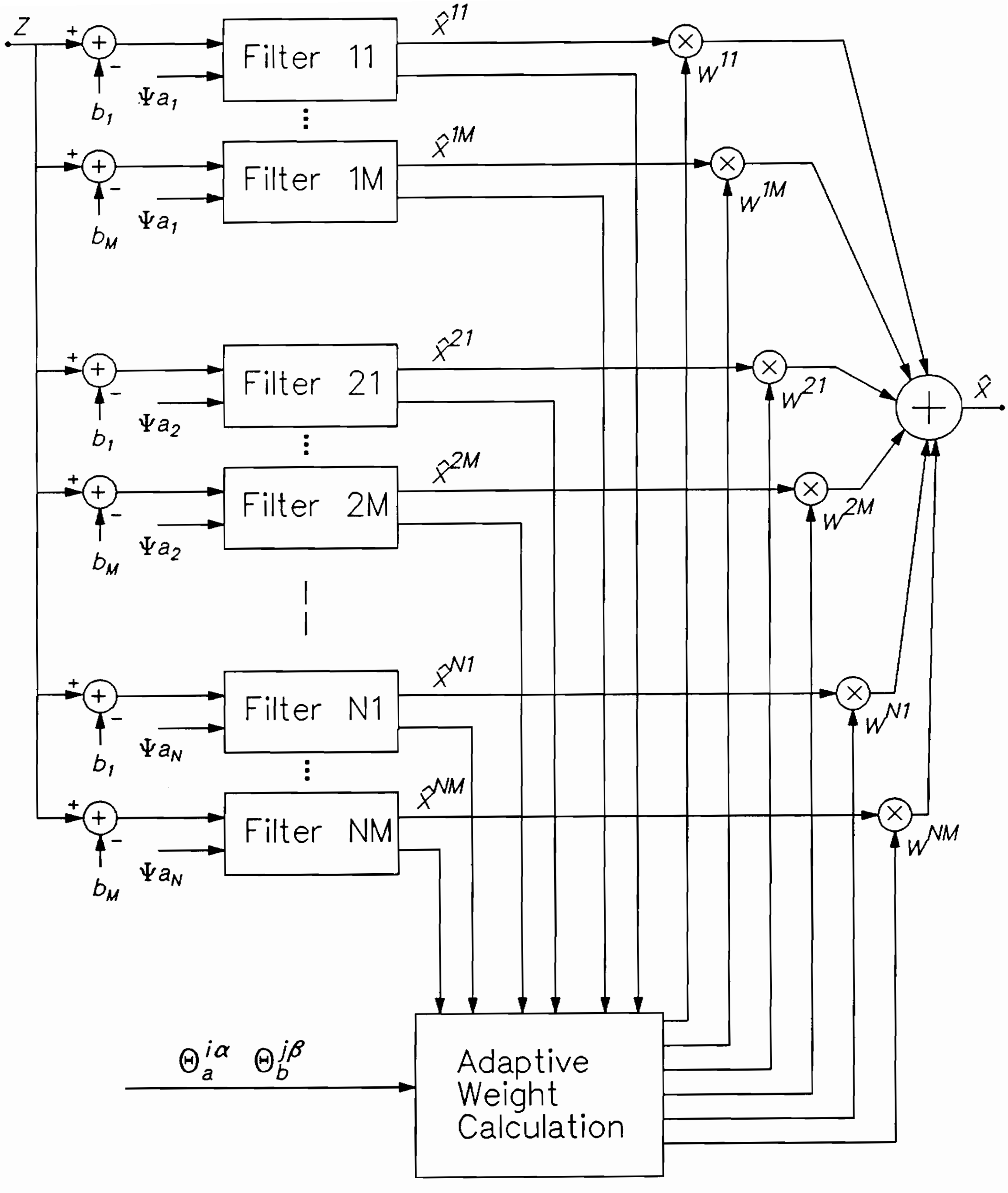


Figure 2. Modified Gaussian sum adaptive filter structure.

3.4 Simulation Example

An example illustrating the modified Gaussian sum estimation technique is now presented. A first-order system is used, modeled by the following discrete-time equations

plant:
$$x_{k+1} = e^{-\alpha T} x_k + (1 - e^{-\alpha T}) u_k + \frac{(1 - e^{-\alpha T})}{\alpha} w_k \tag{3.4.1}$$
$$w_k = a_k + m_k \tag{3.4.2}$$

measurement:
$$z_k = x_k + v_k \tag{3.4.3}$$
$$v_k = b_k + n_k \tag{3.4.4}$$

The value of α is 0.6 and the sample time $T = 1$ second. For simplicity, let the plant noise (3.4.2) be zero mean white Gaussian with variance $Q = 1.0$. Therefore, the randomly-switching plant bias term $a_k = 0$. The actual measurement noise, modeled by (3.4.4), has the Gamma density of (3.2.12), with a mean and variance of 4. The measurement bias term b_k can be randomly selected from the last three bias terms of Table 1, {2.553, 4.555, 6.933}. The first bias term of Table 1, {2.537}, is not used since it is so close in value to the second bias term. The measurement model noise term n_k is zero mean white Gaussian with variance $R = 1.0$. A deterministic input of $u_k = 10$ is used throughout the simulation. Figure 3a shows the measurement and state sequences. Note how the measurement is centered about 14, indicating a mean value for the Gamma density of 4.

The filter is initialized with equally-valued weighting terms

$$w_0^j = \frac{1}{3}, \quad \text{for } j = 1, 2, 3 \tag{3.4.5}$$

A Markov transition probability matrix $\Theta_b^{j\beta}$, consisting of $\theta_b^{j\beta}$ elements, is configured with a high probability that the bias term does not switch from one value to another, and a low probability that the bias term does switch, such as

$$\Theta_b^{j\beta} = \begin{bmatrix} .95 & .025 & .025 \\ .025 & .95 & .025 \\ .025 & .025 & .95 \end{bmatrix} \tag{3.4.6}$$

The initial value of the state and state estimate is $x_0 = \hat{x}_0 = 20$, and the initial value of the variance of the error $x_{k+1} - \hat{x}_{k+1}$ is $P_o = 100$. The overall state estimate and the state are shown in Figure 3b, with the error and overall state estimate shown in Figure 3c. Note how the error appears to be zero mean, thus showing that the MGS adaptive filter removes the bias effect of the nonzero mean Gamma measurement noise.

Figures 4a, b, and c show the weighting terms for each of the b_k bias terms. In order to lessen the noise of the weighting terms, a first-order lowpass filter

$$w_{k+1}^j = \lambda w_k^j + (1-\lambda) w_{k+1}^j \tag{3.4.7}$$

is used to smooth the weighting terms, where $\lambda = 0.7$. Figure 4d can be thought of as the overall measurement bias estimate $\hat{b}_k$ due to the nonzero mean Gamma measurement noise. Using (3.3.35) and (3.3.50), this overall bias estimate is part of the overall state estimate (3.3.12) and is written in this case with $a_k = 0$ as

$$\hat{b}_k = \sum_{j=1}^{M} w_{k+1}^j b_k^j \tag{3.4.8}$$

Note how this overall bias estimate approximately models the mean value of 4 of the Gamma measurement noise.

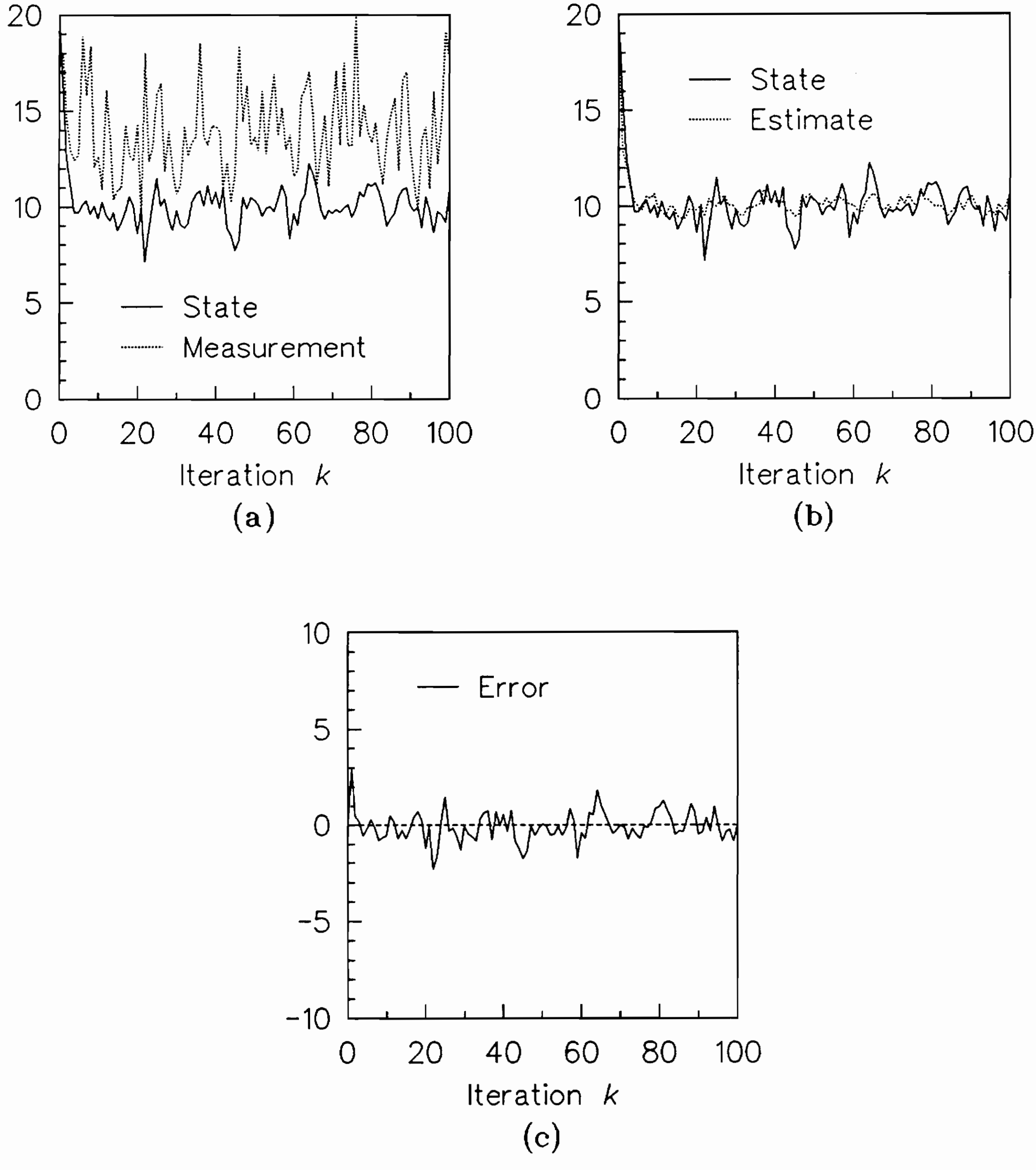


Figure 3. (a) Measurement and state simulation, (b) modified Gaussian sum estimate compared with state, (c) error = state − estimate.

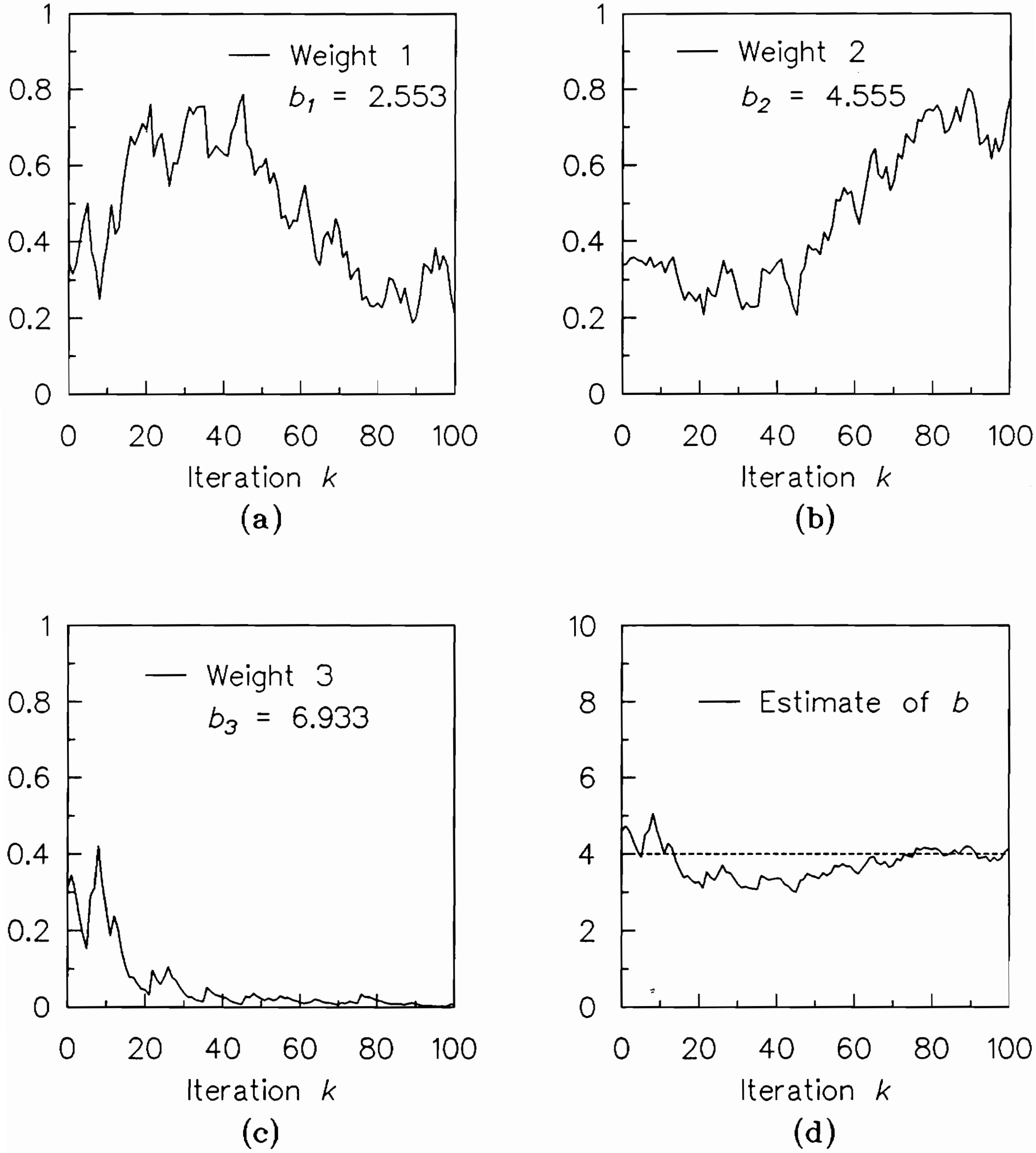


Figure 4. (a) Weight 1, (b) weight 2, (c) weight 3, (d) overall measurement bias estimate.

3.5 Comparison of Gaussian Sum and Modified Gaussian Sum Algorithms

The Gaussian sum (GS) algorithm of Alspach [6], [7] and the modified Gaussian sum (MGS) algorithm developed in this present work exhibit several similarities. Among them are using a Gaussian sum density approximation to model the actual nonGaussian densities, a structure formed by a bank of Kalman filters operating in parallel, and weighting terms that are nonlinear functions of the measurement data. However, there are several key differences between the two algorithms that set them apart.

One difference is that the GS algorithm produces an exponentially increasing number of Gaussian terms at each iteration, while the MGS algorithm produces only a fixed number of Gaussian terms. As a result, the GS algorithm requires an exponentially increasing number of individual Kalman filters operating in parallel, which is impossible to implement for any practical purpose. The MGS algorithm avoids the increasing computational and storage requirements of the GS algorithm.

This leads to a second difference between the two. The GS algorithm is optimal while the MGS algorithm is suboptimal. Tugnait [55] and Raisch [56] point out that in order for the state estimate

$$\hat{x}_{k+1} = E[x_{k+1} \mid Z_{k+1}] = \sum_{i=1}^{N} \sum_{j=1}^{M} \hat{x}_{k+1}^{ij} \; p(a_i, b_j \mid Z_{k+1}) \tag{3.5.1}$$

to be optimal, the full exponentially increasing number of individual Kalman filters is required. This leads, in turn, to the condition that the individual ij^{th} estimate

$$\hat{x}_{k+1}^{ij} = \int_{-\infty}^{\infty} x_{k+1} p(x_{k+1} \mid Z_{k+1}, a_i, b_j) \, dx_{k+1} \tag{3.5.2}$$

is optimal only if the density $p(x_{k+1} \mid Z_{k+1}, a_i, b_j)$ is conditioned on the full measurement sequence $Z_{k+1} = \{z_1, z_2, \ldots, z_{k+1}\}$ *and* the full switching-parameter bias term sequences $A^i_{k+1} = \{a_1, a_2, \ldots, a_{k+1}\}^i$ and $B^j_{k+1} = \{b_1, b_2, \ldots, b_{k+1}\}^j$ of present *and* past bias values for this particular ij^{th} estimate. Therefore, (3.5.2) would be written as

$$\hat{x}^{ij}_{k+1} = \int_{-\infty}^{\infty} x_{k+1} p(x_{k+1} \mid Z_{k+1}, A^i_{k+1}, B^j_{k+1})\, dx_{k+1} \tag{3.5.3}$$

and the overall state estimate (3.5.1) would be

$$\hat{x}_{k+1} = \sum_{i=1}^{N_{k+1}} \sum_{j=1}^{M_{k+1}} \hat{x}^{ij}_{k+1}\; p(A^i_{k+1}, B^j_{k+1} \mid Z_{k+1}) \tag{3.5.4}$$

The weighting term $p(A^i_{k+1}, B^j_{k+1} \mid Z_{k+1})$ is also now a function of the full bias term sequences. The upper limits on the summations are no longer fixed constants, but are now functions of the time iteration. At each iteration, a is randomly selected from a fixed set of N discrete values. This produces N^{k+1} possible A^i_{k+1} sequences at iteration k+1, with i running from $1, 2, \ldots, N^{k+1}$. Similarly, b is randomly selected from a fixed set of M discrete values, producing M^{k+1} possible B^j_{k+1} sequences, with j running from $1, 2, \ldots, M^{k+1}$.

Rewriting the measurement sequence as $Z_{k+1} = \{z_{k+1}, Z_k\}$, and the bias term sequences as $A^i_{k+1} = \{a^i_{k+1}, A^i_k\}$ and $B^j_{k+1} = \{b^j_{k+1}, B^j_k\}$, the density of (3.5.3) is expanded as

$$p(x_{k+1} \mid z_{k+1}, Z_k, a^i_{k+1}, A^i_k, b^j_{k+1}, B^j_k) = \frac{N_1 N_2 N_3 N_4}{D_1 D_2 D_3} \tag{3.5.5}$$

where

$$
\begin{aligned}
N_1 &= p(z_{k+1} \mid x_{k+1}, a^i_{k+1}, A^i_k, b^j_{k+1}, B^j_k, Z_k) \\
N_2 &= p(a^i_{k+1} \mid x_{k+1}, A^i_k, b^j_{k+1}, B^j_k, Z_k) \\
N_3 &= p(b^j_{k+1} \mid x_{k+1}, A^i_k, B^j_k, Z_k) \\
N_4 &= p(x_{k+1} \mid A^i_k, B^j_k, Z_k) \\
D_1 &= p(z_{k+1} \mid a^i_{k+1}, A^i_k, b^j_{k+1}, B^j_k, Z_k) \\
D_2 &= p(a^i_{k+1} \mid A^i_k, b^j_{k+1}, B^j_k, Z_k) \\
D_3 &= p(b^j_{k+1} \mid A^i_k, B^j_k, Z_k)
\end{aligned}
$$

The N_2 term can be reduced to the Markov transition probability $N_2 = p(a^i_{k+1} \mid a^i_k)$ since the A^i_{k+1} is a semi-Markov sequence. The N_3, D_2, and D_3 can be reduced in a similar fashion, producing

$$
\begin{aligned}
N_2 &= p(a^i_{k+1} \mid a^i_k), \qquad N_3 = p(b^j_{k+1} \mid b^j_k) \\
D_2 &= p(a^i_{k+1} \mid a^i_k), \qquad D_3 = p(b^j_{k+1} \mid b^j_k)
\end{aligned}
$$

Since $N_2 = D_2$ and $N_3 = D_3$, they each cancel out in (3.5.5) leaving

$$
p(x_{k+1} \mid z_{k+1}, Z_k, a^i_{k+1}, A^i_k, b^j_{k+1}, B^j_k) =
$$

$$
\frac{p(z_{k+1} \mid x_{k+1}, a^i_{k+1}, A^i_k, b^j_{k+1}, B^j_k, Z_k)\, p(x_{k+1} \mid a^i_k, A^i_{k-1}, b^j_k, B^j_{k-1}, Z_k)}{p(z_{k+1} \mid a^i_{k+1}, A^i_k, b^j_{k+1}, B^j_k, Z_k)} \tag{3.5.6}
$$

The density function of (3.5.2) used in the MGS algorithm is an approximation to (3.5.6), since it eliminates the past bias term sequences in (3.5.6). Using a similar development, the weighting term of (3.5.1) used in the MSG algorithm is an approximation to the weighting term of (3.5.4), since the past bias term sequences are eliminated here also. The MGS algorithm uses only the information contained in the present bias values, thereby leading to a suboptimal state estimate (3.5.1).

Even though the GS algorithm produces an optimal state estimate, it can never be found in practice since the number of GS terms increases exponentially from one iteration to the next according to (3.1.15) and (3.1.22), combined below

$$\xi'_{k+1} = q_k r_k \xi'_k \tag{3.5.7}$$

In order to reduce this growing number to some prescribed fixed number Ξ' at each iteration, Alspach [6] proposes a suboptimal estimate based on observed mechanisms that allow terms to be neglected or combined. At iteration k, he drops weighting terms (3.1.16) and (3.1.23) that fall below some prescribed threshold. He also combines several GS terms of (3.1.11) into one term if their means (3.1.12) and variances (3.1.13) have become approximately equal. These two operations are used until only $\xi'_k = \Xi'$ significant terms remain. Then, at iteration $k+1$, the number of GS terms grows again to the larger number ξ'_{k+1} according to (3.5.7), is reduced by eliminating and combining terms until $\xi'_{k+1} = \Xi'$, and the cycle repeats for all subsequent iterations. This method results in a substantial reduction in the number of GS terms with a negligible effect on the $p(x_{k+1} \mid Z_{k+1})$ density approximation.

On the surface, it appears that Alspach's suboptimal Gaussian sum (SGS) algorithm can achieve performance equal to that of the MGS algorithm, since only a fixed number of terms is used to produce the state estimate. However, in order for the SGS algorithm to produce the fixed number of terms Ξ', the larger number of terms $q_k r_k \Xi'$ must first be generated, requiring $q_k r_k \Xi'$ parallel filters. The MGS algorithm requires only Ξ' filters at each iteration to produce a Ξ' term state estimate. The computational savings of the MGS algorithm is considerable. Even if the plant and measurement noise processes (3.1.5) and (3.1.6) were approximated using the minimum number of two terms each ($q_k = r_k = 2$), the SGS algorithm would require a number of filters $q_k r_k = 4$ times greater than the MGS algorithm.

A third difference is that the GS algorithm must always use time-varying Kalman gains while the MGS algorithm allows the use of steady-state Kalman gains. The growing memory problem of the GS algorithm prevents steady-state Kalman gains to be calculated off-line before running the state estimator. The SGS algorithm must also use time-varying gains because the elimination and combination of terms is performed at each iteration. Therefore, the Kalman gains could never be computed in advance. The MGS algorithm uses the same fixed number of filters at each iteration. With constant covariances Q and R of the plant noise (3.3.2) and measurement noise (3.3.4) used in the gain equations for each ij^{th} filter, steady-state Kalman gains could be calculated off-line. Even if Q and R were different for each ij^{th} filter, they would then be modeled as semi-Markov processes, with values randomly selected from fixed sets of discrete values. Since all the covariance values would be known in advance, steady-state Kalman gains could again be calculated off-line.

A fourth difference between the two algorithms is the use of Markov transition probabilities in the weighting term of the MGS algorithm, restated here for convenience

$$w_{k+1}^{ij} = C_{k+1}^{o} \mathrm{N}[\tilde{z}_{k+1}^{ij}, R_{k+1}] \sum_{\alpha=1}^{N} \sum_{\beta=1}^{M} \theta_a^{i\alpha} \theta_b^{j\beta} w_k^{\alpha\beta} \tag{3.5.8}$$

The weighting term of the GS algorithm at iteration k+1 is

$$\alpha_{k+1_j} = \frac{\alpha'_{k+1_i} \gamma_{k+1_m} W_{k+1_j}}{\sum_{i=1}^{\xi'_{k+1}} \alpha'_{k+1_i} \gamma_{k+1_m} W_{k+1_j}} \tag{3.5.9}$$

Define the scale factor as

$$\lambda^o_{k+1} = \frac{1}{\sum_{i=1}^{\xi'_{k+1}} \alpha'_{k+1_i} \gamma_{k+1_m} W_{k+1_j}} \tag{3.5.10}$$

Using the definition of the α'_{k+1_i} term

$$\alpha'_{k+1_i} = \alpha_{k_l} \beta_{k_n} \tag{3.5.11}$$

and (3.5.10) in (3.5.9) results in

$$\alpha_{k+1_j} = \lambda^o_{k+1} W_{k+1_j} [\gamma_{k+1_m} \beta_{k_n}] \alpha_{k_l} \tag{3.5.12}$$

The GS weighting term (3.5.12) now has a similar recursive form as the MGS weighting term (3.5.8), with λ^o_{k+1} and C^o_{k+1} being the respective scale factors, and W_{k+1_j} and $N[\tilde{z}^{ij}_{k+1}, R_{k+1}]$ being the respective Gaussian density functions based on the measurement data. The main difference between the two weighting terms is the Markov transition probabilities in (3.5.8) and the $\gamma_{k+1_m} \beta_{k_n}$ factor in (3.5.12). The Markov probabilities govern the chances that a bias term is going to switch from one value to another. The $\gamma_{k+1_m} \beta_{k_n}$ factor is formed from the weighting terms of the Gaussian sum approximations of the nonGaussian densities (3.1.5) and (3.1.6), and has no meaning similar to the Markov probabilities. Of course, the current number of α_{k+1_j} terms is larger than the previous number of α_{k_l} terms at the each iteration, while the number of w^{ij}_{k+1} terms remains fixed at each iteration.

4.0 APPLICATION TO A MODAL SYSTEM

The modified Gaussian sum (MGS) adaptive filtering technique of chapter 3 is now applied to a modal system problem. Large, heavily damped structures modeled by a plant with complex modes are considered in [1]. The input to the plant is unknown, but its frequency characteristics are known and the input is always present. The nonGaussian nature of the input signal is modeled using the signal model of chapter 3, and the MGS adaptive filter for this system is developed. A simulation example using the MGS adaptive filter is given and the results are compared to those produced from an augmented Kalman filter based on a system model assuming a narrowband Gaussian input signal. A necessary condition for effective estimation is derived. Alternate estimation procedures are developed to compensate for situations when this condition is not met. Several configurations are simulated and their performance results are analyzed and compared.

4.1 Modal System and Filter Development

The MGS algorithm is applied to a modal system with a nonGaussian input signal and Gaussian plant and measurement noise. The theoretical foundation for this system, which serves as a model for large, heavily damped structures, is found in [1] and briefly outlined in appendix A. Using a zero-order-hold model, the resulting discrete-time system equations with added noise terms are

$$x_{k+1} = \Phi x_k + \Gamma u_k + \Psi w_k \tag{4.1.1}$$

$$z_k = H x_k + D u_k + v_k \tag{4.1.2}$$

$$u_k = b_k + n_k \tag{4.1.3}$$

where

x_{k+1}	is the state vector
z_k	is the measurement vector
w_k	is a zero mean white Gaussian plant noise process with covariance Q
v_k	is a zero mean white Gaussian measurement noise process with covariance R
Φ, Γ, Ψ, H, D	are the respective constant transition matrices
u_k	is the vector Gaussian sum (GS) signal model of the actual nonGaussian input signal, comprised of semi-Markov bias vector b_k, and zero mean white Gaussian noise n_k with semi-Markov covariance S_k

and the random quantities x_k, w_k, v_k, b_k, and n_k are assumed to be mutually statistically independent.

There are two important differences between the modal system of (4.1.1) - (4.1.3) and the general system of chapter 3, (3.3.1) - (3.3.4). The first difference is that in the modal system, the signal u_k is an input not only to the plant, but is also fed through to the measurement z_k. Inserting (4.1.3) into (4.1.1) and (4.1.2) produces modified plant and measurement equations

$$x_{k+1} = \Phi x_k + \Gamma b_k + \Psi w'_k \tag{4.1.4}$$

$$z_k = H x_k + D b_k + v'_k \tag{4.1.5}$$

where

$$w'_k = \Gamma n_k + \Psi w_k \tag{4.1.6}$$

$$v'_k = D n_k + v_k \tag{4.1.7}$$

Because n_k appears in both (4.1.6) and (4.1.7), w'_k and v'_k are correlated Gaussian random processes. At first glance, this correlation may seem to be an obstacle in developing the Kalman filter equations for this model, since the Kalman filter development requires that the plant and measurement noise processes be uncorrelated. However, what is really necessary in this situation is for w'_{k-1} and v'_k to be uncorrelated [57]. Rewriting (4.1.4) for x_k gives

$$x_k = \Phi x_{k-1} + \Gamma b_{k-1} + \Psi w'_{k-1} \tag{4.1.8}$$

and substituting (4.1.8) into (4.1.5) gives

$$z_k = H(\Phi x_{k-1} + \Gamma b_{k-1} + \Psi w'_{k-1}) + D b_k + v'_k \tag{4.1.9}$$

thereby showing that any covariance calculations involving (4.1.9) will involve the correlation between w'_{k-1} and v'_k, and not between w'_k and v'_k. The correlation between w'_{k-1} and v'_k is given by

$$
\begin{aligned}
E[w'_{k-1}v'^{\mathrm{T}}_k] &= E[(\Gamma n_{k-1} + \Psi w_{k-1})(Dn_k + v_k)^{\mathrm{T}}] \\
&= \Gamma E[n_{k-1}n_k^{\mathrm{T}}]D^{\mathrm{T}} + \Gamma E[n_{k-1}v_k^{\mathrm{T}}] + \Psi E[w_{k-1}n_k^{\mathrm{T}}]D^{\mathrm{T}} + \Psi E[w_{k-1}v_k^{\mathrm{T}}] \quad (4.1.10) \\
&= 0
\end{aligned}
$$

since w_k, v_k, and n_k are assumed to be mutually statistically independent.

The second difference is that the covariance S_k of the GS signal model is no longer constant, but can vary with time. The bias vector b_k and covariance S_k are both semi-Markov processes with state transitions governed by a single Markov transition probability matrix. That is, values of b_k and S_k will be paired together and will randomly switch from one pair to another. The switching covariance can be thought of as changing the power of the Gaussian process.

As in chapter 3, the optimal estimate $\hat{x}_{k+1}$ of the state vector is found by minimizing the mean-square error between x_{k+1} and $\hat{x}_{k+1}$. This results in the conditional mean estimate

$$
\hat{x}_{k+1} = E[x_{k+1} \mid Z_{k+1}] = \int_{-\infty}^{\infty} x_{k+1} p(x_{k+1} \mid Z_{k+1})\, dx_{k+1} \tag{4.1.11}
$$

where Z_{k+1} is the current measurement sequence $\{z_1, z_2, \ldots, z_{k+1}\}$. The conditional density function of (4.1.11) can be written as the ratio of the corresponding joint and marginal densities

$$
p(x_{k+1} \mid Z_{k+1}) = \frac{p(x_{k+1}, Z_{k+1})}{p(Z_{k+1})} \tag{4.1.12}
$$

The bias vector b_k and covariance matrix S_k are explicitly brought into (4.1.12) by considering the joint density $p(x_{k+1}, Z_{k+1})$ to be a marginal density found from $p(x_{k+1}, Z_{k+1}, b_k = b_i, S_k = S_i)$ by summing over the b_k and S_k terms

$$p(x_{k+1} \mid Z_{k+1}) = \frac{\sum_{i=1}^{M} p(x_{k+1}, Z_{k+1}, b_k = b_i, S_k = S_i)}{p(Z_{k+1})} \tag{4.1.13}$$

After some additional algebraic manipulation, the optimal estimate of (4.1.11) becomes

$$\hat{x}_{k+1} = \sum_{i=1}^{M} \hat{x}^i_{k+1} p(b_i, S_i \mid Z_{k+1}) \tag{4.1.14}$$

where $\hat{x}^i_{k+1}$ is the conditional mean estimate of x_{k+1} given that $b_k = b_i$ and $S_k = S_i$, denoted by

$$\hat{x}^i_{k+1} = \int_{-\infty}^{\infty} x_{k+1} p(x_{k+1} \mid Z_{k+1}, b_i, S_i)\, dx_{k+1} \tag{4.1.15}$$

and the weighting factor $p(b_i, S_i \mid Z_{k+1})$ is the probability that $b_k = b_i$ and $S_k = S_i$ given the current measurement sequence. The overall estimate of (4.1.14) is then a weighted sum of M individual estimates, each based on a particular pair of parameters b_i and S_i.

Each individual estimate (4.1.15) is found by a modified Kalman filter. Using the same method in chapter 3, the Kalman filter equation for (4.1.15) is

$$\hat{x}^i_{k+1} = \Phi\hat{x}^i_k + \Gamma b_i + K^i_{k+1}[z_{k+1} - Db_i - H(\Phi\hat{x}^i_k + \Gamma b_i)] \tag{4.1.16}$$

with covariance and gain equations

$$M^i_{k+1} = \Phi P^i_k \Phi^{\mathrm{T}} + \Gamma S_i \Gamma^{\mathrm{T}} + \Psi Q \Psi^{\mathrm{T}} \tag{4.1.17}$$

$$K^i_{k+1} = M^i_{k+1} H^{\mathrm{T}} (H M^i_{k+1} H^{\mathrm{T}} + D S_i D^{\mathrm{T}} + R)^{-1} \tag{4.1.18}$$

$$P_{k+1}^i = (I - K_{k+1}^i H) M_{k+1}^i \tag{4.1.19}$$

Note that (4.1.17) - (4.1.19) are now dependent on S_i so that an individual set of these equations has to be computed for each i^{th} estimator (4.1.16), whereas in chapter 3 only one set was computed and used for all individual estimators.

As in chapter 3, the next step in producing the overall estimate (4.1.14) is to find the weighting term $p(b_i, S_i \mid Z_{k+1})$. Using Bayes' theorem, and writing Z_{k+1} as $\{Z_k, z_{k+1}\}$, the weighting term becomes

$$p(b_i, S_i \mid Z_{k+1}) = \frac{p(z_{k+1} \mid b_i, S_i, Z_k) p(b_i, S_i \mid Z_k)}{p(z_{k+1} \mid Z_k)} \tag{4.1.20}$$

The denominator term is independent of i and is a scale factor that ensures that the sum of the weights (4.1.20) at each iteration is equal to one. The first term of the numerator can be modeled as a Gaussian density if the bias and covariance terms switch slowly compared to the sample interval k. This is the same assumption made in chapter 3. This conditional density is

$$p(z_{k+1} \mid b_i, S_i, Z_k) = \mathrm{N}[H(\Phi \hat{x}_k^i + \Gamma b_i) + D b_i,\ H M_{k+1}^i H^{\mathrm{T}} + D S_i D^{\mathrm{T}} + R] \tag{4.1.21}$$

The mean and covariance of (4.1.21) are readily available from the Kalman filter equation (4.1.16) and the Kalman gain (4.1.18).

The second term of the numerator can be rewritten to explicitly include the sample time for the bias and covariance terms

$$p(b_i, S_i \mid Z_k) = p(b_{k+1} = b_i, S_{k+1} = S_i \mid Z_k) \tag{4.1.22}$$

Using Bayes' rule and the definition of marginal densities, (4.1.22) is expanded into

$$p(b_i, S_i \mid Z_k) = \sum_{j=1}^{M} p(b_{k+1}{=}b_i, S_{k+1}{=}S_i \mid b_k{=}b_j, S_k{=}S_j) p(b_k{=}b_j, S_k{=}S_j \mid Z_k) \quad (4.1.23)$$

As in chapter 3, the following notation is used to express the terms of (4.1.23)

$$\theta_{bS}^{ij} = p(b_{k+1} = b_i, S_{k+1} = S_i \mid b_k = b_j, S_k = S_j) \quad (4.1.24)$$

$$w_k^j = p(b_k = b_j, S_k = S_j \mid Z_k) \quad (4.1.25)$$

The term (4.1.24) is a Markov transition probability [22]; that is, θ_{bS}^{ij} is the conditional probability that $b = b_i$ and $S = S_i$ at time $k+1$, given that $b = b_j$ and $S = S_j$ at time k. The term (4.1.25) is of the same form as (4.1.20) and is just the previous weighting term at the previous time interval k.

Combining (4.1.21), (4.1.24), and (4.1.25), the weighting term (4.1.20) is written as

$$w_{k+1}^i = p(b_i, S_i \mid Z_{k+1}) = C_{k+1}^o p(z_{k+1} \mid b_i, S_i, Z_k) \sum_{j=1}^{M} \theta_{bS}^{ij} w_k^j \quad (4.1.26)$$

where C_{k+1}^o is a scale factor determined at each time interval such that the sum of all the weighting terms (4.1.26) is equal to one.

The equations for the modified Gaussian sum adaptive filter for the system of (4.1.1) - (4.1.3) are summarized in Table 3. The structure of the overall adaptive filter is again a bank of Kalman filters operating in parallel, with each individual estimate multiplied by its own corresponding weighting term. The i^{th} estimator based on the bias and covariance pair that most closely matches the actual bias and covariance pair of the modeled system will have a corresponding weighting term that tends closer to one, while the weights of the other mismatched estimators will tend towards zero. A block diagram of the MGS adaptive filter for this modal system is shown in Figure 5.

Table 3.

Modified Gaussian sum adaptive filter equations for a modal system.

System:

$$x_{k+1} = \Phi x_k + \Gamma u_k + \Psi w_k$$
$$z_k = H x_k + D u_k + v_k$$
$$u_k = b_k + n_k$$

Overall estimate:

$$\hat{x}_{k+1} = \sum_{i=1}^{M} \hat{x}^i_{k+1} p(b_i, S_i \mid Z_{k+1})$$

Kalman filter equation:

$$\hat{x}^i_{k+1} = \Phi \hat{x}^i_k + \Gamma b_i + K^i_{k+1}[z_{k+1} - D b_i - H(\Phi \hat{x}^i_k + \Gamma b_i)]$$

Kalman gain equations:

$$M^i_{k+1} = \Phi P^i_k \Phi^{\mathrm{T}} + \Gamma S_i \Gamma^{\mathrm{T}} + \Psi Q \Psi^{\mathrm{T}}$$
$$K^i_{k+1} = M^i_{k+1} H^{\mathrm{T}} (H M^i_{k+1} H^{\mathrm{T}} + D S_i D^{\mathrm{T}} + R)^{-1}$$
$$P^i_{k+1} = (I - K^i_{k+1} H) M^i_{k+1}$$

Weighting term:

$$w^i_{k+1} = p(b_i, S_i \mid Z_{k+1}) = C^o_{k+1} p(z_{k+1} \mid b_i, S_i, Z_k) \sum_{j=1}^{M} \theta^{ij}_{bS} w^j_k$$

with

$$p(z_{k+1} \mid b_i, S_i, Z_k) = \mathrm{N}[H(\Phi \hat{x}^i_k + \Gamma b_i) + D b_i,\ H M^i_{k+1} H^{\mathrm{T}} + D S_i D^{\mathrm{T}} + R]$$
$$\theta^{ij}_{bS} = p(b_{k+1} = b_i, S_{k+1} = S_i \mid b_k = b_j, S_k = S_j)$$
$$w^j_k = p(b_k = b_j, S_k = S_j \mid Z_k)$$

and scale factor C^o_{k+1} such that $\sum_{i=1}^{M} w^i_{k+1} = 1$

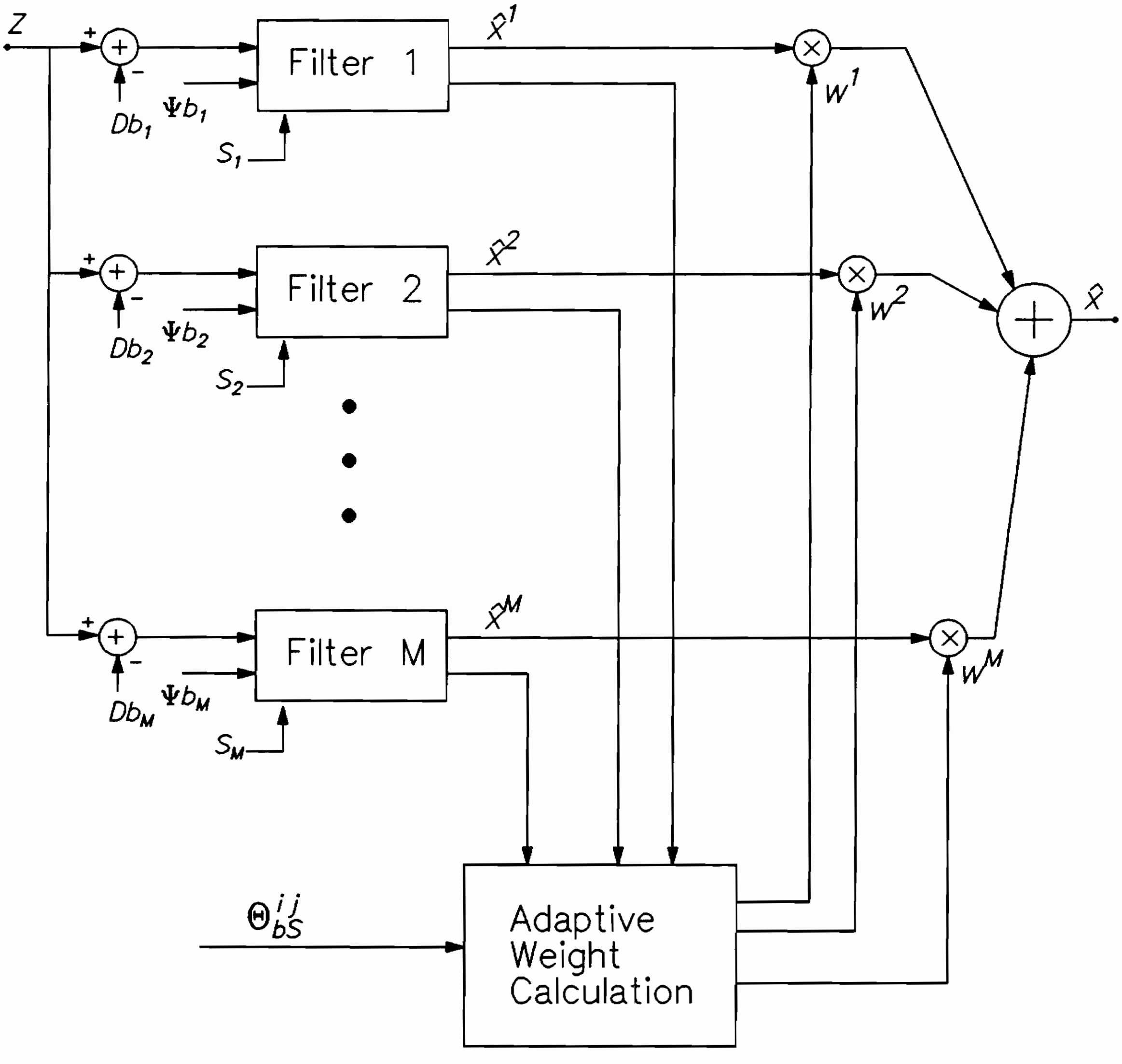


Figure 5. Modified Gaussian sum adaptive filter structure for a modal system.

4.2 Simulation Example

An example applying the MGS algorithm to a modal system with a nonGaussian input signal is now presented. The results are compared to those produced from an augmented system model assuming that the input signal is Gaussian. A second-order system is used, modeled by the transfer function

$$G(s) = \frac{s^2 + cs + d}{s^2 + as + b} \tag{4.2.1}$$

and in continuous-time state space form as

$$\dot{x}(t) = Ax(t) + Bu(t) \tag{4.2.2}$$

$$y(t) = Cx(t) + Du(t) \tag{4.2.3}$$

The system matrices are defined by

$$A = \begin{bmatrix} 0 & 1 \\ -b & -a \end{bmatrix} \quad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad C = \begin{bmatrix} d-b & c-a \end{bmatrix} \quad D = \begin{bmatrix} 1 \end{bmatrix} \tag{4.2.4}$$

where the resonant frequency of the system is $f_o = 20$ Hz and the damping coefficient is $\zeta_o = 0.1$ so that

$$a = 2\zeta_o(2\pi f_o) \qquad b = (2\pi f_o)^2 \qquad c = 5a \qquad d = 5b \tag{4.2.5}$$

Using a zero-order-hold model with a sample time of $T = \frac{1}{60}$ second, the equivalent discrete-time system with added noise terms becomes

$$x_{k+1} = \Phi x_k + \Gamma u_k + \Psi w_k \tag{4.2.6}$$

$$z_k = H x_k + D u_k + v_k \tag{4.2.7}$$

where

$$\Phi = \begin{bmatrix} -0.3271 & 0.0057 \\ -89.2411 & -0.4691 \end{bmatrix} \quad \Gamma = \begin{bmatrix} 0.0001 \\ 0.0057 \end{bmatrix} \quad \Psi = \begin{bmatrix} 1 \\ 100 \end{bmatrix}$$
$$H = \begin{bmatrix} 63165 & 100.53 \end{bmatrix} \quad D = \begin{bmatrix} 1 \end{bmatrix} \tag{4.2.8}$$

The plant noise w_k is zero-mean white Gaussian with covariance $Q = 10^{-10}$, and is uncorrelated with measurement noise v_k, which is also zero-mean white Gaussian with covariance $R = 10^{-2}$.

The nonGaussian input signal is a stochastic FM signal generated by frequency modulating a sinusoid with a Gaussian process. The form of the signal is

$$u(t) = A_u \sin(2\pi f_u t + k_u \int_0^t m(\tau)\, d\tau) \tag{4.2.9}$$

where the amplitude is $A_u = 5$, the carrier frequency is $f_u = 2$ Hz, the modulation index is $k_u = 10$, and $m(t)$ is zero-mean white Gaussian noise with variance 1. Figure 6a shows the FM signal $u(t)$ for 1 second. The probability density function of $u(t)$ has the form [58]

$$p(u) = \begin{cases} \dfrac{1}{\pi\sqrt{A_u^2 - u^2}} & |u| < A_u \\ 0 & |u| \geq A_u \end{cases} \tag{4.2.10}$$

Figure 6b shows a plot of the density function of (4.2.10), while Figure 6c shows a normalized histogram of a 10 second sample of $u(t)$. Note how closely the

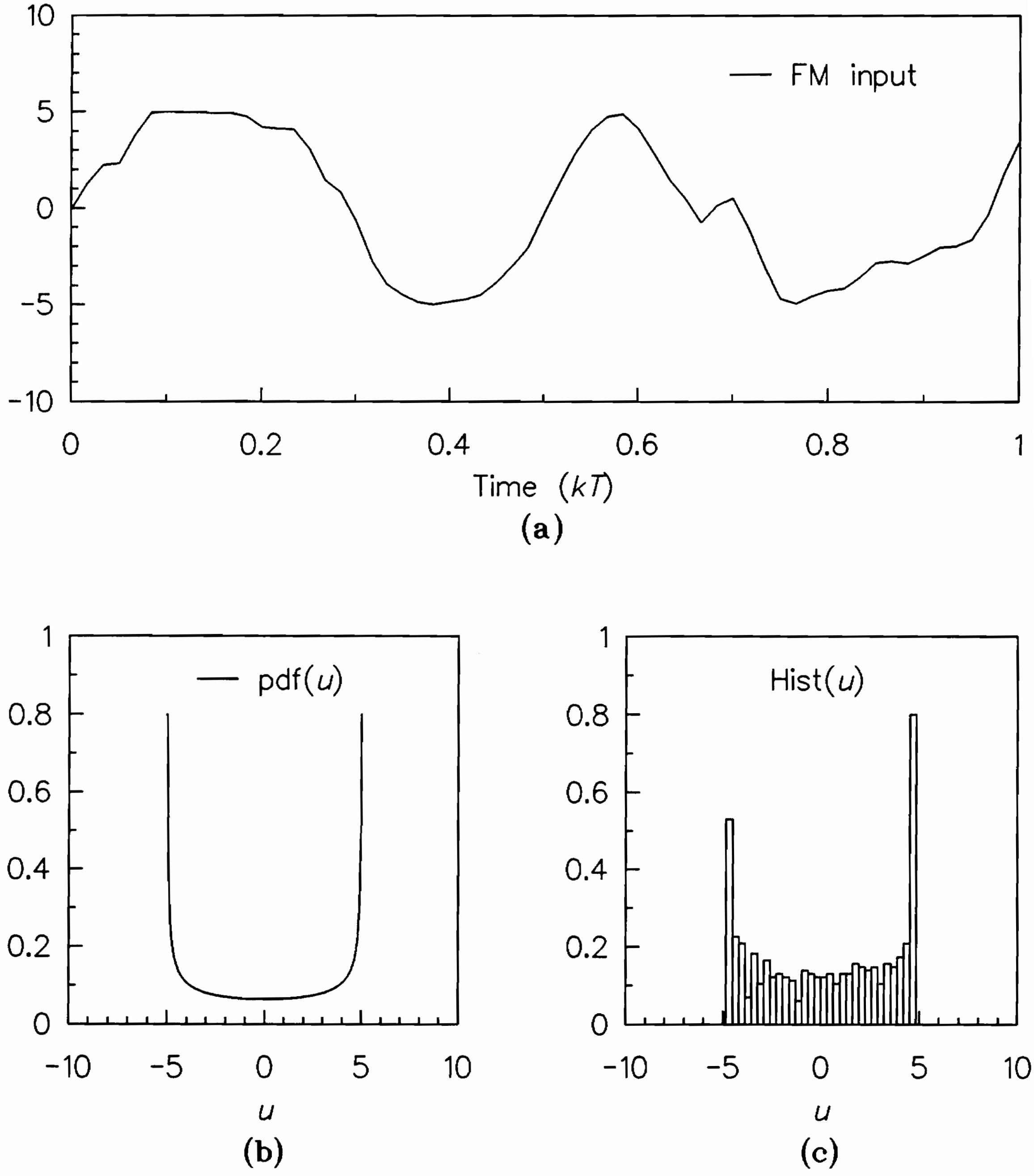


Figure 6. (a) Stochastic FM signal $u(t)$, (b) probability density function of $u(t)$, (c) normalized histogram of $u(t)$.

histogram matches the actual density function.

The nonGaussian input signal $u(t)$ of (4.2.9) will be modeled in discrete-time using the Gaussian sum (GS) signal model described in section 4.1

$$u_k = b_k + n_k \tag{4.2.11}$$

with GS density approximation described in chapter 3

$$p(u) = \sum_{i=1}^{M} P_i \mathrm{N}[b_i, S_i] \tag{4.2.12}$$

Figure 7 compares the nonGaussian density function (4.2.10) of the stochastic FM input signal with a three-term GS density approximation (4.2.12) minimizing the L^2 norm (3.2.11). Table 4 lists the values of P_i, b_i, and S_i used in the Gaussian sum approximation and the MGS adaptive filter.

Table 4.

Gaussian sum P_i, b_i, and S_i values.

i	P_i	b_i	S_i
1	0.1	−5.0	0.04
2	0.1	5.0	0.04
3	0.8	0.0	4.28

It is assumed that the nonGaussian input signal $u(t)$ is much larger than both the Gaussian plant noise w_k and the Gaussian measurement noise v_k. However, allowing the power of $u(t)$ to be larger than the power (covariance Q) of w_k and

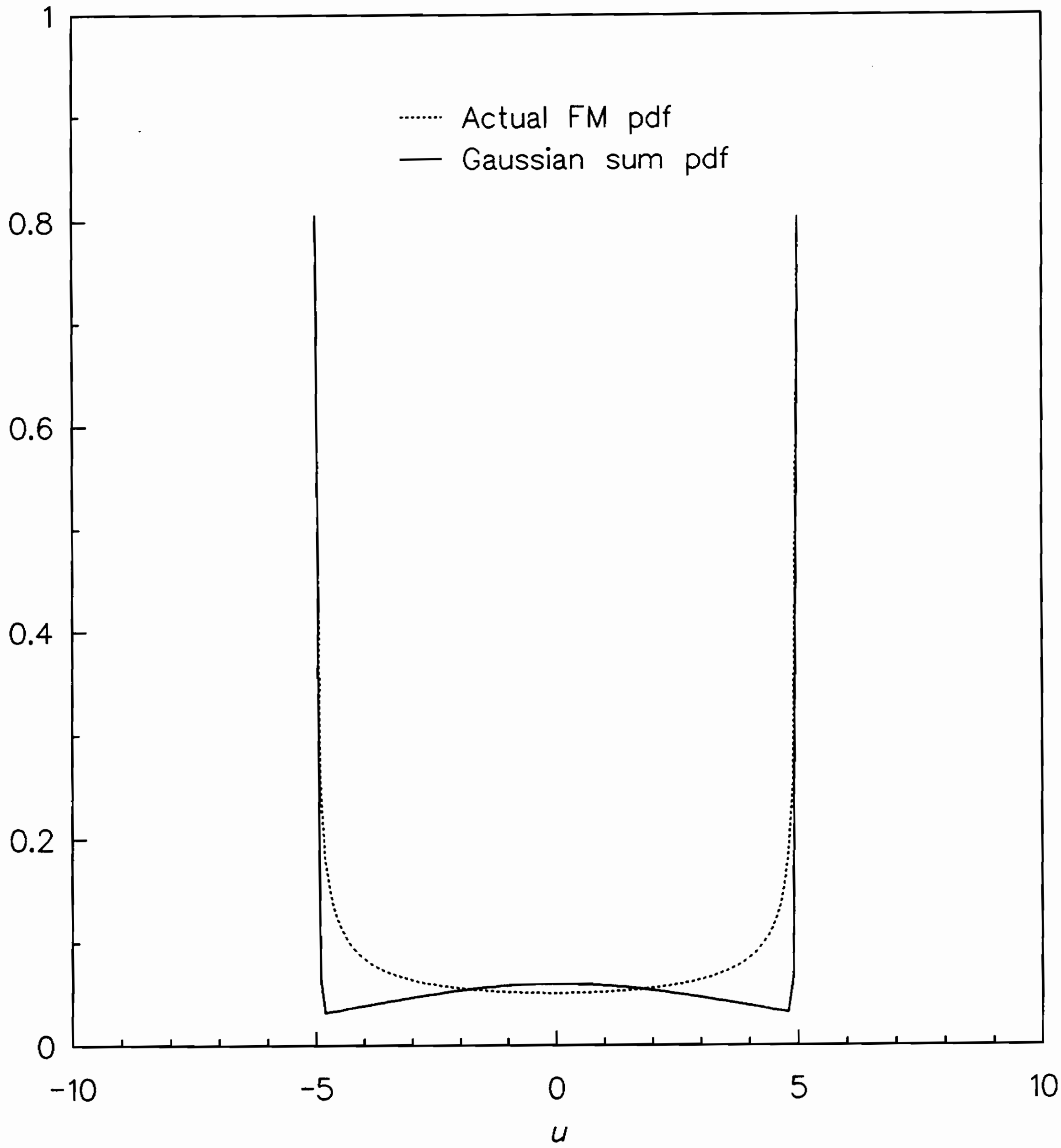


Figure 7. Stochastic FM signal and Gaussian sum probability density functions.

power (covariance R) of v_k is not enough to ensure good MGS simulation results. The nonGaussian input signal $u(t)$ is modeled by u_k (4.2.11), which will take on different b_i, S_i pairs at various times. The lowest value of S_i from Table 4 is $S_i = 0.04$. From covariance equation (4.1.18) and Kalman gain (4.1.19), repeated here for convenience

$$M_{k+1}^i = \Phi P_k^i \Phi^{\mathrm{T}} + \Gamma S_i \Gamma^{\mathrm{T}} + \Psi Q \Psi^{\mathrm{T}}$$
$$K_{k+1}^i = M_{k+1}^i H^{\mathrm{T}} (H M_{k+1}^i H^{\mathrm{T}} + D S_i D^{\mathrm{T}} + R)^{-1}$$

it is apparent that for good MGS simulation results, covariances Q and R must be selected such that the elements of $\Gamma S_i \Gamma^{\mathrm{T}}$ are larger than the elements of $\Psi Q \Psi^{\mathrm{T}}$, and the elements of $D S_i D^{\mathrm{T}}$ are larger than the elements of R. This specification is met using $S_i = 0.04$, $Q = 10^{-10}$, $R = 10^{-2}$, and Γ, Ψ, and D from (4.2.8), giving

$$\Gamma S_i \Gamma^{\mathrm{T}} = \begin{bmatrix} 2.8 \times 10^{-10} & 1.9 \times 10^{-8} \\ 1.9 \times 10^{-8} & 1.3 \times 10^{-6} \end{bmatrix} \qquad \Psi Q \Psi^{\mathrm{T}} = \begin{bmatrix} 1.0 \times 10^{-10} & 1.0 \times 10^{-8} \\ 1.0 \times 10^{-8} & 1.0 \times 10^{-6} \end{bmatrix}$$

$$D S_i D^{\mathrm{T}} = 0.04 \qquad R = 0.01$$

For comparison purposes, the nonGaussian input signal $u(t)$ of (4.2.9) will also be modeled as a narrowband Gaussian signal, generated by passing white Gaussian noise through a linear, time-invariant filter. This shaping filter is designed using the frequency characteristics of $u(t)$. Figure 8 shows the power spectral density of $u(t)$ and the power spectral density of a second-order shaping filter, modeled by the transfer function

$$G_{nb}(s) = \frac{N_{nb} s}{s^2 + a_{nb} s + b_{nb}} \tag{4.2.13}$$

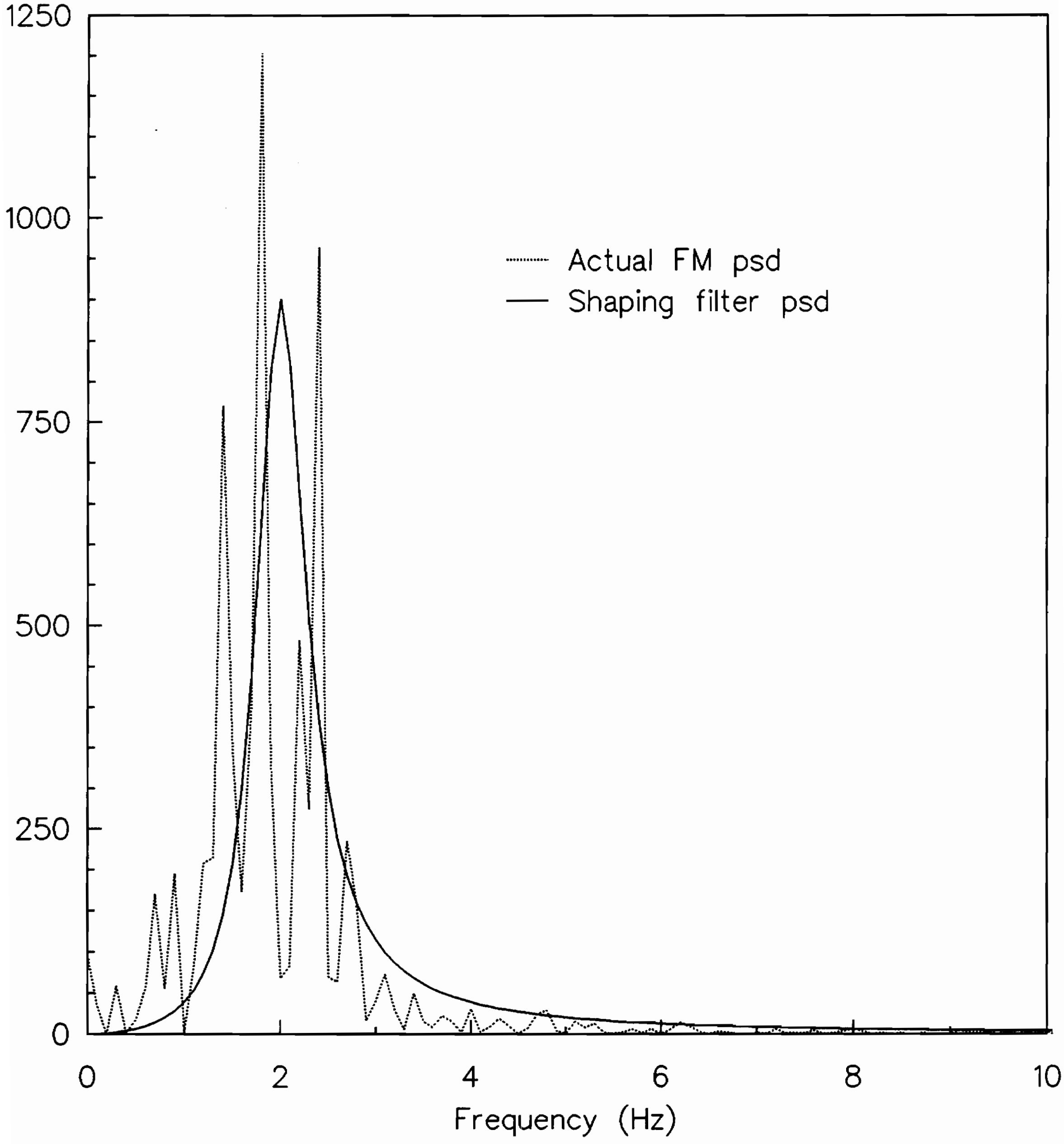


Figure 8. Stochastic FM signal and shaping filter power spectral densities.

and in continuous-time state space form as

$$\dot{x}_{nb}(t) = A_{nb}x_{nb}(t) + B_{nb}w_{nb}(t) \tag{4.2.14}$$

$$u_{nb}(t) = C_{nb}x_{nb}(t) \tag{4.2.15}$$

The system matrices are defined by

$$A_{nb} = \begin{bmatrix} 0 & 1 \\ -b_{nb} & -a_{nb} \end{bmatrix} \quad B_{nb} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad C_{nb} = \begin{bmatrix} 0 & 1 \end{bmatrix} \tag{4.2.16}$$

with resonant frequency $f_{nb} = 2$ Hz and damping coefficient $\zeta_{nb} = \frac{1}{2\pi}$ such that

$$a_{nb} = 2\zeta_{nb}(2\pi f_{nb}) = 4 \qquad b_{nb} = (2\pi f_{nb})^2 = 157.91 \tag{4.2.17}$$

and w_{nb} is zero-mean white Gaussian noise with variance $N_{nb}^2 = (120)^2$.

The Gaussian input model $u_{nb}(t)$ of (4.2.15) is now assumed to model the stochastic FM input $u(t)$ of (4.2.9) by augmenting the system equations (4.2.2) and (4.2.3) with the shaping filter equations (4.2.14) and (4.2.15), giving

$$\dot{x}_a(t) = A_a x_a(t) + B_a w_{nb}(t) \tag{4.2.18}$$

$$y_a(t) = C_a x_a(t) \tag{4.2.19}$$

The system matrices of (4.2.18) and (4.2.19) are defined by

$$x_a(t) = \begin{bmatrix} x(t) \\ x_{nb}(t) \end{bmatrix} \quad A_a = \begin{bmatrix} A & BC_{nb} \\ 0 & A_{nb} \end{bmatrix} \quad B_a = \begin{bmatrix} 0 \\ B_{nb} \end{bmatrix} \quad C_a = \begin{bmatrix} C & DC_{nb} \end{bmatrix} \tag{4.2.20}$$

Using a zero-order-hold model with a sample time of $T = \frac{1}{60}$ second, the equivalent discrete-time augmented system with added noise terms becomes

$$x^a_{k+1} = \Phi_a x^a_k + w^{nb}_k + \Psi_a w_k \tag{4.2.21}$$

$$z^a_k = H_a x^a_k + v_k \tag{4.2.22}$$

where

$$\Phi_a = \begin{bmatrix} -0.3271 & 0.0057 & -0.00009 & 0.00008 \\ -89.2411 & -0.4691 & -0.0129 & 0.00524 \\ 0 & 0 & 0.9786 & 0.0160 \\ 0 & 0 & -2.5275 & 0.9146 \end{bmatrix} \quad \Psi_a = \begin{bmatrix} 1 \\ 100 \\ 0 \\ 0 \end{bmatrix}$$

$$H_a = \begin{bmatrix} 63165 & 100.53 & 0 & 1 \end{bmatrix} \tag{4.2.23}$$

The first plant noise term w^{nb}_k is the discrete-time version of the continuous-time plant noise $w_{nb}(t)$ of (4.2.14), with new covariance found from [5] as

$$Q_{nb} = \int_0^T e^{A_a \tau} B_a N^2_{nb} B^{\mathrm{T}}_a \, e^{A^{\mathrm{T}}_a \tau} \, d\tau$$

or

$$Q_{nb} = \begin{bmatrix} 0.00002 & 0.0024 & 0.0047 & 0.0374 \\ 0.0024 & 0.0329 & 0.0564 & 55.85 \\ 0.0047 & 0.0564 & 1.048 & 92.22 \\ 0.0374 & 55.85 & 92.22 & 11194 \end{bmatrix} \tag{4.2.24}$$

The second plant noise term w_k and the measurement noise v_k are the same as those found in (4.2.6) and (4.2.7). The three processes w^{nb}_k, w_k, and v_k are assumed to be mutually statistically independent.

The Kalman filter based on the augmented system (4.2.21) and (4.2.22), is given by

$$\hat{x}^a_{k+1} = \Phi_a \hat{x}^a_k + K^a_{k+1}[z_{k+1} - H_a \Phi_a \hat{x}^a_k] \tag{4.2.25}$$

with covariance and gain equations

$$M^a_{k+1} = \Phi_a P^a_k \Phi^{\mathrm{T}}_a + Q_{nb} + \Psi_a Q \Psi^{\mathrm{T}}_a \tag{4.2.26}$$

$$K^a_{k+1} = M^a_{k+1} H^{\mathrm{T}}_a (H_a M^a_{k+1} H^{\mathrm{T}}_a + R)^{-1} \tag{4.2.27}$$

$$P^a_{k+1} = (I - K^a_{k+1} H_a) M^a_{k+1} \tag{4.2.28}$$

Note that in (4.2.25) the measurement z_k produced from (4.2.6) and (4.2.7) using the actual nonGaussian input signal (4.2.9) is used instead of the measurement z^a_k from the augmented system (4.2.21) and (4.2.22).

The MGS filter is initialized with equally-valued weighting terms

$$w^j_0 = \frac{1}{3}, \quad \text{for } j = 1, 2, 3 \tag{4.2.29}$$

A Markov transition probability matrix Θ^{ij}_{bS}, consisting of θ^{ij}_{bS} elements, is configured with a high probability that the bias term does not switch from one value to another, and a low probability that the bias term does switch, such as

$$\Theta^{ij}_{bS} = \begin{bmatrix} .95 & .025 & .025 \\ .025 & .95 & .025 \\ .025 & .025 & .95 \end{bmatrix} \tag{4.2.30}$$

The initial values of the state, state estimate, and error covariance for the MGS adaptive filter are

$$x_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \hat{x}_0 = \begin{bmatrix} 0.0004 \\ 0.0135 \end{bmatrix} \quad P_o = \begin{bmatrix} 10000 & 0 \\ 0 & 10000 \end{bmatrix} \tag{4.2.31}$$

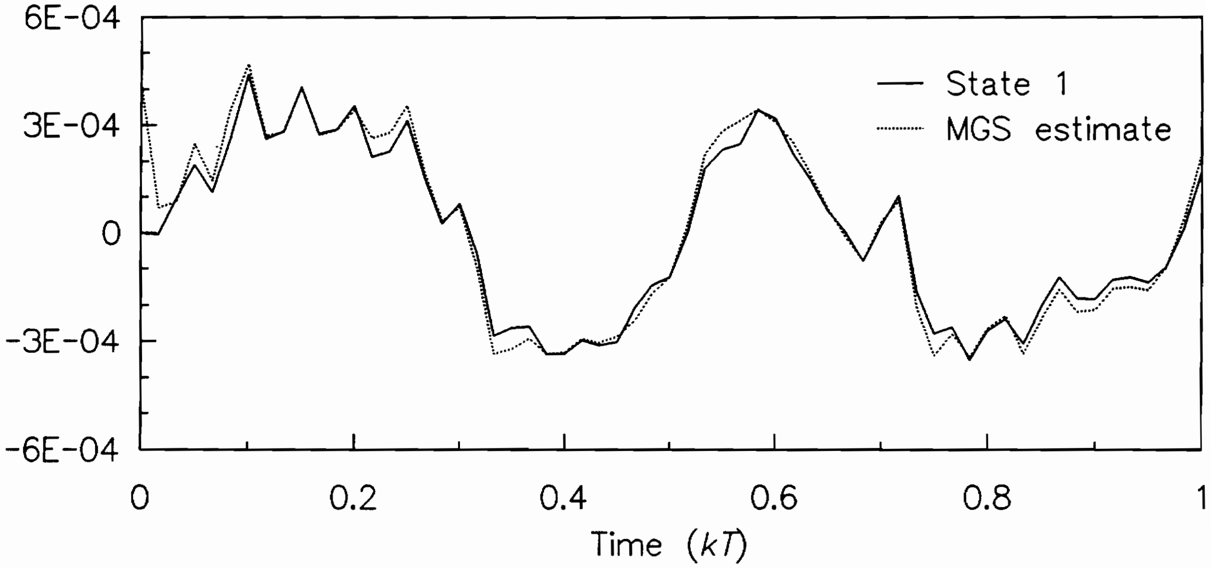
and for the augmented Kalman filter are

$$x_0^a = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad \hat{x}_0^a = \begin{bmatrix} 0.0004 \\ 0.0135 \\ 0 \\ 0 \end{bmatrix} \quad P_o^a = \begin{bmatrix} 10000 & 0 & 0 & 0 \\ 0 & 10000 & 0 & 0 \\ 0 & 0 & 10000 & 0 \\ 0 & 0 & 0 & 10000 \end{bmatrix} \tag{4.2.32}$$

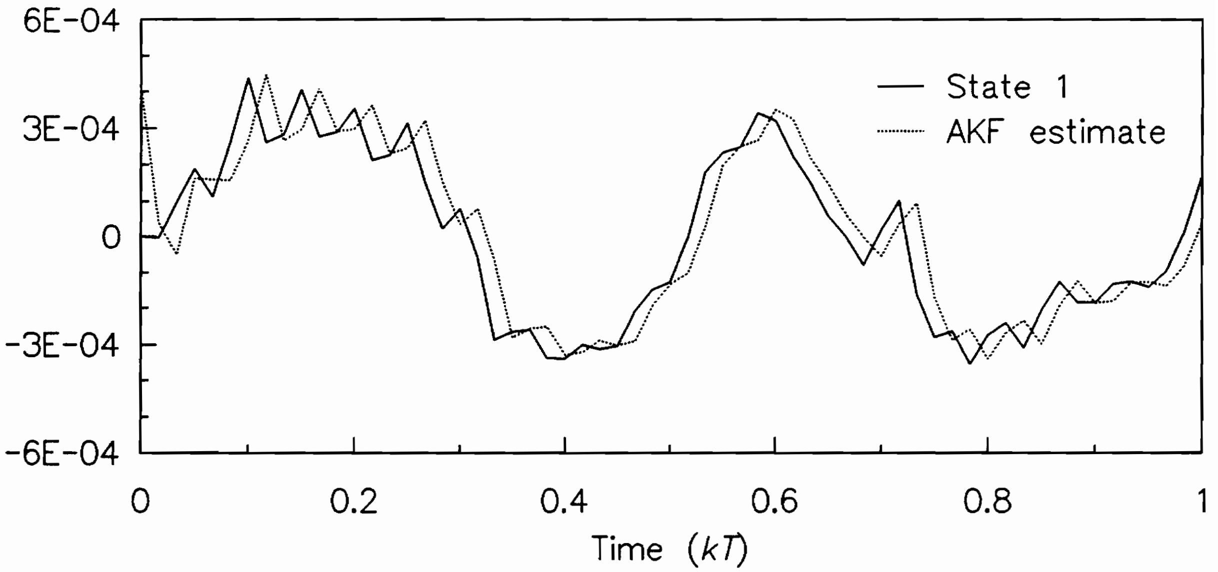
Figure 9a shows the first state x_{k+1}^1 and the overall state estimate $\hat{x}_{k+1}^1$ from the MGS adaptive filter, while Figure 9b shows the first state x_{k+1}^1 and the overall state estimate $\hat{x}_{k+1}^{a1}$ from the augmented Kalman filter. Figure 10a and 10b show similar estimation results for the second state x_{k+1}^2. The MGS adaptive filter performs very well, with its estimates tracking the states almost exactly. The augmented Kalman filter however introduces considerable delay in its estimates. Table 5 shows a normalized mean-square-error percentage measure for the state estimates produced by

$$\text{mse } \% = \frac{E[(x-\hat{x})^2]}{E[x^2]} \times 100 \tag{4.2.33}$$

The mse percentages for the augmented Kalman filter are considerably higher than the mse percentages for the MGS adaptive filter due to the delay in the augmented Kalman filter estimates. Figure 11a shows the weighting terms for each of the three b_i, S_i pairs. Figure 11b compares the three weighting terms with the stochastic FM input signal, showing how the weights switch according to the amplitude of the input signal.



(a)



(b)

Figure 9. (a) State 1 and MGS adaptive filter estimate, (b) state 1 and augmented Kalman filter estimate.

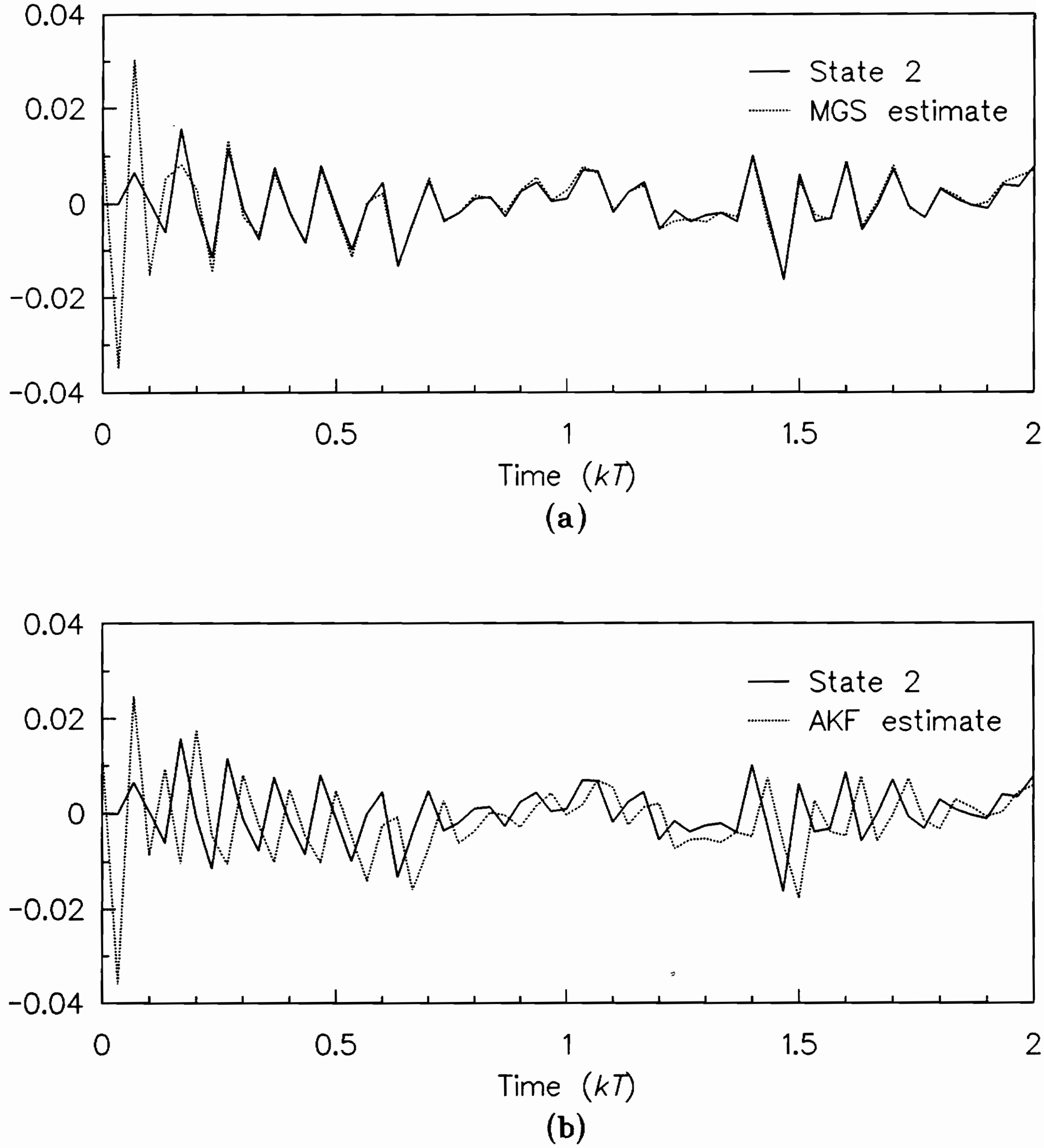


Figure 10. (a) State 2 and MGS adaptive filter estimate,
(b) state 2 and augmented Kalman filter estimate.

Table 5.

Normalized mean-square-error percentage for state estimates.

State	MGS filter	Augmented filter
1	1.85	14.76
2	3.35	227.67

$$\text{mse } \% = \frac{E[(x-\hat{x})^2]}{E[x^2]} \times 100$$

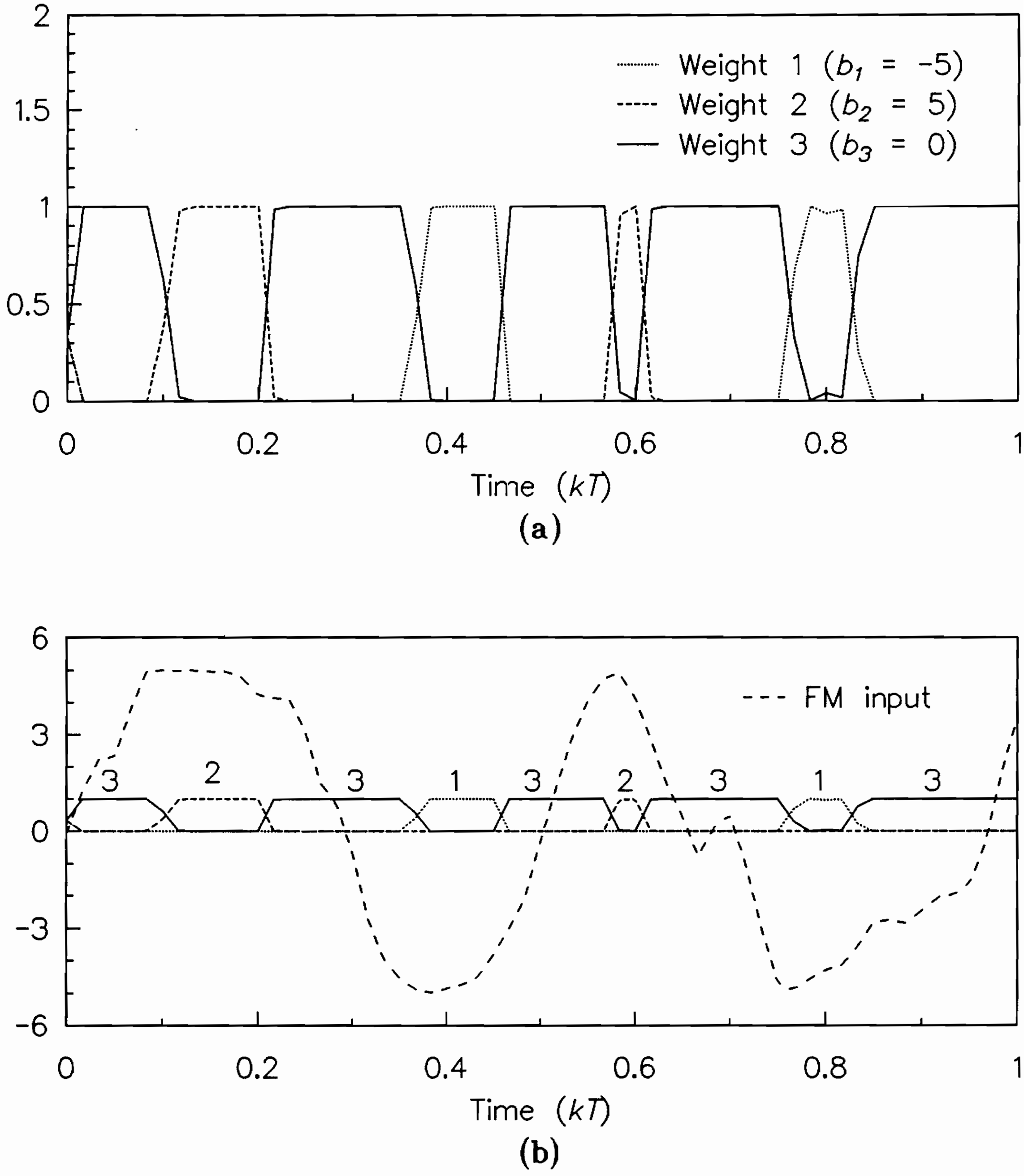


Figure 11. (a) Weighting terms, (b) weighting terms compared with stochastic FM signal.

4.3 Necessary Condition for Effective Estimation

The conditional density (4.1.21), which appears in the weighting term (4.1.27) is of prime importance in determining whether or not the MGS adaptive filter will work properly for a given system. As previously stated, the i^{th} estimator (4.1.16) based on the bias and covariance pair that most closely matches the actual bias and covariance pair of the modeled system will have a corresponding weighting term that tends closer to one, while the weights of the other mismatched estimators will tend towards zero. The bias and covariance pairs influence the weighting terms through the conditional density (4.1.21), rewritten here as the n-dimensional (z_{k+1} is an $n \times 1$ vector) Gaussian conditional density

$$p(z_{k+1} \mid b_i, S_i, Z_k) = \frac{1}{(2\pi)^{n/2} \left| R^i_{k+1} \right|^{1/2}} \, \mathrm{e}^{-\frac{1}{2} r^{i\,\mathrm{T}}_{k+1} R^{i\,-1}_{k+1} r^i_{k+1}} \tag{4.3.1}$$

where

$$R^i_{k+1} = H M^i_{k+1} H^{\mathrm{T}} + D S_i D^{\mathrm{T}} + R \tag{4.3.2}$$

is defined as the i^{th} measurement covariance from Kalman gain (4.1.18) and

$$r^i_{k+1} = z_{k+1} - D b_i - H(\Phi \hat{x}^i_k + \Gamma b_i) \tag{4.3.3}$$

is defined as the i^{th} measurement residual from Kalman estimator (4.1.16). Essentially, the measurement residual of the matched filter (with the bias and covariance pair that most closely matches the actual pair of the modeled system) will have the smallest expected value, typically zero mean. The residuals of all the other mismatched filters will be biased. Under the Gaussian assumption, the probability of the matched filter will be the largest among all the filters.

In the previous simulation examples of chapters 3 and 4, this method of calculating the conditional measurement probabilities (4.3.1) and i^{th} estimator weighting terms (4.1.26) works very well, producing very accurate state estimates. However, in the case of the force to acceleration modal system described in [1] and by (A.5) of appendix A, this method did not perform well. Upon analysis of several simulations, it appeared that the statistical steady-state value (that is, the steady-state value of the expected value) of the measurement residual of *each* i^{th} filter converged to zero. Only the correctly matched filter is supposed to have a statistical steady-state (SSS) value of zero, while the residuals of the mismatched filters are supposed to be biased. Yet, in this case, *none* of the residuals were biased. Therefore, the MGS algorithm could not detect which filter possessed the correct bias and covariance pair, and the weights of all the filters converged to the same value.

Upon investigation, it has been determined that the SSS value of the i^{th} measurement residual (4.3.3) is not only a function of the bias term b_i, but is also a function of the dc gain of the system transfer function from u_k to z_k

$$G(z)\Big|_{z=1} = G_{dc} = H(I-\Phi)^{-1}\Gamma + D \tag{4.3.4}$$

The dc gains of the systems used in the previous simulations of chapters 3 and 4 are nonzero, producing different SSS residual values due to the different bias and covariance pairs. As shown in appendix A, the dc gain of the force to acceleration modal system is zero, producing equal SSS residual values despite differences in the bias and covariance pairs.

In order to show the dependency of the SSS residual value upon the system dc gain, first rearrange the i^{th} Kalman filter equation (4.1.16) as

$$\hat{x}^i_{k+1} = (I - K^i_{k+1}H)(\Phi\hat{x}^i_k + \Gamma b_i) + K^i_{k+1}(z_{k+1} - Db_i) \tag{4.3.5}$$

and take the expected value

$$E[\hat{x}^i_{k+1}] = (I - K^i_{k+1}H)(\Phi E[\hat{x}^i_k] + \Gamma b_i) + K^i_{k+1}(E[z_{k+1}] - Db_i) \tag{4.3.6}$$

Equation (4.3.6) requires the expected value of the measurement model (4.1.2) at iteration k+1, given as

$$E[z_{k+1}] = HE[x_{k+1}] + Du_{k+1} + E[v_{k+1}] \tag{4.3.7}$$

with u_{k+1} being the actual input signal, assumed known for this development. Substituting the plant model (4.1.1) for x_{k+1} in (4.3.7) gives

$$\begin{aligned} E[z_{k+1}] &= HE[\Phi x_k + \Gamma u_k + \Psi w_k] + Du_{k+1} + E[v_{k+1}] \\ &= H\Phi E[x_k] + H\Gamma u_k + H\Psi E[w_k] + Du_{k+1} + E[v_{k+1}] \\ &= H\Phi E[x_k] + H\Gamma u_k + Du_{k+1} \end{aligned} \tag{4.3.8}$$

since w_k and v_{k+1} are each zero mean. Substituting (4.3.8) into (4.3.6) produces

$$E[\hat{x}^i_{k+1}] = (I - K^i_{k+1}H)(\Phi E[\hat{x}^i_k] + \Gamma b_i) + K^i_{k+1}(H\Phi E[x_k] + H\Gamma u_k + Du_{k+1} - Db_i) \tag{4.3.9}$$

Now, let (4.3.9) reach steady-state with $E[\hat{x}^i_{k+1}] = E[\hat{x}^i_k] = \hat{x}^i_{ss}$ and $u_{k+1} = u_k = u$.

The SSS value of (4.3.9) is then

$$\hat{x}^i_{ss} = (I - K^i_{ss}H)(\Phi\hat{x}^i_{ss} + \Gamma b_i) + K^i_{ss}(H\Phi x_{ss} + H\Gamma u + Du - Db_i) \qquad (4.3.10)$$

To find x_{ss}, the SSS value of the state model (4.1.1), first find the expected value

$$\begin{aligned} E[x_{k+1}] &= \Phi E[x_k] + \Gamma u_k + \Psi E[w_k] \\ &= \Phi E[x_k] + \Gamma u_k \end{aligned}$$

and let $E[x_{k+1}] = E[x_k] = x_{ss}$ and $u_k = u$ at steady-state

$$x_{ss} = \Phi x_{ss} + \Gamma u$$

Solving for x_{ss} gives

$$x_{ss} = (I - \Phi)^{-1}\Gamma u \qquad (4.3.11)$$

and substituting (4.3.11) into (4.3.10) produces

$$\hat{x}^i_{ss} = (I - K^i_{ss}H)(\Phi\hat{x}^i_{ss} + \Gamma b_i) + K^i_{ss}H\Phi(I - \Phi)^{-1}\Gamma u + K^i_{ss}(H\Gamma u + Du - Db_i) \qquad (4.3.12)$$

Now, the $\Phi(I - \Phi)^{-1}$ term of (4.3.12) can be rewritten as the identity

$$\Phi(I - \Phi)^{-1} = (I - \Phi)^{-1} - I \qquad (4.3.13)$$

since post-multiplying both sides of (4.3.13) by $(I - \Phi)$ shows the equality

$$\Phi(I-\Phi)^{-1}(I-\Phi) = (I-\Phi)^{-1}(I-\Phi) - (I-\Phi)$$
$$\Phi = I - I + \Phi$$
$$\Phi = \Phi$$

Substituting (4.3.13) into (4.3.12) and expanding gives

$$\begin{aligned}\hat{x}^i_{ss} = (I - K^i_{ss}H)\Phi\hat{x}^i_{ss} + (I - K^i_{ss}H)\Gamma b_i + K^i_{ss}H(I-\Phi)^{-1}\Gamma u \\ - K^i_{ss}H\Gamma u + K^i_{ss}H\Gamma u + K^i_{ss}Du - K^i_{ss}Db_i\end{aligned}$$

$$\begin{aligned}\hat{x}^i_{ss} = (I - K^i_{ss}H)\Phi\hat{x}^i_{ss} + (I - K^i_{ss}H)\Gamma b_i + K^i_{ss}H(I-\Phi)^{-1}\Gamma u \\ + K^i_{ss}Du - K^i_{ss}Db_i\end{aligned}$$

$$\hat{x}^i_{ss} = (I - K^i_{ss}H)\Phi\hat{x}^i_{ss} + (I - K^i_{ss}H)\Gamma b_i - K^i_{ss}Db_i + K^i_{ss}[H(I-\Phi)^{-1}\Gamma + D]u$$

$$[I - (I - K^i_{ss}H)\Phi]\hat{x}^i_{ss} = (\Gamma - K^i_{ss}H\Gamma - K^i_{ss}D)b_i + K^i_{ss}[H(I-\Phi)^{-1}\Gamma + D]u$$

or

$$\hat{x}^i_{ss} = [I - (I - K^i_{ss}H)\Phi]^{-1}\{[\Gamma - K^i_{ss}(H\Gamma + D)]b_i + K^i_{ss}[H(I-\Phi)^{-1}\Gamma + D]u\} \quad (4.3.14)$$

With x_{ss} (4.3.11) and $\hat{x}^i_{ss}$ (4.3.14) known, the SSS value of the i^{th} measurement residual (4.3.3) can now be found. Taking the expected value of (4.3.3) and using (4.3.8) produces

$$E[r^i_{k+1}] = H\Phi E[x_k] + H\Gamma u_k + Du_{k+1} - Db_i - H\Phi E[\hat{x}^i_k] - H\Gamma b_i \qquad (4.3.15)$$

Letting (4.3.15) reach steady-state, with $E[r^i_{k+1}] = r^i_{ss}$ and $u_{k+1} = u_k = u$, the result is

$$r^i_{ss} = H\Phi x_{ss} + H\Gamma u + Du - Db_i - H\Phi \hat{x}^i_{ss} - H\Gamma b_i \tag{4.3.16}$$

Using (4.3.11) and (4.3.13), $H\Phi x_{ss}$ is rewritten as

$$H\Phi x_{ss} = H\Phi(I-\Phi)^{-1}\Gamma u = H(I-\Phi)^{-1}\Gamma u - H\Gamma u \tag{4.3.17}$$

Substituting (4.3.17) into (4.3.16) produces a cancellation of terms

$$\begin{aligned} r^i_{ss} &= H(I-\Phi)^{-1}\Gamma u - H\Gamma u + H\Gamma u + Du - Db_i - H\Phi\hat{x}^i_{ss} - H\Gamma b_i \\ &= H(I-\Phi)^{-1}\Gamma u + Du - Db_i - H\Phi\hat{x}^i_{ss} - H\Gamma b_i \\ &= [H(I-\Phi)^{-1}\Gamma + D]u - (H\Gamma + D)b_i - H\Phi\hat{x}^i_{ss} \end{aligned} \tag{4.3.18}$$

Using (4.3.14) in (4.3.18) gives

$$\begin{aligned} r^i_{ss} &= [H(I-\Phi)^{-1}\Gamma + D]u - (H\Gamma + D)b_i \\ &\quad - H\Phi[I-(I-K^i_{ss}H)\Phi]^{-1}\{[\Gamma - K^i_{ss}(H\Gamma+D)]b_i + K^i_{ss}[H(I-\Phi)^{-1}\Gamma + D]u\} \end{aligned} \tag{4.3.19}$$

The dc gain G_{dc} (4.3.4) appears in two places in (4.3.19) rewritten as

$$r^i_{ss} = G_{dc}u - (H\Gamma + D)b_i - H\Phi[I-(I-K^i_{ss}H)\Phi]^{-1}\{[\Gamma - K^i_{ss}(H\Gamma+D)]b_i + K^i_{ss}G_{dc}u\} \tag{4.3.20}$$

However, (4.3.20) can be simplified further using additional algebraic manipulation.

The $H\Gamma + D$ term can be rewritten by combining the dc gain (4.3.4) and the identity (4.3.13) as

$$\begin{aligned} G_{dc} &= H[\Phi(I-\Phi)^{-1} + I]\Gamma + D \\ &= H\Phi(I-\Phi)^{-1}\Gamma + H\Gamma + D \end{aligned}$$

so that

$$H\Gamma + D = G_{dc} - H\Phi(I-\Phi)^{-1}\Gamma \tag{4.3.21}$$

Substituting (4.3.21) into (4.3.20) and rearranging terms gives

$$\begin{aligned} r^i_{ss} = {} & G_{dc}u - (G_{dc} - H\Phi(I-\Phi)^{-1}\Gamma)b_i \\ & - H\Phi[I-(I-K^i_{ss}H)\Phi]^{-1}\{[\Gamma - K^i_{ss}(G_{dc} - H\Phi(I-\Phi)^{-1}\Gamma)]b_i + K^i_{ss}G_{dc}u\} \end{aligned}$$

$$\begin{aligned} r^i_{ss} = {} & G_{dc}u - G_{dc}b_i + H\Phi(I-\Phi)^{-1}\Gamma b_i \\ & - H\Phi[I-(I-K^i_{ss}H)\Phi]^{-1}\{\Gamma b_i - K^i_{ss}G_{dc}b_i + K^i_{ss}H\Phi(I-\Phi)^{-1}\Gamma b_i + K^i_{ss}G_{dc}u\} \end{aligned}$$

$$\begin{aligned} r^i_{ss} = {} & G_{dc}(u-b_i) + H\Phi(I-\Phi)^{-1}\Gamma b_i \\ & - H\Phi[I-(I-K^i_{ss}H)\Phi]^{-1}\{[I + K^i_{ss}H\Phi(I-\Phi)^{-1}]\Gamma b_i + K^i_{ss}G_{dc}(u-b_i)\} \end{aligned} \tag{4.3.22}$$

Next, the $[I-(I-K^i_{ss}H)\Phi]^{-1}$ term can be rearranged as

$$[I-(I-K^i_{ss}H)\Phi]^{-1} = [(I-\Phi) + K^i_{ss}H\Phi]^{-1} \tag{4.3.23}$$

and factoring out an $(I-\Phi)$ term produces

$$[(I-\Phi)+K_{ss}^{i}H\Phi]^{-1}=[(I+K_{ss}^{i}H\Phi(I-\Phi)^{-1})(I-\Phi)]^{-1} \qquad (4.3.24)$$

Using the identity

$$(AB)^{-1}=B^{-1}A^{-1}$$

if A and B are nonsingular square matrices, setting $A=I+K_{ss}^{i}H\Phi(I-\Phi)^{-1}$ and $B=(I-\Phi)$, (4.3.24) is equal to

$$[(I+K_{ss}^{i}H\Phi(I-\Phi)^{-1})(I-\Phi)]^{-1}=(I-\Phi)^{-1}(I+K_{ss}^{i}H\Phi(I-\Phi)^{-1})^{-1} \qquad (4.3.25)$$

Combining (4.3.23) - (4.3.25) therefore results in

$$[I-(I-K_{ss}^{i}H)\Phi]^{-1}=(I-\Phi)^{-1}(I+K_{ss}^{i}H\Phi(I-\Phi)^{-1})^{-1} \qquad (4.3.26)$$

Substituting (4.3.26) into (4.3.22) gives

$$\begin{aligned} r_{ss}^{i} &= G_{dc}(u-b_i)+H\Phi(I-\Phi)^{-1}\Gamma b_i \\ &\quad -H\Phi(I-\Phi)^{-1}(I+K_{ss}^{i}H\Phi(I-\Phi)^{-1})^{-1}\{[I+K_{ss}^{i}H\Phi(I-\Phi)^{-1}]\Gamma b_i+K_{ss}^{i}G_{dc}(u-b_i)\} \end{aligned}$$

and multiplying through gives a cancellation of terms

$$\begin{aligned} r_{ss}^{i} &= G_{dc}(u-b_i)+H\Phi(I-\Phi)^{-1}\Gamma b_i-H\Phi(I-\Phi)^{-1}\Gamma b_i \\ &\qquad -H\Phi(I-\Phi)^{-1}(I+K_{ss}^{i}H\Phi(I-\Phi)^{-1})^{-1}K_{ss}^{i}G_{dc}(u-b_i) \\ &= G_{dc}(u-b_i)-H\Phi(I-\Phi)^{-1}(I+K_{ss}^{i}H\Phi(I-\Phi)^{-1})^{-1}K_{ss}^{i}G_{dc}(u-b_i) \\ &= [I-H\Phi(I-\Phi)^{-1}(I+K_{ss}^{i}H\Phi(I-\Phi)^{-1})^{-1}K_{ss}^{i}]G_{dc}(u-b_i) \qquad (4.3.27) \end{aligned}$$

Substituting (4.3.26) into (4.3.27), r^i_{ss} becomes finally

$$r^i_{ss} = [I - H\Phi[I - (I - K^i_{ss}H)\Phi]^{-1}K^i_{ss}]G_{dc}(u - b_i) \tag{4.3.28}$$

The i^{th} SSS residual value (4.3.28) is clearly a function of the dc gain G_{dc} as well as the bias term b_i. It also is a function of the covariance term S_i through the steady state Kalman gain K^i_{ss}. As a result, a nonzero dc system gain becomes a necessary condition for effective estimation.

If G_{dc} is nonzero, then r^i_{ss} will be different for each bias and covariance pair. The filter with the matched pair that causes r^i_{ss} to be the smallest will also cause r^i_{k+1} (4.3.3) to have the smallest expected value of all the filters. The probability (4.3.1) of this matched filter will then be the largest among all the filters, and will produce the largest weighting term (4.1.26). If G_{dc} is zero, then r^i_{ss} is also zero for all the filters. The differences between the bias and covariance pairs are masked, causing the probabilities (4.3.1), and therefore the weighting terms (4.1.26), of all the filters to converge to the same value. At this point, the MGS algorithm becomes an ineffective estimation scheme. However, the alternate estimation procedures described in the next section may provide a solution to this problem.

4.4 Alternate Estimation Procedures

A possible solution to the problem of zero dc system gain is to determine if an alternate measurement provides a nonzero dc gain. For example, if in a particular zero dc gain system acceleration measurements are taken, changing to velocity measurements may provide a nonzero dc gain. If the actual sensors producing these measurements cannot be changed, integrating the acceleration measurement data to produce *approximate*-velocity measurement data may allow the MGS algorithm to work effectively.

As shown in [1] and by (A.5) of appendix A, the force to acceleration modal system produces acceleration measurement data at sensor grid points and has a zero dc gain. This system is termed the acceleration modal system, or AMS. Changing this AMS to produce velocity measurement data at sensor grid points results in a new transfer function with a nonzero dc gain. This new system is termed the velocity modal system, or VMS. Typically, changing the actual system cannot be done in practice. However, given a large enough signal to noise ratio between the input signal u_k and measurement noise v_k, a good approximation of the velocity measurement data can be generated by integrating the acceleration measurement data from the AMS using a discrete-time integrator. Redesigning the i^{th} measurement covariance (4.3.2) and residual (4.3.3) to be based on the VMS while actually using the integrated acceleration (*approximate*-velocity) measurement data from the AMS produces proper weighting terms (4.1.26) that allow the MGS algorithm to work effectively.

In order to illustrate this alternate estimation procedure, the system and filter equations for the VMS are developed below. Using a zero-order-hold model for the continuous-time force input to velocity output system described by (A.6) of appendix A, the resulting discrete-time system with added noise terms is

$$x_{k+1} = \Phi x_k + \Gamma u_k + \Psi w_k \tag{4.4.1}$$

$$y_k = H_v x_k + v_k^v \tag{4.4.2}$$

$$u_k = b_k + n_k \tag{4.4.3}$$

where x_{k+1} is the state vector

y_k is the velocity measurement vector

w_k is a zero mean white Gaussian plant noise process with covariance Q

v_k^v is a zero mean white Gaussian measurement noise process with covariance R_v

Φ, Γ, Ψ, H_v are the respective constant transition matrices

u_k is the vector GS signal model of the actual nonGaussian input signal, comprised of semi-Markov bias vector b_k, and zero mean white Gaussian noise n_k with semi-Markov covariance S_k

and the random quantities x_k, w_k, v_k^v, b_k, and n_k are assumed to be mutually statistically independent. The state model (4.4.1) is exactly the same as the model used previously (4.1.1), so that the states of the new VMS are exactly the same as the states of the previous AMS. The Gaussian sum signal model (4.4.3) is also the same as the model used previously (4.1.3). The measurement model (4.4.2) is different than (4.1.2) because y_k is a velocity measurement while z_k is an acceleration measurement. Other differences in (4.4.2) include a new transition matrix H_v and a new measurement noise process v_k^v with new covariance R_v.

Following the same procedures in section 4.1, the MGS algorithm is applied to the system of (4.4.1) - (4.4.3), producing the overall state estimate

$$\hat{x}_{k+1}^v = \sum_{i=1}^{M} \hat{x}_{k+1}^{vi} p(b_i, S_i \mid Y_{k+1}) \tag{4.4.4}$$

where Y_{k+1} is the current velocity measurement sequence $\{y_1, y_2, \ldots, y_{k+1}\}$, and $\hat{x}_{k+1}^{vi}$

is given by the Kalman filter equation

$$\hat{x}^{vi}_{k+1} = \Phi\hat{x}^{vi}_k + \Gamma b_i + K^{vi}_{k+1}[y_{k+1} - H_v(\Phi\hat{x}^{vi}_k + \Gamma b_i)] \tag{4.4.5}$$

with covariance and gain equations

$$M^{vi}_{k+1} = \Phi P^{vi}_k \Phi^{\mathrm{T}} + \Gamma S_i \Gamma^{\mathrm{T}} + \Psi Q \Psi^{\mathrm{T}} \tag{4.4.6}$$

$$K^{vi}_{k+1} = M^{vi}_{k+1} H^{\mathrm{T}}_v (H_v M^{vi}_{k+1} H^{\mathrm{T}}_v + R_v)^{-1} \tag{4.4.7}$$

$$P^{vi}_{k+1} = (I - K^{vi}_{k+1} H_v) M^{vi}_{k+1} \tag{4.4.8}$$

and weighting terms

$$w^{vi}_{k+1} = p(b_i, S_i \mid Y_{k+1}) = C^v_{k+1} p(y_{k+1} \mid b_i, S_i, Y_k) \sum_{j=1}^{M} \theta^{ij}_{bS} w^{vj}_k \tag{4.4.9}$$

with

$$p(y_{k+1} \mid b_i, S_i, Y_k) = \mathrm{N}[H_v(\Phi\hat{x}^{vi}_k + \Gamma b_i),\ H_v M^{vi}_{k+1} H^{\mathrm{T}}_v + R_v] \tag{4.4.10}$$

$$\theta^{ij}_{bS} = p(b_{k+1} = b_i, S_{k+1} = S_i \mid b_k = b_j, S_k = S_j) \tag{4.4.11}$$

$$w^{vj}_k = p(b_k = b_j, S_k = S_j \mid Y_k) \tag{4.4.12}$$

where C^v_{k+1} is the scale factor determined at each time interval such that the sum of all the weighting terms (4.4.9) is equal to one.

The conditional density (4.4.10) of the VMS is a function of the i^{th} velocity measurement covariance

$$R^{vi}_{k+1} = H_v M^{vi}_{k+1} H^{\mathrm{T}}_v + R_v \tag{4.4.13}$$

and the i^{th} velocity measurement residual

$$r_{k+1}^{vi} = y_{k+1} - H_v(\Phi \hat{x}_k^{vi} + \Gamma b_i) \tag{4.4.14}$$

which requires velocity measurement data y_{k+1}. This data is not available, since the actual system being simulated is the AMS which produces acceleration measurement data z_{k+1}. In order to generate the proper probabilities (4.4.10) needed by the weighting terms (4.4.9) for effective MGS estimation, *approximate*-velocity measurement data $\widetilde{y}_{k+1}$, generated from the integral of z_{k+1} of the AMS, is used in place of y_{k+1} in (4.4.14) and (4.4.5). This requires a new Kalman filter

$$\hat{x}_{k+1}^{\widetilde{v}i} = \Phi \hat{x}_k^{\widetilde{v}i} + \Gamma b_i + K_{k+1}^{vi}[\widetilde{y}_{k+1} - H_v(\Phi \hat{x}_k^{\widetilde{v}i} + \Gamma b_i)] \tag{4.4.15}$$

and new weighting terms

$$w_{k+1}^{\widetilde{v}i} = p(b_i, S_i \mid \widetilde{Y}_{k+1}) = C_{k+1}^{\widetilde{v}} p(\widetilde{y}_{k+1} \mid b_i, S_i, \widetilde{Y}_k) \sum_{j=1}^{M} \theta_{bS}^{ij} w_k^{\widetilde{v}j} \tag{4.4.16}$$

requiring

$$p(\widetilde{y}_{k+1} \mid b_i, S_i, \widetilde{Y}_k) = \mathrm{N}[H_v(\Phi \hat{x}_k^{\widetilde{v}i} + \Gamma b_i),\ H_v M_{k+1}^{vi} H_v^{\mathrm{T}} + R_v] \tag{4.4.17}$$

$$w_k^{\widetilde{v}j} = p(b_k = b_j, S_k = S_j \mid \widetilde{Y}_k) \tag{4.4.18}$$

The Kalman gain and covariance equations (4.4.6) - (4.4.8) and Markov probabilities (4.4.11) can be used without modification since they are not dependent on the *approximate*-velocity measurement data $\widetilde{y}_{k+1}$.

The overall estimate can be calculated by three different methods. The first method produces

$$\hat{x}^{\tilde{a}}_{k+1} = \sum_{i=1}^{M} \hat{x}^{i}_{k+1} p(b_i, S_i \mid \tilde{Y}_{k+1}) \tag{4.4.19}$$

where $\hat{x}^{i}_{k+1}$ is the i^{th} state estimate (4.1.16) from the AMS. There was never a problem in calculating these AMS state estimates. It was the problem in calculating the weighting terms that necessitated an alternate estimation scheme. However, (4.4.19) requires two *separate* parallel banks of Kalman filters: one using (4.1.16) to calculate $\hat{x}^{i}_{k+1}$, and a second using (4.4.15) to calculate $w^{\tilde{v}i}_{k+1} = p(b_i, S_i \mid \tilde{Y}_{k+1})$.

The second method produces

$$\hat{x}^{\tilde{v}}_{k+1} = \sum_{i=1}^{M} \hat{x}^{\tilde{v}i}_{k+1} p(b_i, S_i \mid \tilde{Y}_{k+1}) \tag{4.4.20}$$

where $\hat{x}^{\tilde{v}i}_{k+1}$ is the i^{th} state estimate (4.4.15) using *approximate*-velocity measurement data with the VMS. This overall state estimate (4.4.20) requires only one bank of parallel Kalman filters using (4.4.15) to calculate both $\hat{x}^{\tilde{v}i}_{k+1}$ and $w^{\tilde{v}i}_{k+1} = p(b_i, S_i \mid \tilde{Y}_{k+1})$. This provides good computational savings over calculating (4.4.19). However, (4.4.19) uses the actual AMS measurement data z_k to calculate $\hat{x}^{i}_{k+1}$, while (4.4.20) never uses z_k in its direct form.

The third method uses the same state equation (4.4.1) and GS signal model (4.4.3) as the AMS and VMS, but requires the use of a second new measurement formed by combining z_k (4.1.2) and y_k (4.4.2) into

$$\zeta_k = \begin{bmatrix} z_k \\ y_k \end{bmatrix} = H_\zeta x_k + D_\zeta u_k + v^\zeta_k \tag{4.4.21}$$

where

$$H_\zeta = \begin{bmatrix} H \\ H_v \end{bmatrix} \quad D_\zeta = \begin{bmatrix} D \\ 0 \end{bmatrix} \quad v_k^\zeta = \begin{bmatrix} v_k \\ v_k^v \end{bmatrix} \tag{4.4.22}$$

Again, velocity measurement data y_k is not available, so *approximate*-velocity measurement data $\widetilde{y}_k$ is used instead producing

$$\widetilde{\zeta}_k = \begin{bmatrix} z_k \\ \widetilde{y}_k \end{bmatrix} \tag{4.4.23}$$

The resulting Kalman filter equation is

$$\hat{x}_{k+1}^{\widetilde{\zeta}_i} = \Phi\hat{x}_k^{\widetilde{\zeta}_i} + \Gamma b_i + K_{k+1}^{\zeta_i}[\widetilde{\zeta}_{k+1} - D_\zeta b_i - H_\zeta(\Phi\hat{x}_k^{\widetilde{\zeta}_i} + \Gamma b_i)] \tag{4.4.24}$$

with covariance and gain equations (independent of $\widetilde{y}_k$)

$$M_{k+1}^{\zeta_i} = \Phi P_k^{\zeta_i}\Phi^{\mathrm{T}} + \Gamma S_i \Gamma^{\mathrm{T}} + \Psi Q \Psi^{\mathrm{T}} \tag{4.4.25}$$

$$K_{k+1}^{\zeta_i} = M_{k+1}^{\zeta_i} H_\zeta^{\mathrm{T}} (H_\zeta M_{k+1}^{\zeta_i} H_\zeta^{\mathrm{T}} + D_\zeta S_i D_\zeta^{\mathrm{T}} + R_\zeta)^{-1} \tag{4.4.26}$$

$$P_{k+1}^{\zeta_i} = (I - K_{k+1}^{\zeta_i} H_\zeta) M_{k+1}^{\zeta_i} \tag{4.4.27}$$

where the covariance of v_k^ζ is given by

$$R_\zeta = \begin{bmatrix} R & 0 \\ 0 & R_v \end{bmatrix} \tag{4.4.28}$$

and the resulting weighting terms are

$$w_{k+1}^{\widetilde{\zeta}_i} = p(b_i, S_i \mid \widetilde{\mathfrak{Z}}_{k+1}) = C_{k+1}^{\widetilde{\zeta}} p(\widetilde{\zeta}_{k+1} \mid b_i, S_i, \widetilde{\mathfrak{Z}}_k) \sum_{j=1}^{M} \theta_{bS}^{ij} w_k^{\widetilde{\zeta}_j} \tag{4.4.29}$$

where $\tilde{\mathfrak{Z}}_{k+1}$ is the current measurement sequence $\{\tilde{\zeta}_1, \tilde{\zeta}_2, \ldots, \tilde{\zeta}_k\}$, θ_{bS}^{ij} is given by (4.4.11), and

$$p(\tilde{\zeta}_{k+1} \mid b_i, S_i, \tilde{\mathfrak{Z}}_k) = \mathrm{N}[H_\zeta(\Phi \hat{x}_k^{\tilde{\zeta}i} + \Gamma b_i) + D_\zeta b_i,\ H_\zeta M_{k+1}^{\zeta i} H_\zeta^{\mathrm{T}} + D_\zeta S_i D_\zeta^{\mathrm{T}} + R_\zeta] \quad (4.4.30)$$

$$w_k^{\tilde{\zeta}j} = p(b_k = b_j, S_k = S_j \mid \tilde{\mathfrak{Z}}_k) \quad (4.4.31)$$

Forming measurement data $\tilde{\zeta}_k$ (4.4.23) allows both the state estimate (4.4.24) and weighting term (4.4.29) to be influenced by both z_k and $\tilde{y}_k$ at the same time, while requiring only one parallel bank of Kalman filters. The overall estimate produced by using this third method is

$$\hat{x}_{k+1}^{\tilde{\zeta}} = \sum_{i=1}^{M} \hat{x}_{k+1}^{\tilde{\zeta}i} p(b_i, S_i \mid \tilde{\mathfrak{Z}}_{k+1}) \quad (4.4.32)$$

These alternate estimation procedures will be compared and evaluated in the next section.

4.5 Comparison of the Alternate Estimation Procedures

The alternate estimation procedures described in the previous section are compared and evaluated using a simulation example. The overall state estimate $\hat{x}_{k+1}^{v}$ (4.4.4) based on the VMS using true velocity measurement data y_k will also be calculated as a benchmark for comparison to the three alternate overall state estimates. For convenience, Table 6 provides a summary of the overall state estimates used in the comparison. A force input to acceleration output modal system consisting of three modes serves as the simulation example. The modes are calculated using (A.7) - (A.14) of appendix A. The parameter values of α_i, $\angle c_i$, x, and f_i used in the three mode, six state model are given in Table 7.

Table 6.

Overall state estimates to be compared.

a.	$\hat{x}^{v}_{k+1} = \sum_{i=1}^{M} \hat{x}^{vi}_{k+1} p(b_i, S_i \mid Y_{k+1})$	Benchmark estimate. i^{th} state estimate and weighting term based on VMS using true velocity measurement data.
b.	$\hat{x}^{\widetilde{a}}_{k+1} = \sum_{i=1}^{M} \hat{x}^{i}_{k+1} p(b_i, S_i \mid \widetilde{Y}_{k+1})$	i^{th} state estimate based on AMS. Weighting term based on VMS using *approximate*-velocity measurement data.
c.	$\hat{x}^{\widetilde{v}}_{k+1} = \sum_{i=1}^{M} \hat{x}^{\widetilde{v}i}_{k+1} p(b_i, S_i \mid \widetilde{Y}_{k+1})$	i^{th} state estimate and weighting term based on VMS using *approximate*-velocity measurement data.
d.	$\hat{x}^{\widetilde{\zeta}}_{k+1} = \sum_{i=1}^{M} \hat{x}^{\widetilde{\zeta}i}_{k+1} p(b_i, S_i \mid \widetilde{\mathfrak{Z}}_{k+1})$	i^{th} state estimate and weighting term based on combined AMS and VMS using acceleration measurement data and *approximate*-velocity measurement data.

Table 7.

Parameter values for the three mode model.

i	α_i	$\angle c_i$ (rad)	x (m)	f_i (Hz)
1	1.446813	5.686	17.5	95.63
2	4.934421	0.705	35.0	67.47
3	13.721530	1.598	52.5	28.02

The system transition matrices are found from appendix A equations (A.5) for the AMS and (A.6) for the VMS. Using a zero-order-hold model, a sampling time of $T = 0.001$ seconds, and the values in Table 7, the discrete-time matrices for the AMS are

$$\Phi = \begin{bmatrix} 0.8250 & 0.0009 & 0 & 0 & 0 & 0 \\ -339.2108 & 0.8223 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.9118 & 0.0010 & 0 & 0 \\ 0 & 0 & -173.5593 & 0.9022 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.9846 & 0.0010 \\ 0 & 0 & 0 & 0 & -30.6059 & 0.9577 \end{bmatrix} \tag{4.5.1}$$

$$\Gamma = \begin{bmatrix} 0.0069 & 0.0033 & -0.0015 \\ -2.7667 & -4.5899 & -4.9021 \\ 0.0014 & -0.0040 & -0.0071 \\ -3.0180 & -2.5293 & -0.5295 \\ -0.0059 & 0.0023 & 0.0038 \\ -0.0314 & 0.9420 & -0.8696 \end{bmatrix} \tag{4.5.2}$$

$$\Psi = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{4.5.3}$$

$$H = \begin{bmatrix} -2807.66 & 16.1818 & -5625.67 & 6.0323 & -729.65 & -11.5741 \\ -7693.42 & 10.9286 & -5815.17 & -5.2798 & 2099.60 & 2.6883 \\ -10046.6 & 2.0778 & -2530.33 & -13.4374 & -1278.69 & 9.0026 \end{bmatrix} \tag{4.5.4}$$

$$D = \begin{bmatrix} 110.7665 & -45.3759 & -166.8950 \\ -45.3759 & -166.8950 & -41.4010 \\ -166.8950 & -41.4010 & -55.5049 \end{bmatrix} \tag{4.5.5}$$

The H_v transition matrix needed for the VMS is

$$H_v = \begin{bmatrix} 16.2043 & 0.0078 & 6.3411 & 0.0313 & -10.9321 & 0.0234 \\ 10.9903 & 0.0213 & -4.9605 & 0.0324 & 0.8409 & -0.0673 \\ 2.1583 & 0.0278 & -13.2985 & 0.0141 & 10.1277 & 0.0410 \end{bmatrix} \tag{4.5.6}$$

The measurement model of the combined AMS and VMS requires H_ζ and D_ζ matrices formed using (4.4.22). The plant noise w_k of the state model is zero-mean white Gaussian with covariance

$$Q = \begin{bmatrix} 1.4142 & 0 & 0 \\ 0 & 0.2828 & 0 \\ 0 & 0 & 0.0344 \end{bmatrix} \tag{4.5.7}$$

and is uncorrelated with AMS measurement noise v_k and VMS measurement noise v_k^v. Both are zero-mean white Gaussian with respective covariances

$$R = \begin{bmatrix} 26.36 & 0 & 0 \\ 0 & 19.77 & 0 \\ 0 & 0 & 20.41 \end{bmatrix} \quad R_v = \begin{bmatrix} 0.0264 & 0 & 0 \\ 0 & 0.0198 & 0 \\ 0 & 0 & 0.0204 \end{bmatrix} \tag{4.5.8}$$

The measurement noise covariance R_ζ of the combined AMS and VMS is given by (4.4.28).

The nonGaussian input signal is the stochastic FM signal generated by

$$u(t) = A_u \sin(2\pi f_u t + k_u \int_0^t m(\tau)\, d\tau) \tag{4.5.9}$$

where the amplitude is $A_u = 5$, the carrier frequency is $f_u = 10$ Hz, the modulation index is $k_u = 100$, and $m(t)$ is zero-mean white Gaussian noise with variance 1. Figure 12 shows the FM signal for 0.2 second. The probability density function of (4.5.9) has the same form as (4.2.10).

The nonGaussian input signal (4.5.9) will be modeled in discrete-time using the Gaussian sum (GS) signal model described in section 4.1

$$u_k = b_k + n_k \tag{4.5.10}$$

with GS density approximation (4.2.12). For simplicity, the nonGaussian input signal is applied only to mode 1.

Three filters are used in each MGS adaptive filter of Table 6. Using the values of Table 4, the b_i and S_i parameters for the three filters are

$$b_1 = \begin{bmatrix} -5 \\ 0 \\ 0 \end{bmatrix} \quad b_2 = \begin{bmatrix} 5 \\ 0 \\ 0 \end{bmatrix} \quad b_3 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \tag{4.5.11}$$

$$S_1 = S_2 = \begin{bmatrix} 0.04 & 0 & 0 \\ 0 & 0.04 & 0 \\ 0 & 0 & 0.04 \end{bmatrix} \quad S_3 = \begin{bmatrix} 4.28 & 0 & 0 \\ 0 & 4.28 & 0 \\ 0 & 0 & 4.28 \end{bmatrix} \tag{4.5.12}$$

Three of the MGS adaptive filters of Table 6 require *approximate*-velocity measurements. A discrete-time integrator implementing the trapezoidal rule is used, given by [59]

$$\tilde{y}_{k+1} = \tilde{y}_k + \frac{T}{2}(z_{k+1} + z_k) \tag{4.5.13}$$

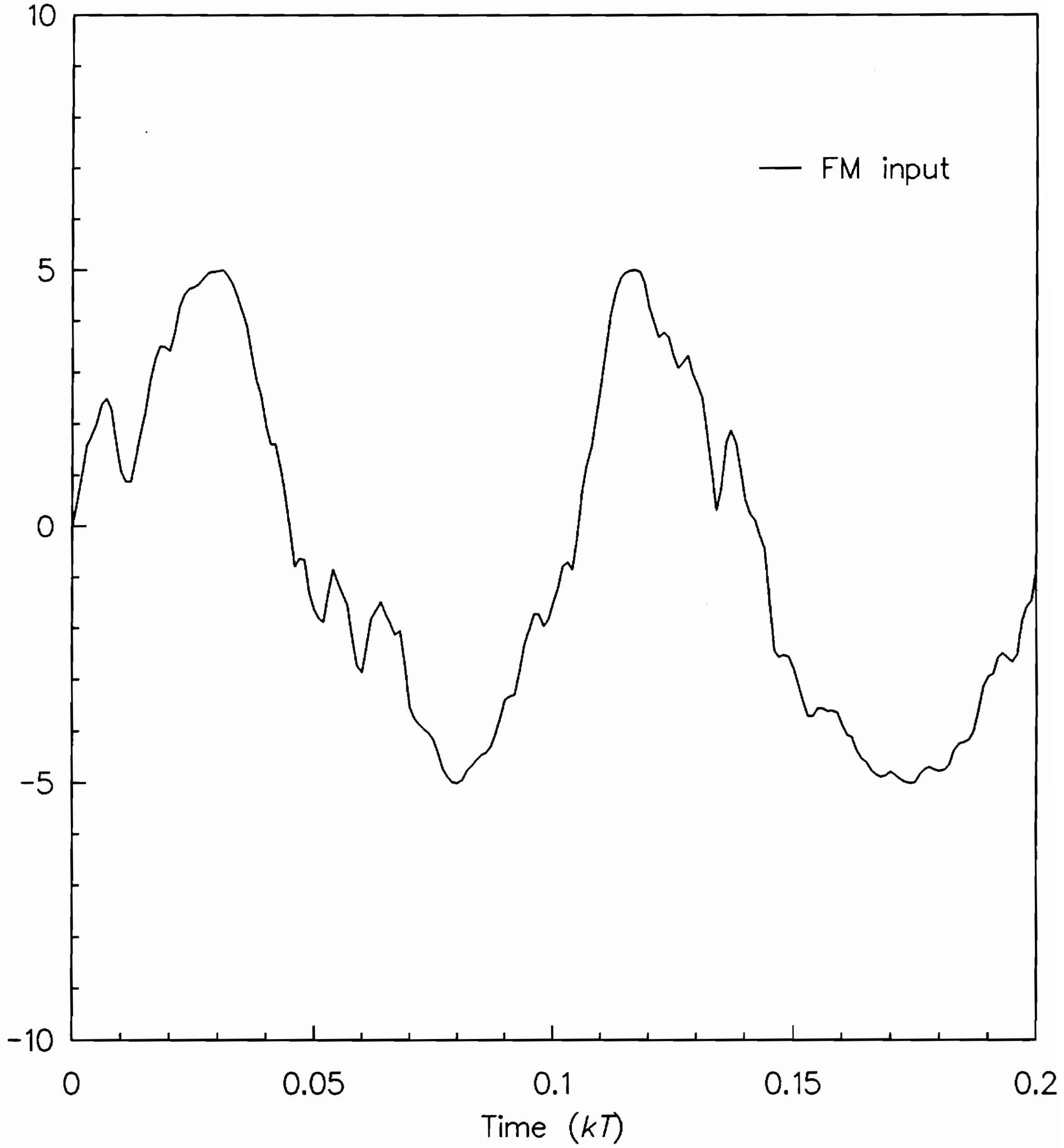


Figure 12. Stochastic FM input signal.

Figure 13a shows the mode 1 acceleration measurement z_k, while Figure 13b shows both the mode 1 actual velocity measurement y_k and *approximate*-velocity measurement $\tilde{y}_k$. The two velocity measurements track fairly closely under low noise conditions.

Each MGS adaptive filter is initialized with equally-valued weighting terms

$$w_0^j = \frac{1}{3}, \quad \text{for } j = 1,2,3 \tag{4.5.14}$$

A Markov transition probability matrix Θ_{bS}^{ij}, consisting of θ_{bS}^{ij} elements, is configured with a high probability that the bias term does not switch from one value to another, and a low probability that the bias term does switch, such as

$$\Theta_{bS}^{ij} = \begin{bmatrix} .85 & .075 & .075 \\ .075 & .85 & .075 \\ .075 & .075 & .85 \end{bmatrix} \tag{4.5.15}$$

The initial values of the state, state estimate, and error covariance for each MGS adaptive filter are

$$x_0 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad \hat{x}_0 = \begin{bmatrix} 0.2 \\ 152 \\ 0.2 \\ 62 \\ 0.2 \\ 37 \end{bmatrix} \quad P_o = \begin{bmatrix} 10000 & 0 & 0 & 0 & 0 & 0 \\ 0 & 10000 & 0 & 0 & 0 & 0 \\ 0 & 0 & 10000 & 0 & 0 & 0 \\ 0 & 0 & 0 & 10000 & 0 & 0 \\ 0 & 0 & 0 & 0 & 10000 & 0 \\ 0 & 0 & 0 & 0 & 0 & 10000 \end{bmatrix} \tag{4.5.16}$$

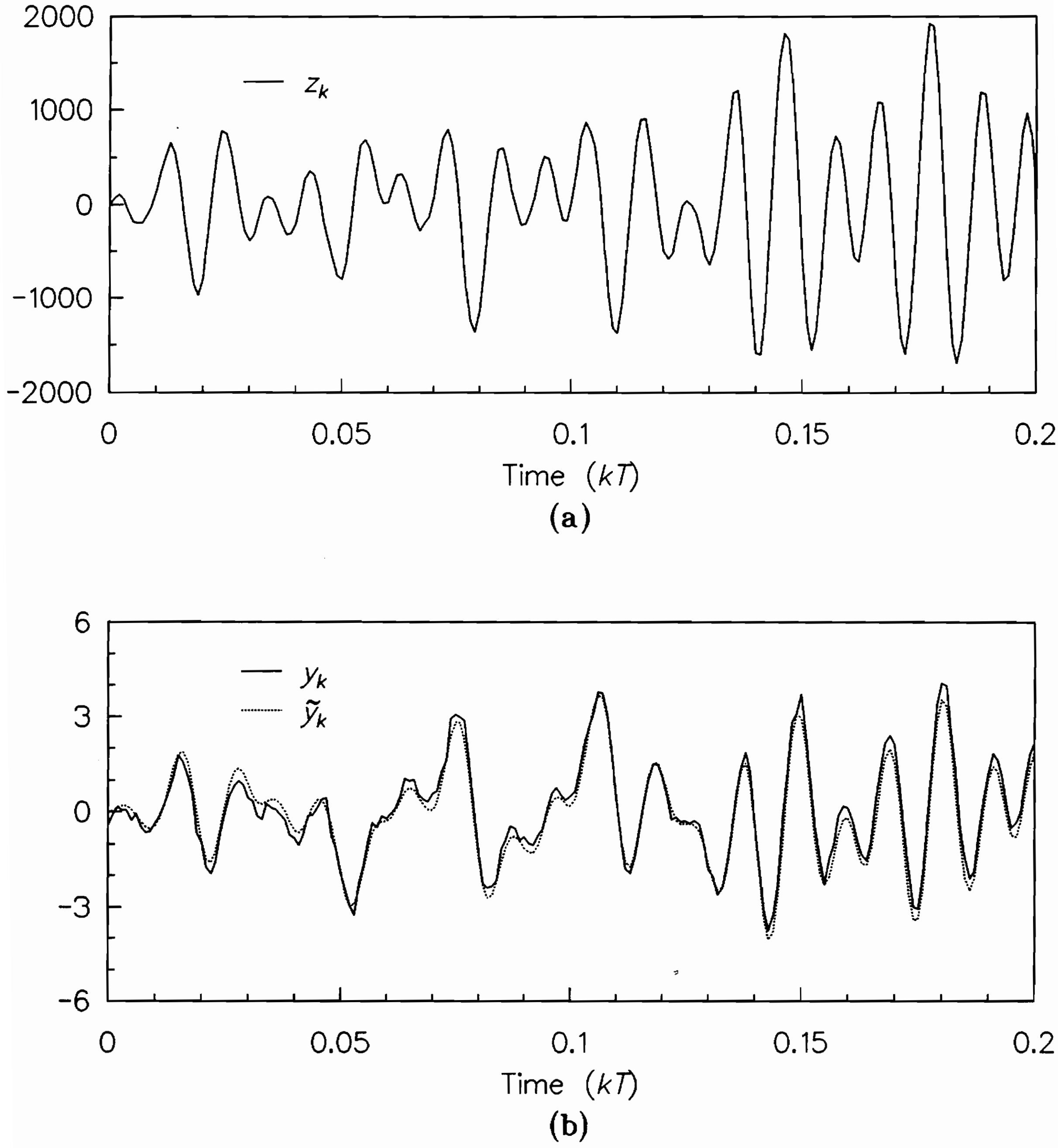


Figure 13. (a) Mode 1 acceleration measurement z_k, (b) mode 1 actual velocity measurement y_k and *approximate*-velocity measurement $\tilde{y}_k$.

Each plot in Figure 14 shows the position state from mode 1 and an overall position state estimate from a particular MGS adaptive filter. Figures 14a, b, c, and d follow the designations of Table 6: estimator a, $\hat{x}^{v}_{k+1}$; estimator b, $\hat{x}^{\tilde{a}}_{k+1}$; estimator c, $\hat{x}^{\tilde{v}}_{k+1}$; estimator d, $\hat{x}^{\tilde{\zeta}}_{k+1}$. Each plot in Figure 15 shows the velocity state from mode 1 and an overall velocity state estimate from a particular MGS adaptive filter. The states and state estimates are shown for mode 2 in Figures 16 and 17, and for 0.4 seconds of mode 3 in Figures 18 and 19. Apart from benchmark estimator a, the best estimates appear to come from estimators d and c, while estimator b deviates the most from the actual state.

Figure 20a shows the slightly noisy weighting terms used in estimator a. Figure 20b shows the weighting terms used in estimators b and c, which are smoother due to the effect of integrating the acceleration measurements. Figure 20c shows the weighting terms used in estimator d, which are also smoother than those of estimator a, and reach and remain at their peak values more so than those of estimators b and c.

Table 8 shows a normalized mean-square-error percentage measure for the state estimators, calculated using (4.2.33). The mse percentages are the lowest for benchmark estimator a, followed by estimators d, c, and a.

Figure 21a shows the three mode 1 residuals (4.2.3) produced from the MGS adaptive filter using the zero dc gain force to acceleration modal system (A.5) of appendix A. The three residuals become equal once steady-state is reached. Figure 21b shows the three mode 1 residuals produced from the MGS adaptive filter using the nonzero dc gain force to velocity modal system (A.6) of appendix A. The residuals remain biased throughout the simulation.

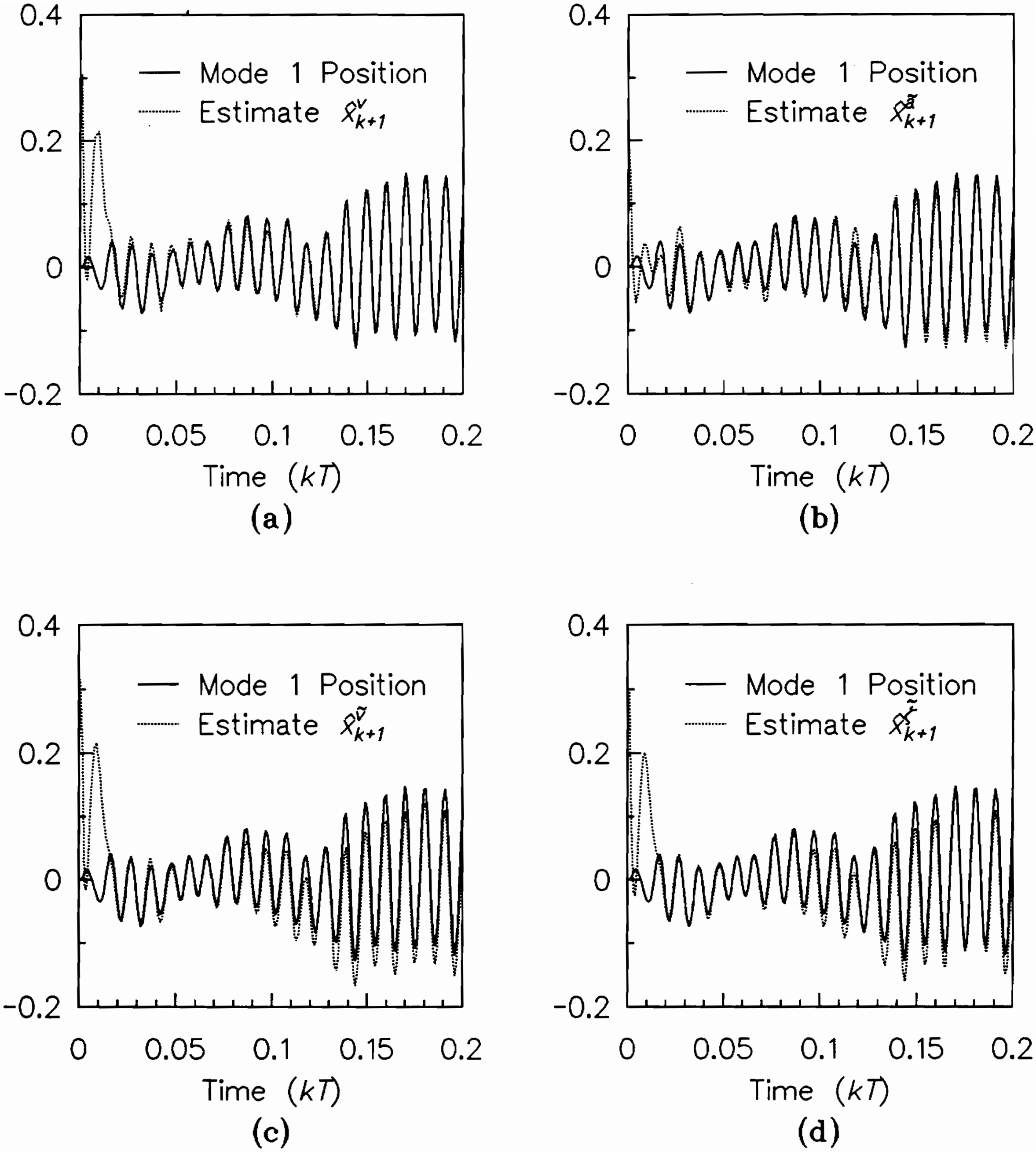


Figure 14. Mode 1 position: (a) State and estimate $\hat{x}^{v}_{k+1}$, (b) state and estimate $\hat{x}^{\tilde{a}}_{k+1}$, (c) state and estimate $\hat{x}^{\tilde{v}}_{k+1}$, (d) state and estimate $\hat{x}^{\tilde{\zeta}}_{k+1}$.

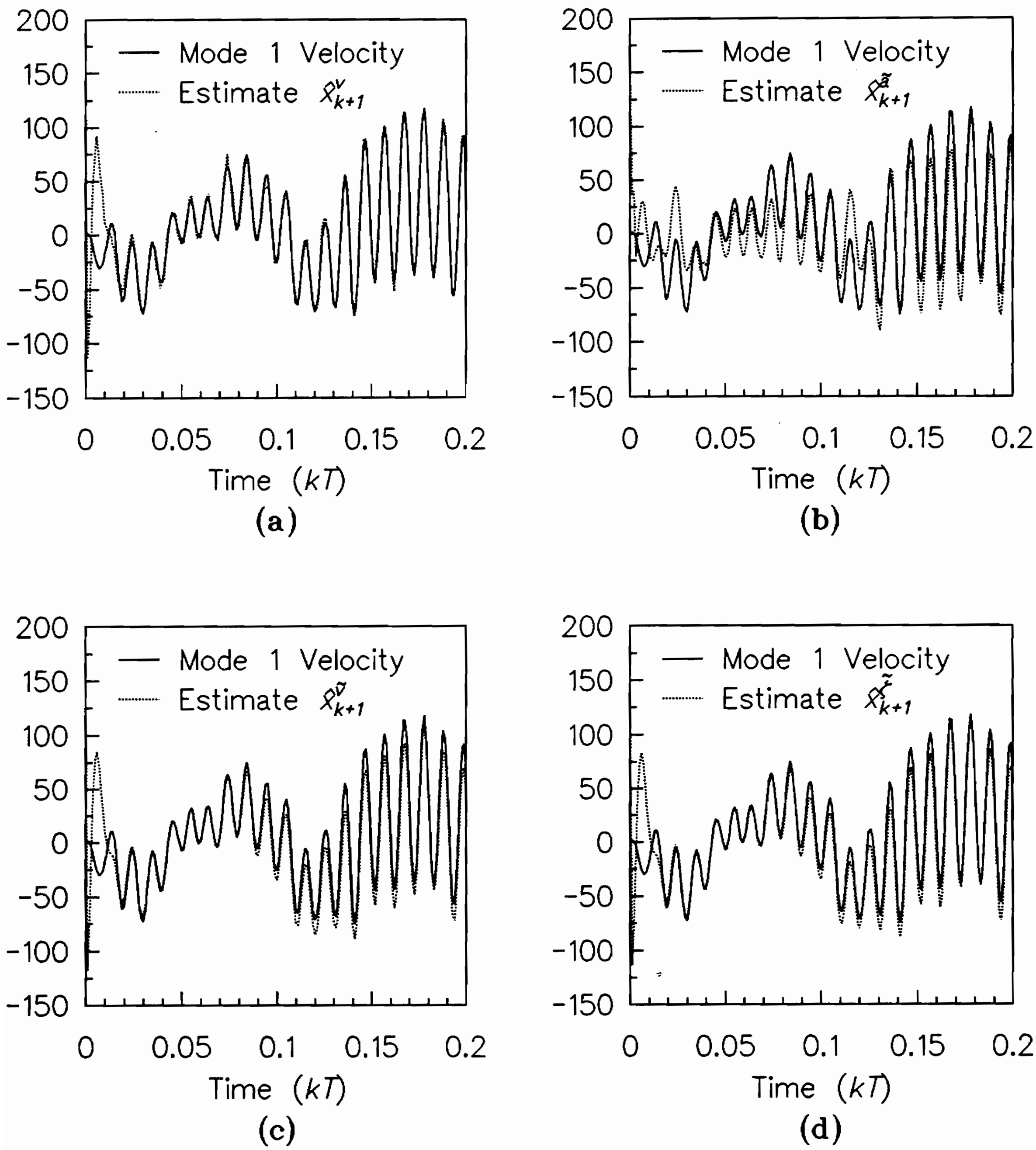


Figure 15. Mode 1 velocity: (a) State and estimate $\hat{x}^{v}_{k+1}$, (b) state and estimate $\hat{x}^{\tilde{a}}_{k+1}$, (c) state and estimate $\hat{x}^{\tilde{v}}_{k+1}$, (d) state and estimate $\hat{x}^{\tilde{\zeta}}_{k+1}$.

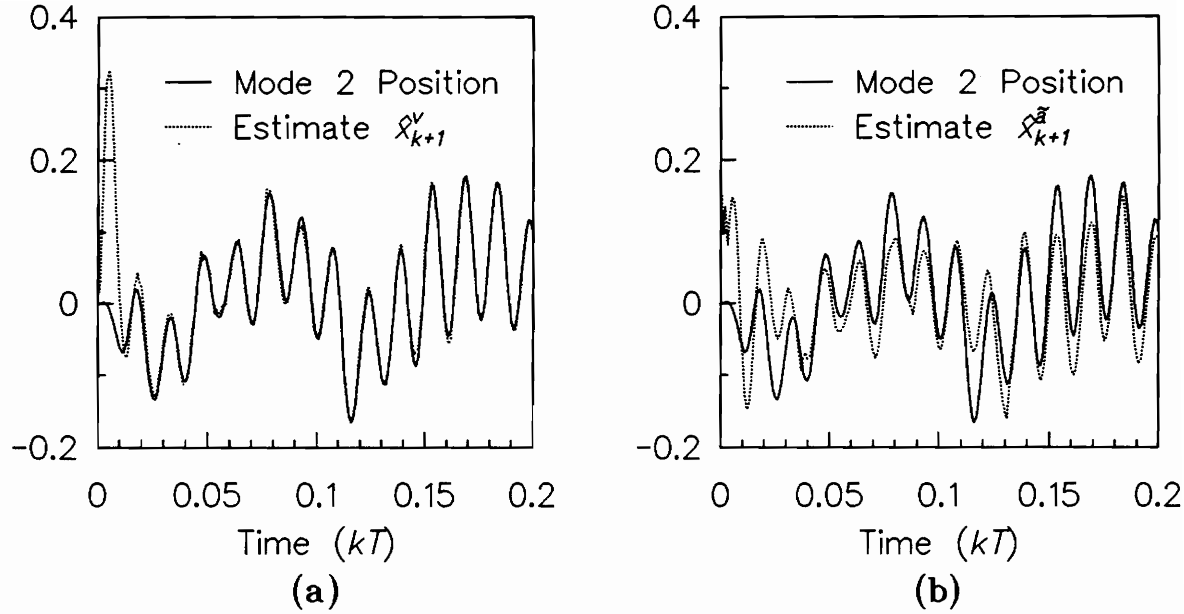


Figure 16. Mode 2 position: (a) State and estimate $\hat{x}^{v}_{k+1}$, (b) state and estimate $\hat{x}^{\tilde{a}}_{k+1}$, (c) state and estimate $\hat{x}^{\tilde{v}}_{k+1}$, (d) state and estimate $\hat{x}^{\tilde{\zeta}}_{k+1}$.

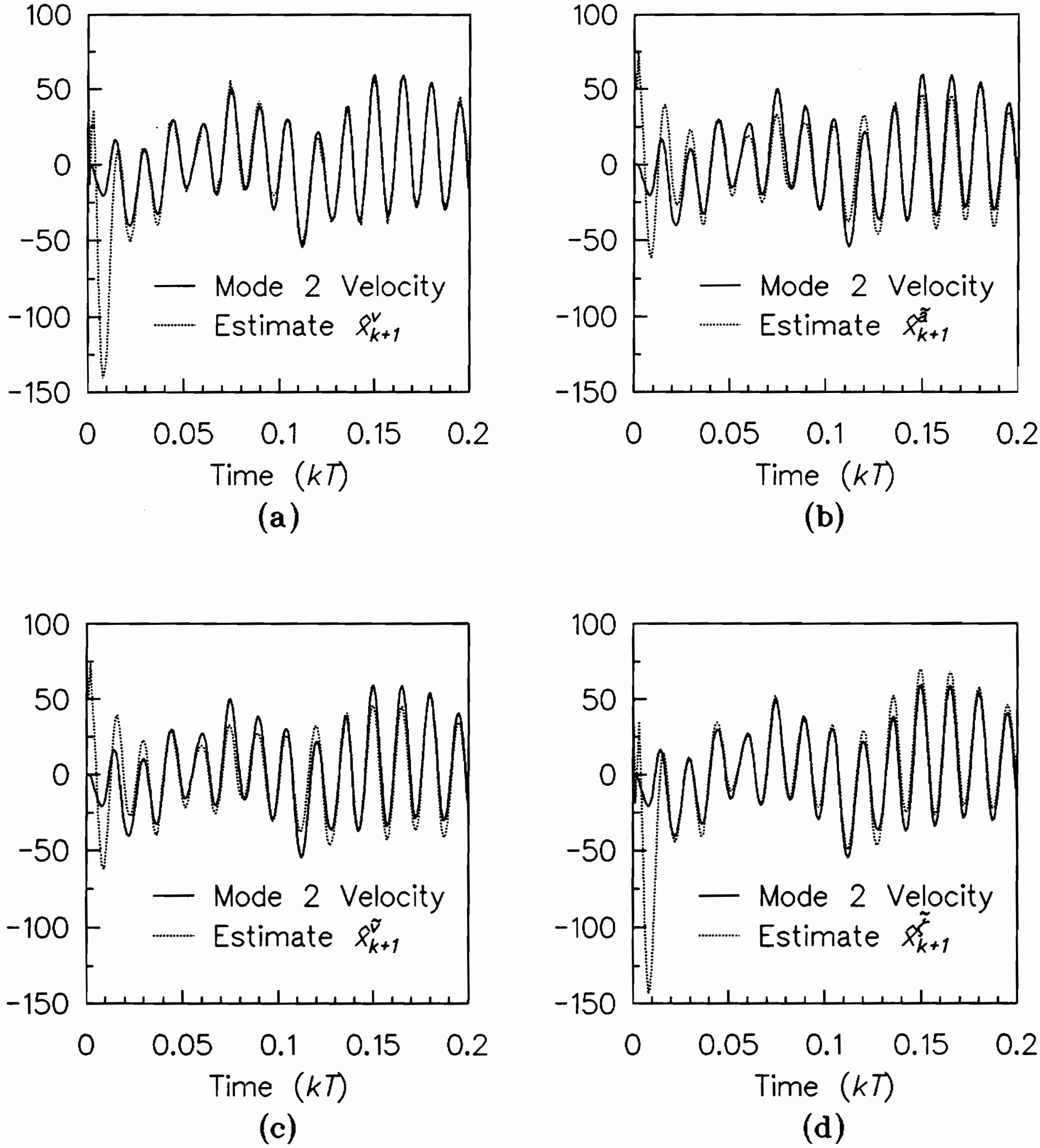


Figure 17. Mode 2 velocity: (a) State and estimate $\hat{x}^{v}_{k+1}$, (b) state and estimate $\hat{x}^{\tilde{a}}_{k+1}$, (c) state and estimate $\hat{x}^{\tilde{v}}_{k+1}$, (d) state and estimate $\hat{x}^{\tilde{\zeta}}_{k+1}$.

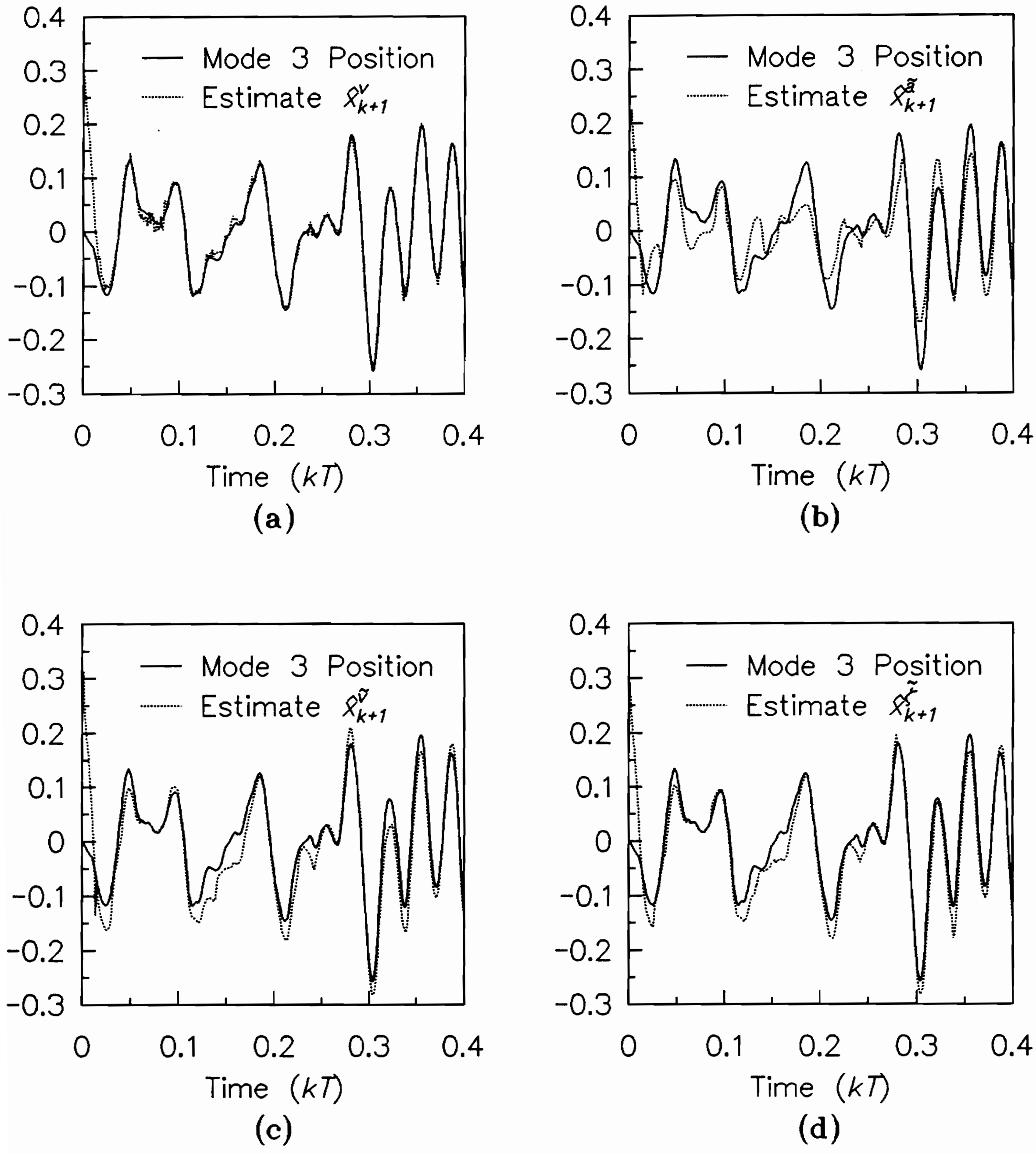


Figure 18. Mode 3 position: (a) State and estimate $\hat{x}^{v}_{k+1}$, (b) state and estimate $\hat{x}^{\tilde{a}}_{k+1}$, (c) state and estimate $\hat{x}^{\tilde{v}}_{k+1}$, (d) state and estimate $\hat{x}^{\tilde{\zeta}}_{k+1}$.

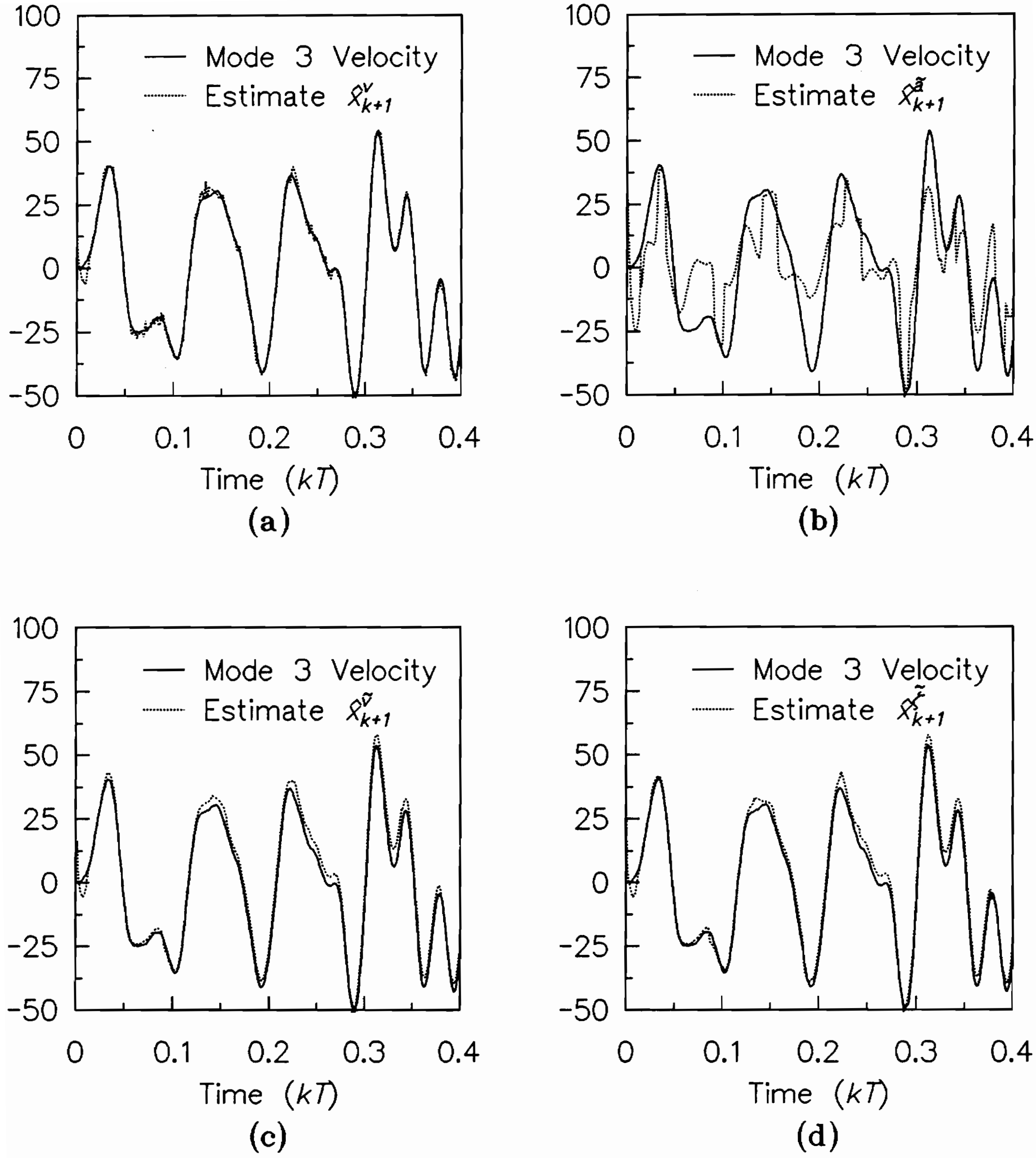


Figure 19. Mode 3 velocity: (a) State and estimate $\hat{x}^{v}_{k+1}$, (b) state and estimate $\hat{x}^{\tilde{a}}_{k+1}$, (c) state and estimate $\hat{x}^{\tilde{v}}_{k+1}$, (d) state and estimate $\hat{x}^{\tilde{\zeta}}_{k+1}$.

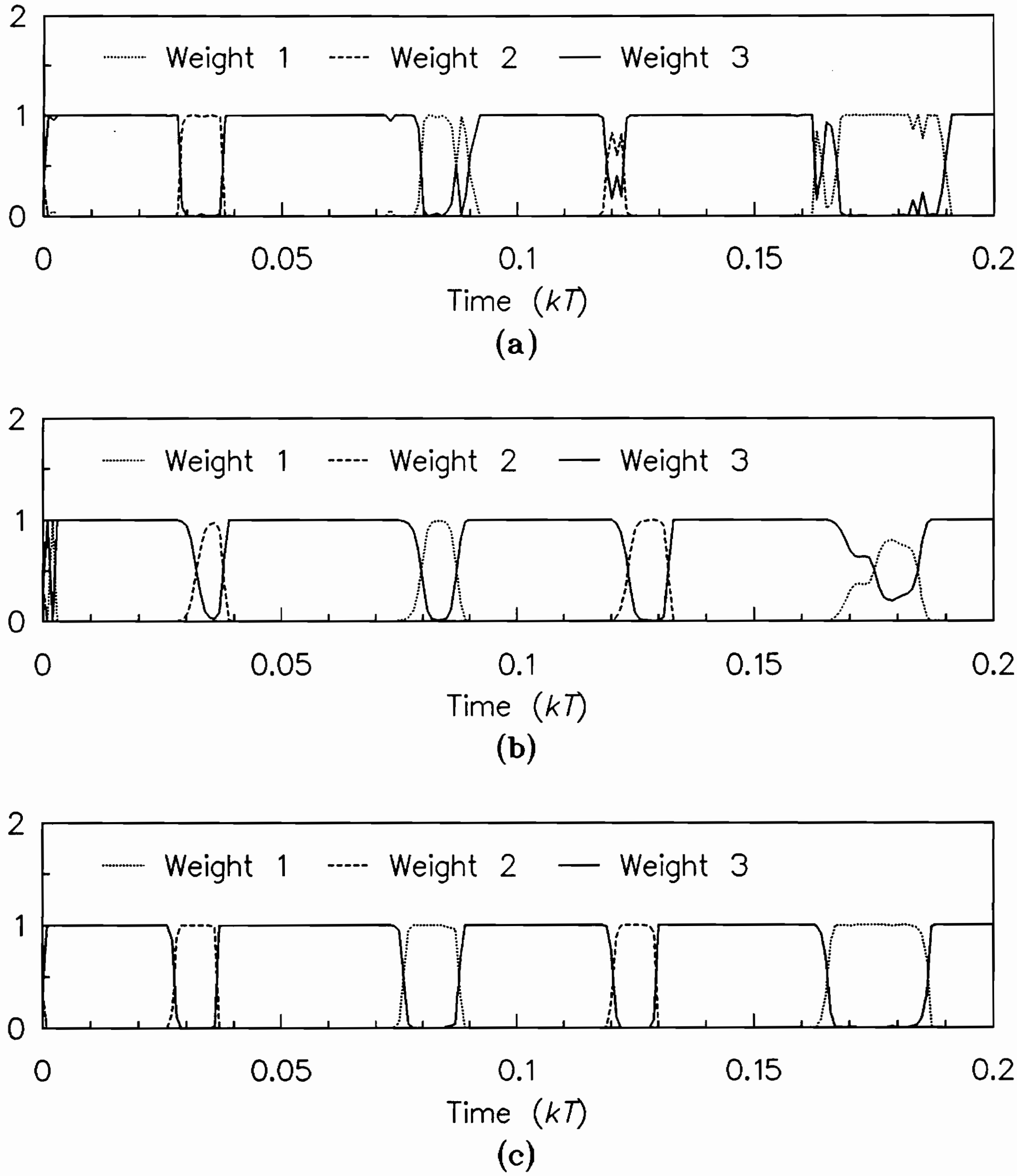


Figure 20. (a) Weighting terms, estimator a, (b) weighting terms, estimators b and c, (c) weighting terms, estimator d.

Table 8.

Normalized mean-square-error percentage for state estimates.

State	Estimator a $\hat{x}^{v}_{k+1}$	Estimator b $\hat{x}^{\tilde{a}}_{k+1}$	Estimator c $\hat{x}^{\tilde{v}}_{k+1}$	Estimator d $\hat{x}^{\tilde{\zeta}}_{k+1}$
Mode 1 position	1.60	3.62	17.04	11.71
Mode 1 velocity	0.87	26.70	6.93	5.93
Mode 2 position	1.02	33.83	15.32	12.20
Mode 2 velocity	1.74	11.21	6.09	4.18
Mode 3 position	0.94	24.73	16.70	13.06
Mode 3 velocity	0.39	39.03	0.34	0.38

$$\text{mse }\% = \frac{E[(x-\hat{x})^2]}{E[x^2]} \times 100$$

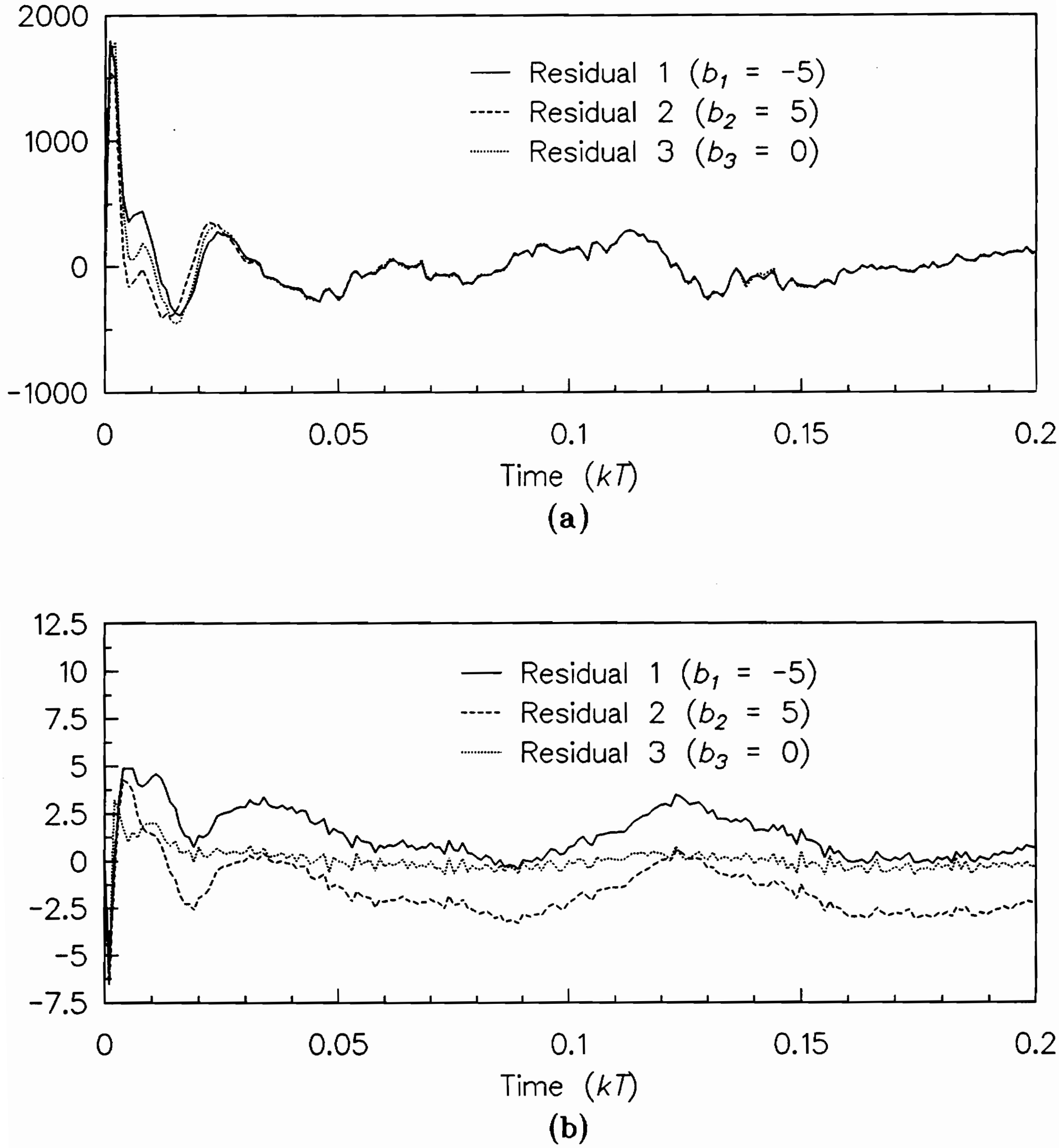


Figure 21. (a) Residuals using the zero dc gain force to acceleration modal system, (b) residuals using the nonzero dc gain force to velocity modal system.

5.0 PARAMETER UPDATING METHODS

As described in chapter 4, the input to the modal system is unknown, but its frequency characteristics are known and the input is always present. It is also not directly measurable by the sensors. The nonGaussian nature of the stochastic FM input signal is modeled using the b_i and S_i parameters of the Gaussian sum signal model. These parameters are closely related to the amplitude of this signal, due to the nature of its probability density function (4.2.10). Two of the b_i bias terms are equal to the positive and negative peak amplitudes of the FM signal, with small corresponding variances. The S_i term associated with the $b_i = 0$ bias term has a value of roughly $(\frac{2}{5}A_u)^2$, where A_u is the peak value of the input.

The stochastic FM signals used in the simulation examples in chapter 4 had constant amplitudes. In practice, the amplitude may shift over a period of time. For good performance under these conditions, the MGS adaptive filter must have some approximate knowledge of the amplitude of the FM signal in order to update the b_i and S_i parameters. Two methods of updating the parameters of the MGS adaptive filter are described in this chapter. One method involves processing the measurement residual of the filter incorporating the $b_i = 0$ (zero-bias) term. The second method involves the use of a Gaussian double-sum to detect a shift in amplitude. A simulation example is used to examine the performance of these two methods.

5.1 Zero-Bias Measurement Residual Method

One method of updating the b_i and S_i parameters of the MGS adaptive filter involves processing the measurement residual of the filter in the parallel bank of Kalman filters incorporating the $b_i = 0$ (zero-bias) term. Using the modal system model of chapter 4 and the MGS adaptive filter equations of Table 3, the i^{th} measurement residual is

$$r^i_{k+1} = z_{k+1} - Db_i - H(\Phi\hat{x}^i_k + \Gamma b_i) \tag{5.1.1}$$

For notational clarity, set $i = 0$ and $b_0 = 0$ in (5.1.1) to produce the zero-bias measurement residual

$$r^0_{k+1} = z_{k+1} - H\Phi\hat{x}^0_k \tag{5.1.2}$$

Substituting measurement equation (4.1.2) at time iteration k+1 into (5.1.2) gives

$$r^0_{k+1} = Hx_{k+1} + Du_{k+1} + v_{k+1} - H\Phi\hat{x}^0_k \tag{5.1.3}$$

and substituting plant equation (4.1.1) into (5.1.3) gives

$$r^0_{k+1} = H(\Phi x_k + \Gamma u_k + \Psi w_k) + Du_{k+1} + v_{k+1} - H\Phi\hat{x}^0_k \tag{5.1.4}$$

Rearranging (5.1.4) produces

$$r_{k+1}^0 = H\Phi(x_k - \hat{x}_k^0) + H\Gamma u_k + Du_{k+1} + H\Psi w_k + v_{k+1} \tag{5.1.5}$$

Now, assuming that the input signal u_k is much larger than both w_k and v_{k+1} such that these two terms can be neglected, assuming that the input signal changes slowly enough such that u_k is approximately equal to u_{k+1}, and setting the error of the zero-bias filter to be $e_k^0 = (x_k - \hat{x}_k^0)$, r_{k+1}^0 of (5.1.5) becomes approximately equal to

$$r_{k+1}^0 \simeq H\Phi e_k^0 + (H\Gamma + D)u_k \tag{5.1.6}$$

Premultiplying both sides of (5.1.6) by the inverse of $(H\Gamma + D)$ if $(H\Gamma + D)$ is square, or by its pseudo-inverse if non-square, and taking the expected value, the zero-bias measurement residual is processed as

$$\tilde{u}_{k+1} = (H\Gamma + D)^{-1}E[r_{k+1}^0] \simeq (H\Gamma + D)^{-1}H\Phi E[e_k^0] + E[u_k] \tag{5.1.7}$$

Assuming that the expected value of the error is $E[e_k^0] = 0$, and given that $E[u_k] = u_k$, the processed zero-bias measurement residual is approximately equal to the input signal

$$\tilde{u}_{k+1} \simeq u_k \tag{5.1.8}$$

Information contained in $\tilde{u}_{k+1}$ is used to periodically update the b_i and S_i parameters of the MGS adaptive filter. The zero-bias residual (5.1.2) is computed and stored at each iteration. After n iterations, the residual sequence is processed using (5.1.7), producing (5.1.8). The positive and negative peak values of (5.1.8) are found from the sequence, and the power of the sequence is calculated. A combination of these values is used to update the b_i and S_i parameters in order to reduce errors related to the assumptions used in the development of (5.1.8). Due to the sinusoidal nature of the input signal, the b_1 (negative-bias) and b_2 (positive-bias) parameters can be updated using

$$b_1 = \beta P_n - (1-\beta)\sqrt{2P_{sin}} \tag{5.1.9}$$

$$b_2 = \beta P_p + (1-\beta)\sqrt{2P_{sin}} \tag{5.1.10}$$

where P_n is the negative peak value of the sequence, P_p is the positive peak value of the sequence, P_{sin} is the power of the sequence, and β is a weighting factor between zero and one. Once b_2 (5.1.10) is updated, the S_3 (variance of the zero-bias filter) parameter is updated using

$$S_3 = (0.4b_2)^2 \tag{5.1.11}$$

After all the parameters are updated, the sequence is cleared and a new sequence is initiated for the next n-sample update.

A simulation example illustrating the use of the zero-bias measurement residual method of updating the b_i and S_i parameters is now given. A modified first-order discrete-time system from [60] is used, given by

$$x_{k+1} = \Phi x_k + \Gamma u_k + \Psi w_k \tag{5.1.12}$$

$$z_k = Hx_k + Du_k + v_k \tag{5.1.13}$$

with values

$$\Phi = 0.7 \qquad \Gamma = 1.0 \qquad \Psi = 1.0 \qquad H = 0.5 \qquad D = 2.0 \tag{5.1.14}$$

The plant noise w_k is zero-mean white Gaussian with covariance $Q = 0.05$, and is uncorrelated with measurement noise v_k, which is also zero-mean white Gaussian with covariance $R = 0.05$.

The nonGaussian stochastic FM input signal (4.2.9) in this example has an amplitude of $A_u = 5$ for the first 60 iterations and an amplitude of $A_u = 2$ for the remaining 140 iterations. The input is modeled using the Gaussian sum signal model described in section 4.1

$$u_k = b_k + n_k \tag{5.1.15}$$

Three filters of the form found in Table 3 are used in the simulation. The initial b_i and S_i parameters for the three filters are

$$b_1 = -5 \qquad b_2 = 5 \qquad b_3 = 0 \tag{5.1.16}$$

$$S_1 = 0.04 \qquad S_2 = 0.04 \quad S_3 = 4.28 \tag{5.1.17}$$

The MGS adaptive filter is initialized with equally-valued weighting terms

$$w_0^j = \frac{1}{3}, \quad \text{for } j = 1,2,3 \tag{5.1.18}$$

A Markov transition probability matrix Θ_{bS}^{ij}, consisting of θ_{bS}^{ij} elements, is configured with a high probability that the bias term does not switch from one value to another, and a low probability that the bias term does switch, such as

$$\Theta_{bS}^{ij} = \begin{bmatrix} .90 & .05 & .05 \\ .05 & .90 & .05 \\ .05 & .05 & .90 \end{bmatrix} \tag{5.1.19}$$

The initial values of the state, state estimate, and error covariance for the MGS adaptive filter are

$$x_0 = 0 \qquad \hat{x}_0 = 16 \qquad P_o = 10000 \tag{5.1.20}$$

Figure 22 compares the stochastic FM signal with the processed zero-bias measurement residual (5.1.8). The two are quite similar in appearance. Figure 23a shows the state x_{k+1} and the state estimate $\hat{x}_{k+1}$ without using the parameter updating method. The initial estimates are good since the parameters are properly matched to the input. However, after the input switches amplitude from $A_u = 5$ to $A_u = 2$, the estimates degrade since the parameters are now mismatched. Figure 24a shows the state x_{k+1} and the state estimate $\hat{x}_{k+1}$ using the parameter updating method every 30 iterations. The initial estimates track well. Then the estimator goes through a period of learning a new set of parameter values as the amplitude of the input changes. Once the correct parameters are found, the estimator tracks with minimal error. Figure 24a shows the weighting terms of the MGS adaptive filter with no parameter updating, and Figure 24b shows the weighting terms of the MGS adaptive filter with parameter updating. Note how only the MGS filter with parameter updating develops proper weighting terms for the estimates during the time the signal has an amplitude of $A_u = 2$.

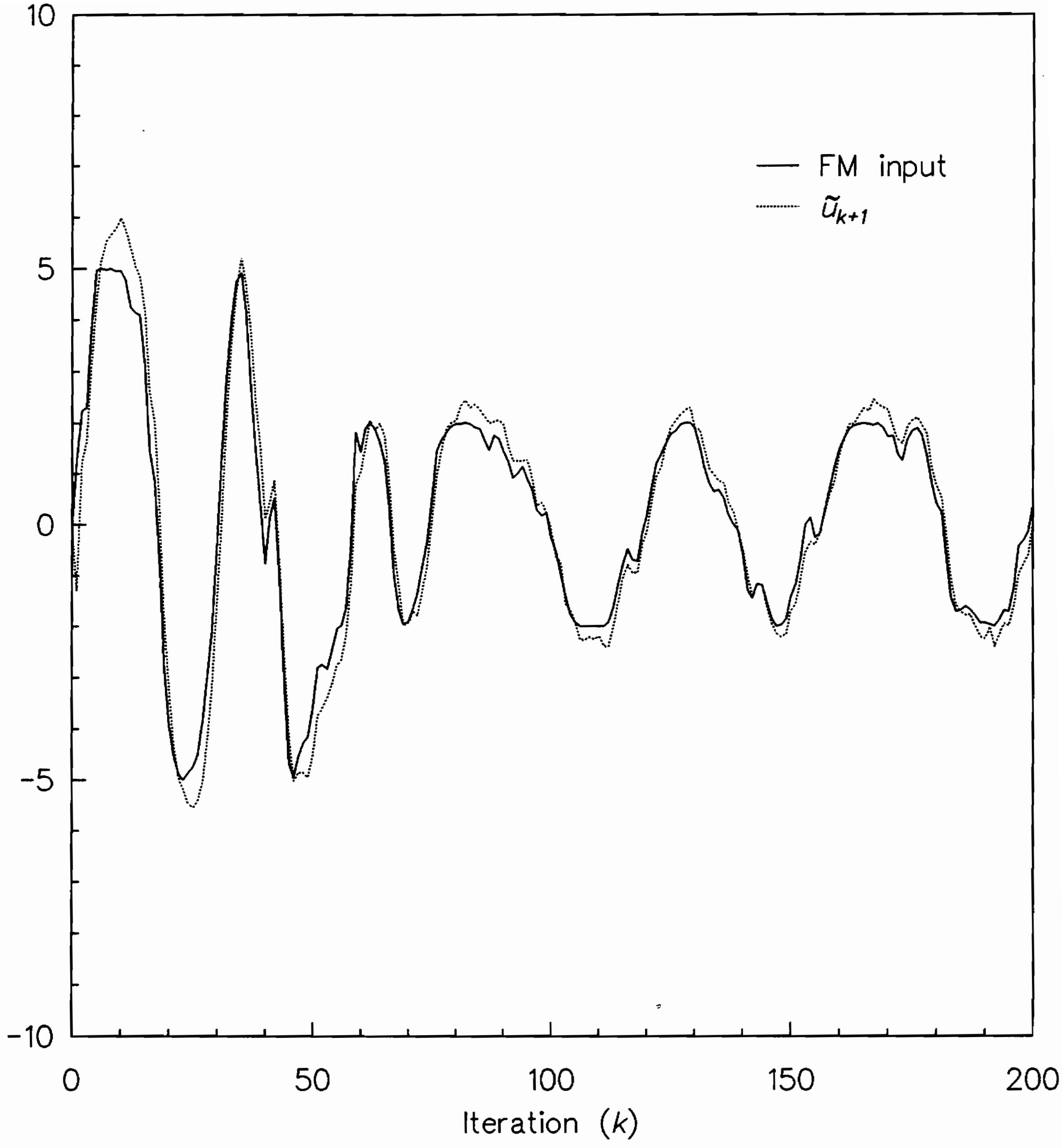


Figure 22. Stochastic FM signal and processed zero-bias measurement residual.

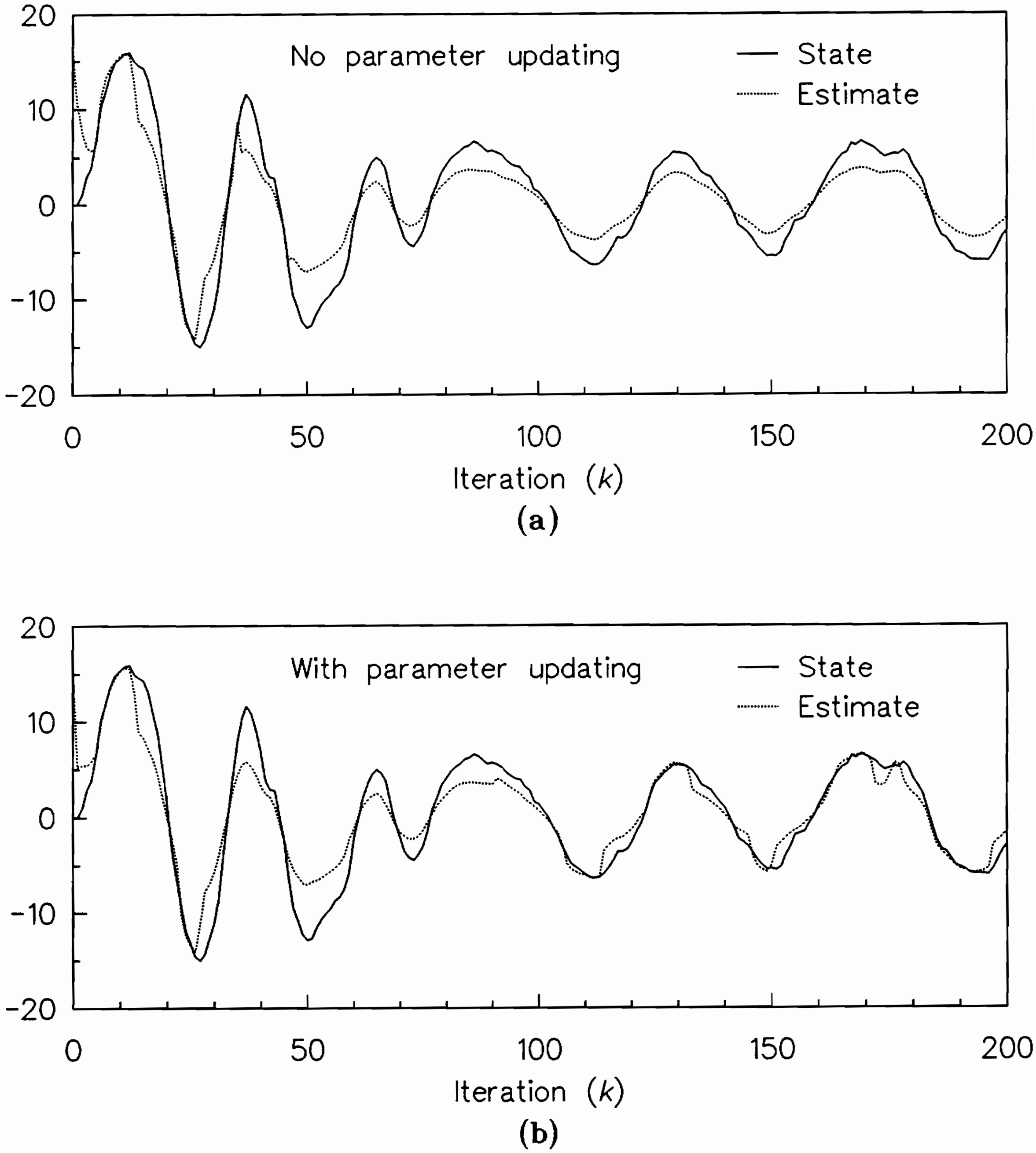
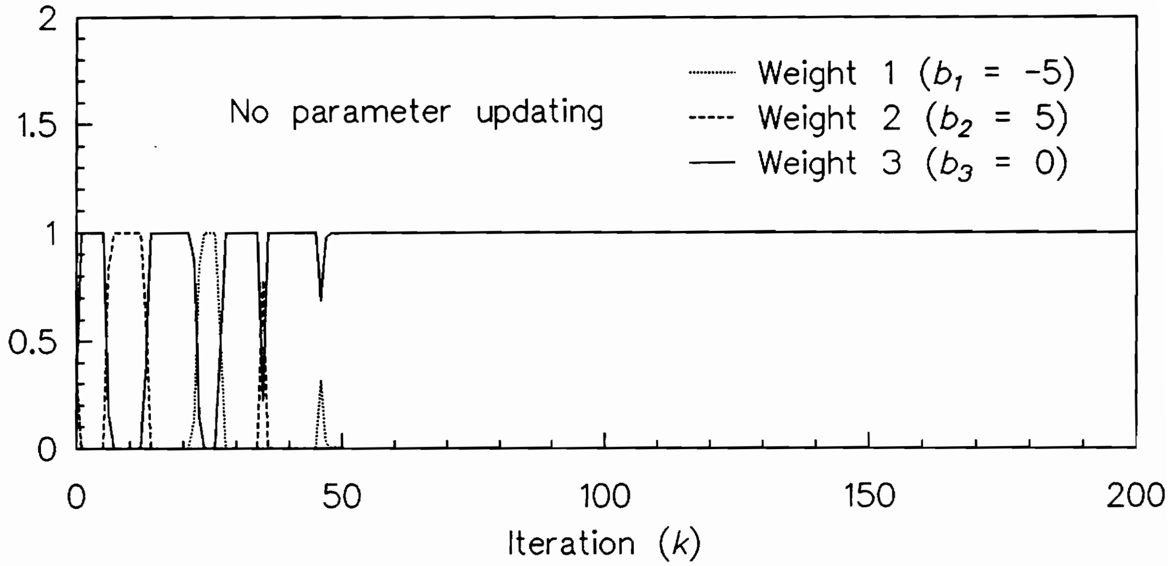
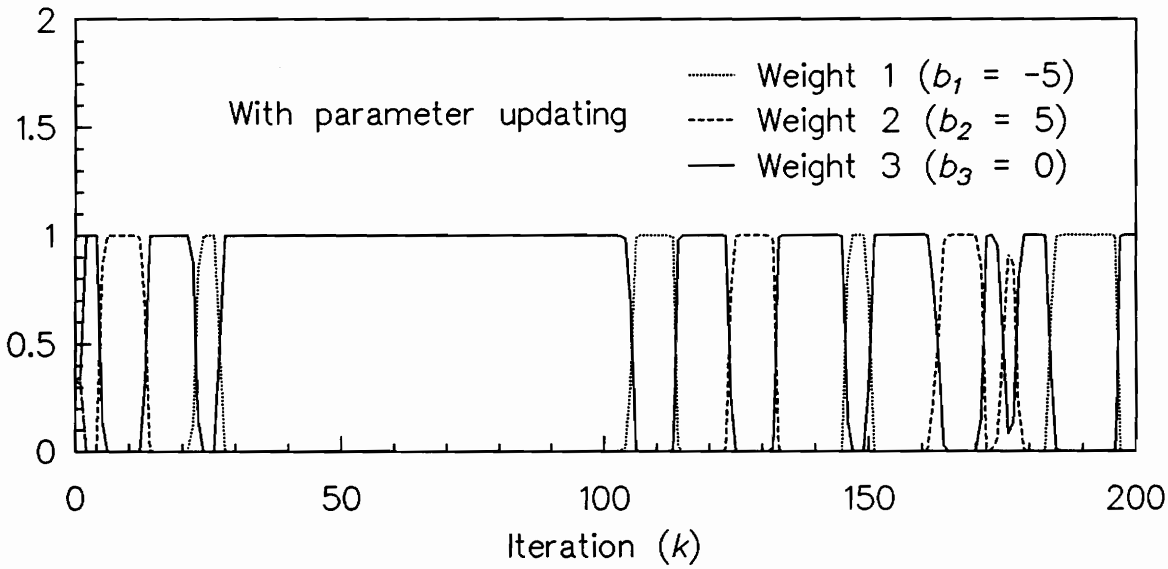


Figure 23. (a) State and state estimate with no parameter updating, (b) state and state estimate using zero-bias measurement residual parameter updating method.



(a)



(b)

Figure 24. (a) Weighting terms with no parameter updating, (b) weighting terms using zero-bias measurement residual parameter updating method.

5.2 Gaussian Double-Sum Method

A second method of updating the parameters of the MGS adaptive filter involves the use of a Gaussian double-sum to detect a shift in amplitude. The nonGaussian signal model (5.1.12) has a Gaussian sum probability density function of the form

$$p(u) = \sum_{i=1}^{M} P_i \mathrm{N}[b_i, S_i] \tag{5.2.1}$$

Now, a new nonGaussian input signal model is generated by multiplying the original model by a constant gain, producing

$$s = Au = A(b+n) \tag{5.2.2}$$

The gain A is a semi-Markov process with values randomly selected from a fixed set of discrete values, characterized by a delta probability density function

$$p(A) = \sum_{j=1}^{N} P_j^A \delta(A - A_j) \tag{5.2.3}$$

with

$$\sum_{j=1}^{N} P_j^A = 1; \quad P_j^A \geq 0 \ \text{ for } j = 1, 2, \ldots, N \tag{5.2.4}$$

This process can be thought of as a randomly-switching gain, each gain value A_j having probability P_j^A. It is assumed that A and u are statistically independent random processes.

The probability density function of the new nonGaussian input signal model (5.2.2) is found using a general result from [61]

$$p(s) = \int_{-\infty}^{\infty} \frac{1}{|A|} \, p_u(s/A) p_A(A) dA \tag{5.2.5}$$

substituting (5.2.3) into (5.2.5) gives

$$p(s) = \int_{-\infty}^{\infty} \frac{1}{|A|} \, p_u(s/A) \sum_{j=1}^{N} P_j^A \delta(A - A_j) dA \tag{5.2.6}$$

Interchanging integration and summation

$$p(s) = \sum_{j=1}^{N} P_j^A \left[\int_{-\infty}^{\infty} \frac{1}{|A|} \, p_u(s/A) \delta(A - A_j) dA \right] \tag{5.2.7}$$

Using the sifting property of the delta function [54]

$$f(a) = \int_{-\infty}^{\infty} f(x) \delta(x - a) dx \tag{5.2.8}$$

the integral of (5.2.7) is evaluated as

$$p(s) = \sum_{j=1}^{N} P_j^A \left[\frac{1}{|A_j|} \, p_u(s/A_j) \right] \tag{5.2.9}$$

The Gaussian density $\mathrm{N}[b_i, S_i]$ of (5.2.1) has the form

$$\mathrm{N}[b_i, S_i] = \frac{1}{\sqrt{2\pi S_i}} \, \mathrm{e}^{-\frac{1}{2}\left[\frac{(u - b_i)^2}{S_i}\right]} \tag{5.2.10}$$

Solving (5.2.2) for $u = s/A_j$ and substituting into (5.2.10) produces

$$\mathrm{N}[b_i, S_i] = \frac{1}{\sqrt{2\pi S_i}}\, \mathrm{e}^{-\frac{1}{2}\left[\frac{\left(\frac{s}{A_j} - b_i\right)^2}{S_i}\right]} \tag{5.1.11}$$

and rearranging the exponent gives

$$\mathrm{N}[b_i, S_i] = \frac{1}{\sqrt{2\pi S_i}}\, \mathrm{e}^{-\frac{1}{2}\left[\frac{(s - A_j b_i)^2}{A_j^2 S_i}\right]} \tag{5.2.12}$$

Combining (5.2.12) and (5.2.1) and substituting into (5.2.9) produces

$$p(s) = \sum_{i=1}^{M} P_i \sum_{j=1}^{N} P_j^A \frac{1}{\sqrt{2\pi A_j^2 S_i}}\, \mathrm{e}^{-\frac{1}{2}\left[\frac{(s - A_j b_i)^2}{A_j^2 S_i}\right]} \tag{5.2.13}$$

or in abbreviated notation

$$p(s) = \sum_{i=1}^{M} \sum_{j=1}^{N} P_i P_j^A \, \mathrm{N}[A_j b_i, A_j^2 S_i] \tag{5.2.14}$$

Thus, the nonGaussian density function of s can be modeled as the Gaussian double-sum (5.2.14). The A_j parameter is used to account for any shift in the amplitude of the actual nonGaussian input signal. Several values of A_j are selected to cover a range of possible input amplitudes. If the actual amplitude changes, the MGS algorithm will react by selecting the filters using the most properly matched A_j value. However, a total of MN filters is now required to implement the MGS algorithm, rather than only M filters previously required. Following a similar development given in chapter 3, the modified Gaussian double-sum (MGDS) adaptive filter equations for a modal system are summarized in Table 9. A block diagram of the MGDS adaptive filter is shown in Figure 25.

The simulation example of section 5.1 will be used again, this time illustrating the use of the Gaussian double-sum method of updating the MGS adaptive filter parameters. Three b_i and S_i parameter values are used again

$$b_1 = -1 \qquad b_2 = 1 \qquad b_3 = 0 \tag{5.2.15}$$

$$S_1 = 0.04 \qquad S_2 = 0.04 \quad S_3 = 0.16 \tag{5.2.16}$$

In effect, these parameters remain fixed. Choosing the proper A_j parameter causes the MGDS adaptive filter values to change. The S_3 parameter is set to $(\frac{2}{5})^2 = 0.16$ to allow for the correct scaling when multiplied by A_j^2. The S_1 and S_2 values are not scaled by A_j^2 since they model the steep peaks at either end of the stochastic FM probability density function (4.2.10). These two parameters properly model these peaks for all amplitude values.

Two A_j values are used in the MGDS adaptive filter to cover the range of the input signal amplitudes used in the simulation. They are

$$A_1 = 2 \qquad A_2 = 5 \tag{5.2.17}$$

The MGDS adaptive filter is initialized with equally-valued weighting terms

$$w_0^{ij} = \frac{1}{6}, \quad \text{for } i = 1,2,3 \text{ and } j = 1,2 \tag{5.2.18}$$

Table 9.

Modified Gaussian double-sum adaptive filter equations for a modal system.

System:
$$x_{k+1} = \Phi x_k + \Gamma s_k + \Psi w_k$$
$$z_k = H x_k + D s_k + v_k$$
$$s_k = A_k(b_k + n_k)$$

Overall estimate:
$$\hat{x}_{k+1} = \sum_{i=1}^{M}\sum_{j=1}^{N} \hat{x}_{k+1}^{ij} p(b_i, S_i, A_j \mid Z_{k+1})$$

Kalman filter equation:
$$\hat{x}_{k+1}^{ij} = \Phi\hat{x}_k^{ij} + \Gamma A_j b_i + K_{k+1}^{ij}[z_{k+1} - DA_j b_i - H(\Phi\hat{x}_k^{ij} + \Gamma A_j b_i)]$$

Kalman gain equations:
$$M_{k+1}^{ij} = \Phi P_k^{ij}\Phi^{\mathrm{T}} + \Gamma A_j^2 S_i \Gamma^{\mathrm{T}} + \Psi Q \Psi^{\mathrm{T}}$$
$$K_{k+1}^{ij} = M_{k+1}^{ij} H^{\mathrm{T}}(H M_{k+1}^{ij} H^{\mathrm{T}} + D A_j^2 S_i D^{\mathrm{T}} + R)^{-1}$$
$$P_{k+1}^{ij} = (I - K_{k+1}^{ij} H) M_{k+1}^{ij}$$

Weighting term:
$$w_{k+1}^{ij} = p(b_i, S_i, A_j \mid Z_{k+1}) = C_{k+1}^{o} p(z_{k+1} \mid b_i, S_i, A_j, Z_k) \sum_{\alpha=1}^{M}\sum_{\beta=1}^{N} \theta_{bS}^{i\alpha}\theta_A^{j\beta} w_k^{\alpha\beta}$$

with
$$p(z_{k+1} \mid b_i, S_i, A_j, Z_k) = \mathrm{N}[H(\Phi\hat{x}_k^{ij} + \Gamma A_j b_i) + DA_j b_i,\ HM_{k+1}^{ij}H^{\mathrm{T}} + DA_j^2 S_i D^{\mathrm{T}} + R]$$
$$\theta_{bS}^{i\alpha} = p(b_{k+1} = b_i, S_{k+1} = S_i \mid b_k = b_\alpha, S_k = S_\alpha)$$
$$\theta_A^{j\beta} = p(A_{k+1} = A_j \mid A_k = A_\beta)$$
$$w_k^{\alpha\beta} = p(b_k = b_\alpha, S_k = S_\alpha, A_k = A_\beta \mid Z_k)$$

and scale factor C_{k+1}^{o} such that $\sum_{i=1}^{M}\sum_{j=1}^{N} w_{k+1}^{ij} = 1$

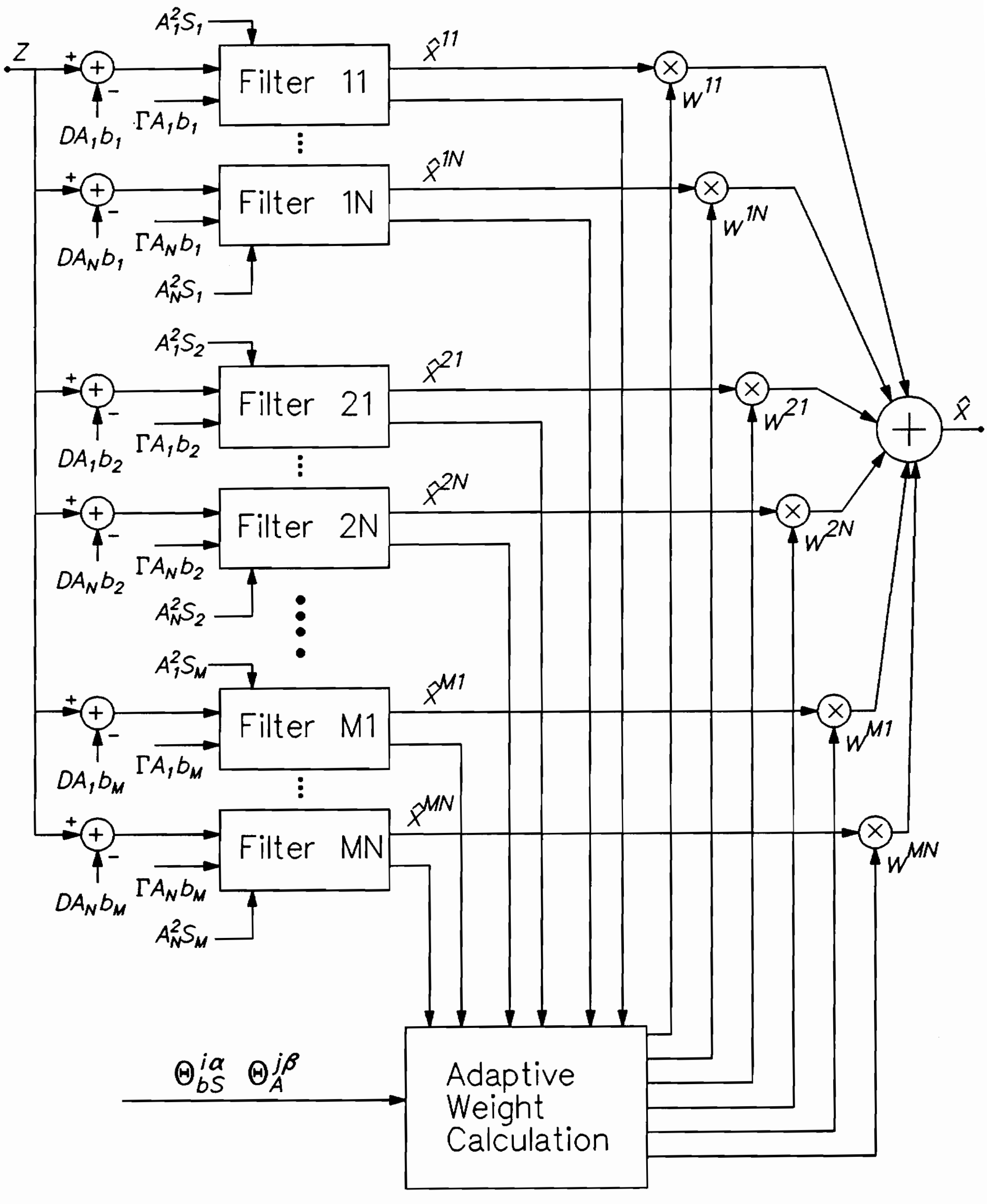


Figure 25. Modified Gaussian double-sum adaptive filter structure.

Two Markov transition probability matrices, $\Theta_{bS}^{i\alpha}$ and $\Theta_{A}^{j\beta}$, are used in the MGS adaptive filter, given by

$$\Theta_{bS}^{i\alpha} = \begin{bmatrix} .90 & .05 & .05 \\ .05 & .90 & .05 \\ .05 & .05 & .90 \end{bmatrix} \qquad \Theta_{A}^{j\beta} = \begin{bmatrix} .95 & .05 \\ .05 & .95 \end{bmatrix} \tag{5.2.19}$$

Figure 26a shows the state x_{k+1} and the same state estimate $\hat{x}_{k+1}$ generated previously without parameter updating. Figure 26b shows the state x_{k+1} and the state estimate $\hat{x}_{k+1}$ using the Gaussian double-sum parameter updating method. The MGDS estimates track well throughout the simulation, without any noticeable learning time.

Figure 27a shows weighting terms 1-3 of the MGDS adaptive filter with $A_j = 2$, and Figure 27b shows weighting terms 4-6 with $A_j = 5$. The input signal starts out with an amplitude of $A_u = 5$, so weighting terms 4-6 initially are in effect. After the amplitude switches to $A_u = 2$, weighting terms 1-3 take over, with one exception. Weight 3 (for $b_i = 0$) should be larger than weight 6 (also for $b_i = 0$) since $A_j = 2$ currently matches the actual input amplitude. However, the measurement residuals for corresponding filters 3 and 6 are equal, because their b_i terms are each zero. This causes the difference in their S_i variance terms to be the deciding factor in calculating the weights. The Gaussian density function in Table 9 used to calculate the weights has a larger variance for weight 6 ($A_j = 5$) than for weight 3 ($A_j = 2$). Therefore, the probability calculated for weight 6 will be larger than the probability calculated for weight 3, so that weight 6 governs the overall state estimate more so than weight 3.

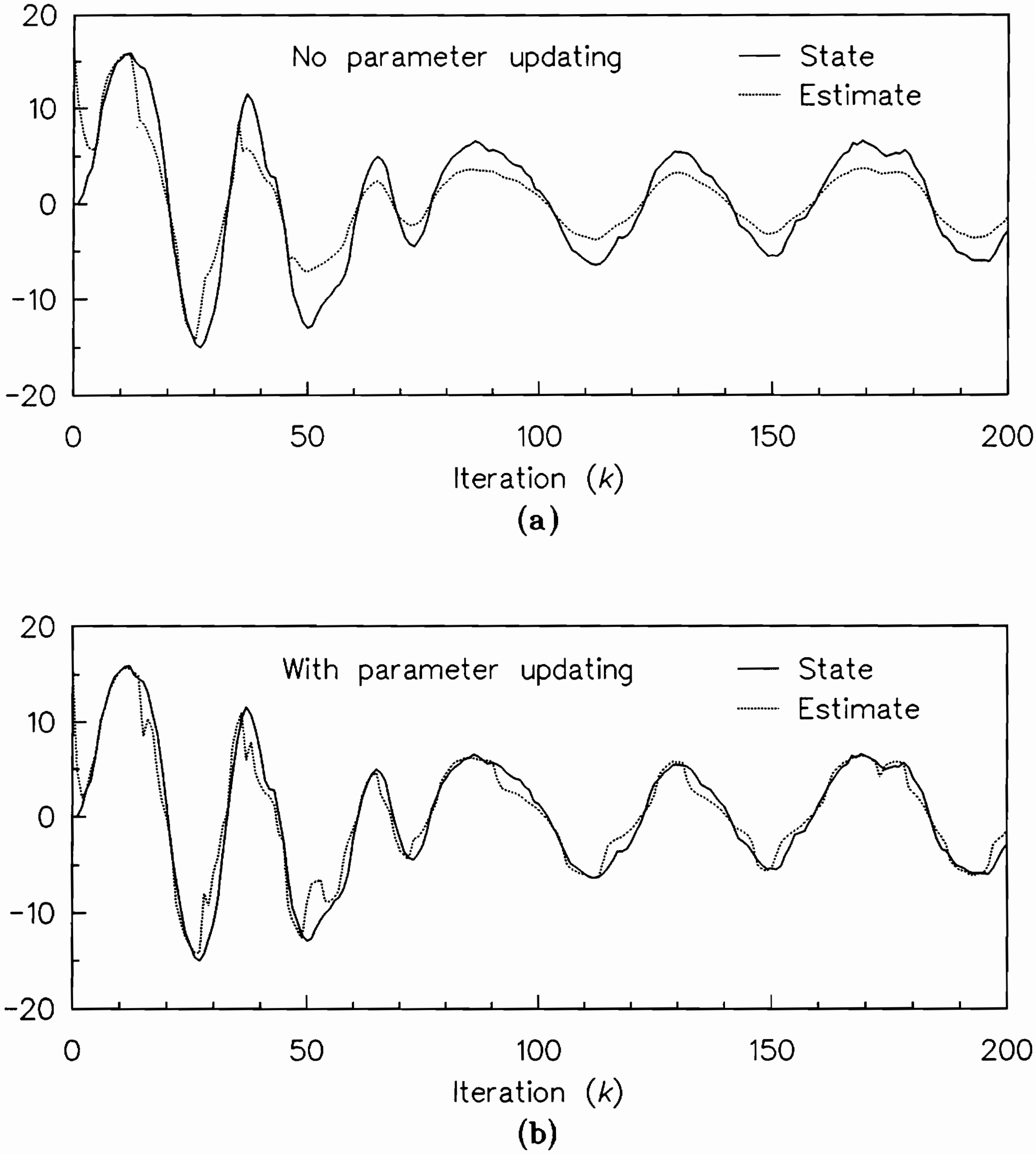


Figure 26. (a) State and state estimate with no parameter updating, (b) state and state estimate using Gaussian double-sum parameter updating method.

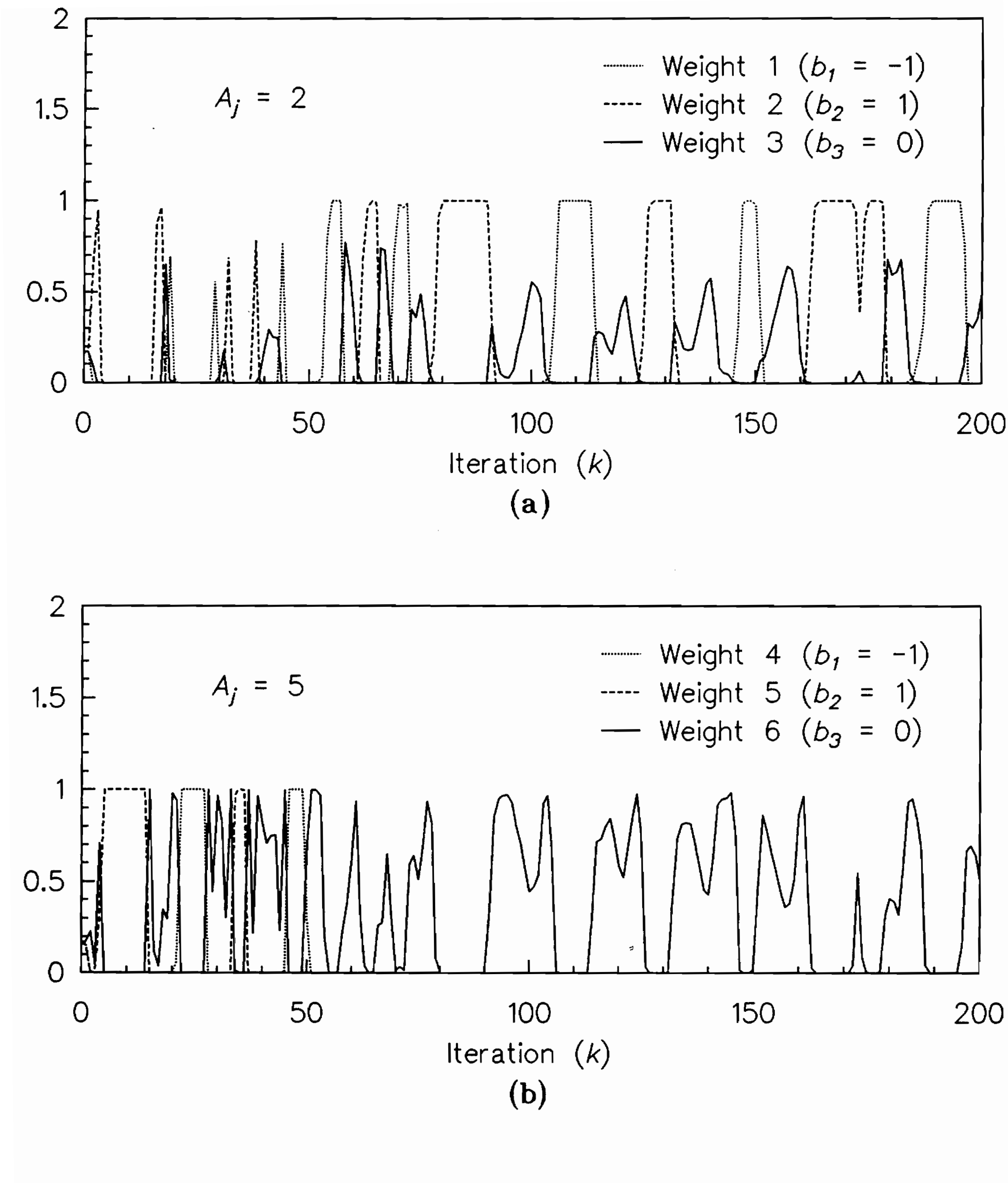


Figure 27. (a) Weighting terms 1-3 with $A_j = 2$, (b) weighting terms 4-6 with $A_j = 5$.

6.0 CONCLUSIONS

This investigation is concerned with effective state estimation of a system driven by an unknown nonGaussian input with additive white Gaussian noise, and observed by measurements containing feedthrough of the same nonGaussian input and corrupted by additional white Gaussian noise. A Gaussian sum (GS) approach has previously been developed [6-8] which can cope with the nonGaussian nature of the input signal. Due to a serious growing memory problem in this approach, a modified Gaussian sum (MGS) estimation technique is developed that avoids the growing memory problem while providing effective state estimation. Several differences between the MGS and GS algorithms are examined, showing the MGS algorithm to be a better performer.

An MGS adaptive filter is derived for a general system and a modal system, with simulation examples performed using a nonGaussian input signal. The modal system simulation results are compared to those produced from an augmented Kalman filter based on an augmented modal system model assuming a narrowband Gaussian input signal. The results show that the MGS adaptive filter provides better estimates than the augmented Kalman filter.

A necessary condition for effective MGS estimation is derived, namely that the system must have a nonzero dc gain. Three alternate estimation procedures are developed to compensate for situations when this condition is not satisfied. The alternate MGS adaptive filters are simulated and their performance results are analyzed and compared.

Two methods of monitoring and updating key parameters of the MGS filter are developed. One is the zero-bias measurement residual method and the second is the Gaussian double-sum method. Simulation results are analyzed to investigate the performance of these methods, with the Gaussian double-sum method proving to be a better performer at the expense of increased computational burden.

Several contributions to the field of applied estimation theory are made from this investigation. These include:

1. The development of a modified Gaussian sum algorithm with nongrowing memory based on a nonGaussian signal model with a Gaussian sum probability density function. Parameters from this model are used directly in the modified Gaussian sum adaptive filter structure.

2. A comparison between the GS filter of [6] and the MGS adaptive filter. The two are similar, but the comparison shows the MGS adaptive filter to be a good improvement to the GS filter.

3. An examination of a necessary condition for effective MGS estimation. This condition provides a simple test to determine if the MGS adaptive filter will work properly for a given system.

4. An alternate configuration of the MGS adaptive filter when the necessary

condition of 3 above is not met. This configuration is applied in several ways, and each is evaluated on a performance basis.

5. Two methods of monitoring and updating key parameters of the MGS adaptive filter. These allow the estimator to react to changes in the input signal level which cause the signal to be nonstationary over long periods of time.

Some suggested directions for future development include a closer examination of the assumptions used in the MGS algorithm to alleviate the increasing computational and storage requirements of the growing memory GS algorithm, applying the MGS algorithm to a larger variety of nonGaussian input signals, and examining how to implement the algorithm in the case of nonlinear systems with nonGaussian inputs.

APPENDIX A

The transfer matrix from actuator force inputs to sensor acceleration outputs for the modal model described in [1] is

$$G(s) = s\sum_{i=1}^{m}\left[\frac{c_i\phi_i\phi_i^{\mathrm{T}}}{s-\lambda_i} + \frac{c_i^*\phi_i^*\phi_i^{*\mathrm{T}}}{s-\lambda_i^*}\right] \tag{A.1}$$

where

m	is the total number of modeled (complex) modes
λ_i	is the i^{th} eigenvalue of the transfer matrix
ϕ_i	is the i^{th} mode shape vector (eigenvector)
c_i	is the i^{th} modal participation factor

An equivalent continuous-time state representation of (A.1) has the form

$$\dot{x}(t) = Ax(t) + Bu(t) \tag{A.2}$$

$$y(t) = Cx(t) + Du(t) \tag{A.3}$$

where

$$x(t) \equiv \begin{bmatrix} q_1(t) & \dot{q}_1(t) & \cdots & q_m(t) & \dot{q}_m(t) \end{bmatrix}^{\mathrm{T}} \tag{A.4}$$

are the modal positions and velocities, $u(t)$ is the m-length vector of actuator input forces, and $y(t)$ is the m-length vector of accelerations at sensor grid points.

Clearly, the dc gain of (A.1) is zero since for $s = 0$, $G(s) = 0$. Differentiating (A.1) once with respect to s removes the s in front of the summation, producing a new transfer matrix from actuator force inputs to sensor velocity outputs. This new transfer matrix has a nonzero dc gain. The discrete-time equations (4.1.1) and (4.1.2) are generated using a zero-order-hold model. Therefore, the discrete-time dc gain (4.2.4) from u_k to z_k is equal to the continuous-time dc gain.

The A, B, C, and D matrices of the state space representation of the force to acceleration transfer matrix are computed as follows. Define

$$\begin{aligned} \Lambda &\equiv \mathrm{diag}\{\lambda_1, \lambda_1^*, \ldots, \lambda_m, \lambda_m^*\} \\ \mathfrak{C} &\equiv \mathrm{diag}\{c_1, c_1^*, \ldots, c_m, c_m^*\} \\ \Phi &\equiv \begin{bmatrix} \phi_1 & \phi_1^* & \cdots & \phi_m & \phi_m^* \end{bmatrix} \end{aligned}$$

Then

$$A = \Lambda, \quad B = \Lambda \mathfrak{C}^{\frac{1}{2}} \Phi^{\mathrm{T}}, \quad C = \Phi \mathfrak{C}^{\frac{1}{2}}, \quad D = \Phi \mathfrak{C} \Phi^{\mathrm{T}} \tag{A.5}$$

The A, B, C, and D matrices of the state space representation of the force to velocity transfer matrix are

$$A = \Lambda, \quad B = \Lambda \mathfrak{C}^{\frac{1}{2}} \Phi^{\mathrm{T}}, \quad C = \Phi \mathfrak{C}^{\frac{1}{2}} \Lambda^{-1}, \quad D = 0 \tag{A.6}$$

The A and B matrices of (A.6) are the same as those from (A.5). This allows the states (A.4) to remain unchanged when either acceleration or velocity outputs are used.

Numerical values for the parameters in (A.5) and (A.6) are generated using the algorithm below. In these equations the subscript i represents the corresponding mode number, and values are generated for $i = 1, 2, \ldots, m$. The eigenvalues are given by

$$\lambda_i = -\alpha_i \pm j\omega_i \tag{A.7}$$

where the α_i are randomly selected with uniform distributions over three intervals: $[1,2]$, $[2,5]$, and $[5,15]$, generally with the highest frequency of occurrence in $[1,2]$, followed by $[5,15]$, and then by $[2,5]$. The ω_i are given by

$$\omega_i = 2\pi f_i \tag{A.8}$$

$$f_i = 100\ \mathrm{e}^{-0.1(\alpha_i - 1)} \tag{A.9}$$

The eigenfunctions are given by

$$\phi_i(x) = \frac{\mathrm{e}^{jk_i x}}{\sqrt{m}} \tag{A.10}$$

where x is the $m \times 1$ vector of grid point locations distributed over a length of 70 meters, and

$$k_i = \frac{\omega_i}{v_i} \tag{A.11}$$

$$v_i = 3000\ \mathrm{e}^{-.025 f_i} \tag{A.12}$$

The magnitudes of the modal participation factors are given by

$$|c_i| = 100\ \mathrm{e}^{.01(f_i - 20)} \tag{A.13}$$

and the phase angles are

$$\angle c_i = \text{random values, uniformly distributed on the interval } [0, 2\pi] \tag{A.14}$$

Using the algorithm above, the parameters of (A.5) and (A.6) are completely specified by α_i, x, and $\angle c_i$ for $i = 1, 2, \ldots, m$ and some positive integer m.

LIST OF REFERENCES

[1] R. L. Moose, H. F. VanLandingham, M. J. Caputi, S. H. Jones, and E. N. Khoury, "Modal control of damped structures," Annu. Rep., ONR Contr. N00014-88-C-0586, Mar. 1990.

[2] R. E. Kalman, "A new approach to linear filtering and prediction problems," *ASME Trans., J. Basic Eng.*, series 82D, pp. 35-45, Mar. 1960.

[3] R. E. Kalman and R. S. Bucy, "New results in linear filtering and prediction theory," *ASME Trans., J. Basic Eng.*, series 83D, pp. 95-108, Mar. 1961.

[4] B. D. O. Anderson and J. B. Moore, *Optimal Filtering.* Englewood Cliffs, NJ: Prentice-Hall, Inc., 1979.

[5] R. F. Stengel, *Stochastic Optimal Control.* New York: John Wiley and Sons, Inc., 1986.

[6] D. L. Alspach, "A Bayesian approximation technique for estimation and control of time-discrete stochastic systems," Ph.D. dissertation, Univ. Calif., San Diego, 1970.

[7] H. W. Sorenson and D. L. Alspach, "Recursive Bayesian estimation using Gaussian sums," *Automatica*, vol. 7, pp. 465-479, 1971.

[8] D. L. Alspach and H. W. Sorenson, "Nonlinear Bayesian estimation using Gaussian sum approximations," *IEEE Trans. Automat. Contr.*, vol. AC-17, no. 4, pp. 439-448, Aug. 1972.

[9] D. L. Alspach, "A parallel processing solution to the adaptive Kalman filtering problem with vector measurements," *Comput. & Elect. Engng.*, vol. 1, pp. 83-94, 1973.

[10] D. L. Alspach, "A parallel filtering algorithm for linear systems with unknown time varying noise statistics," *IEEE Trans. Automat. Contr.*, vol. AC-19, no. 5, pp. 552-556, Oct. 1974.

[11] D. L. Alspach, L. L. Scharf, and A. Abiri, "A Bayesian solution to the problem of state estimation in an unknown noise environment," *Int. J. Control*, vol. 19, no. 2, pp. 265-287, 1974.

[12] D. L. Alspach and J. LaGrotta, "A parallel processing Gaussian sum approach to the Dayton problem," *Proc. 4th Symp. Nonlinear Estimation Theory and Its Applications*, San Diego, CA., pp. 1-16, 1973.

[13] D. L. Alspach, "A Gaussian sum approach to the multi-target identification-tracking problem," *Automatica*, vol. 11, pp. 285-296, 1975.

[14] D. L. Alspach and H. W. Sorenson, "Stochastic optimal control for linear but non-Gaussian systems," *Int. J. Control*, vol. 13, no. 6, pp. 1169-1181, 1971.

[15] D. L. Alspach, "Dual control based on approximate *a posteriori* density functions," *IEEE Trans. Automat. Contr.*, vol. AC-17, no. 5, pp. 689-693, Oct. 1972.

[16] D. T. Magill, "Optimal adaptive estimation of sampled stochastic processes," *IEEE Trans. Auto. Cont.*, vol. AC-10, no. 4, pp. 434-439, Oct. 1965.

[17] G. A. Ackerson and K. S. Fu, "On state estimation in switching environments," *IEEE Trans. Automat. Contr.*, vol. AC-15, no. 15, pp. 10-17, Feb. 1970.

[18] A. G. Jaffer and S. C. Gupta, "Optimal sequential estimation of discrete processes with Markov interrupted observations," *IEEE Trans. Automat. Contr.*, vol. AC-16, no. 5, pp. 471-475, Oct. 1971.

[19] S. Fujita and T. Fukao, "Optimal stochastic control for discrete-time linear system with interrupted observations," *Automatica*, vol. 8, pp. 425-432, 1972.

[20] J. M. H. Bruckner, R. W. Scott, and F. G. Rea, "Analysis of multimodal systems," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-9, no. 6, pp. 883-888, Nov. 1973.

[21] R. A. Howard, "System analysis of semi-Markov processes," *IEEE Trans. Mil. Electron.*, vol. MIL-8, pp. 114-124, Apr. 1964.

[22] R. A. Howard, *Dynamic Probabilistic Systems, Volume II: Semi-Markov and Decision Processes.* New York: John Wiley and Sons, Inc., 1971.

[23] R. L. Moose, "An adaptive estimator with learning for a plant containing semi-Markov switching parameters," Ph.D. dissertation, Duke Univ., 1971.

[24] R. L. Moose and P. P. Wang, "An adaptive estimator with learning for a plant containing semi-Markov switching parameters," *IEEE Trans. on Sys., Man, and Cyber.*, vol. SMC-3, no. 3, pp. 277-281, May 1973.

[25] R. L. Moose, "An adaptive state estimation solution to the maneuvering target problem," *IEEE Trans. Automat. Contr.*, vol. AC-20, no. 3, pp. 359-362, June 1975.

[26] N. H. Gholson and R. L. Moose, "Maneuvering target tracking using adaptive state estimation," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-13, no. 3, pp. 310-317, May 1977.

[27] R. L. Moose, H. F. VanLandingham, and D. H. McCabe, "Modeling and estimation for tracking maneuvering targets," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-15, no. 3, pp. 448-456, May 1979.

[28] D. H. McCabe and R. L. Moose, "Passive source tracking using sonar time delay data," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-29, no. 3, pp. 614-617, June 1981.

[29] R. L. Moose and T. E. Dailey, "Adaptive underwater target tracking using passive multipath time-delay measurements," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-33, no. 4, pp. 777-787, Aug. 1985.

[30] R. L. Moose and P. M. Godiwala, "Passive depth tracking of underwater maneuvering targets," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-33, no. 4, pp. 1041-1044, Aug. 1985.

[31] R. L. Moose, M. K. Sistanizadeh, and G. Skagfjord, "Adaptive estimation for a system with unknown measurement bias," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-22, no. 6, pp. 732-738, Nov. 1986.

[32] R. L. Moose, "Passive range estimation of an underwater maneuvering target," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-35, no. 3, pp. 274-285, Mar. 1987.

[33] J. K. Tugnait and A. H. Haddad, "A detection-estimation scheme for state estimation in switching environments," *Automatica*, vol. 15, pp. 477-481, 1979.

[34] J. K. Tugnait and A. H. Haddad, "Adaptive estimation in linear systems with unknown Markovian noise statistics," *IEEE Trans. Inform. Theory*, vol. IT-26, no. 1, pp. 66-78, Jan. 1980.

[35] H. Akash and H. Kumamoto, "Random sampling approach to state estimation in switching environments," *Automatica*, vol. 13, pp. 429-434, 1977.

[36] P. K. S. Tam and J. B. Moore, "A Gaussian sum approach to phase and frequency estimation," *IEEE Trans. Comm.*, vol. COM-25, no. 9, pp. 935-942, Sept. 1977.

[37] S. P. Dmitriev and L. I. Shimelevich, "Generalized Kalman filter with repeated linearization and its application in geophysical field navigation," *Automation & Remote Control*, vol. 39, no. 4, pt. 1, pp.505-509, Apr. 1978.

[38] H. R. Sirisena and E. P. M. Brown, "Representation of non-Gaussian probability distributions in stochastic load-flow studies by the method of Gaussian sum approximations," *IEE Proc.*, vol. 130, pt. C, no. 4, pp. 165-171, July 1983.

[39] T. Namera and A. R. Stubberud, "Gaussian sum approximation for non-linear fixed-point prediction," *Int. J. Control*, vol. 38, no. 5, pp. 1047-1053, 1983.

[40] M. Gauvrit, "Bayesian adaptive filter for tracking with measurements of uncertain origin," *Automatica*, vol. 20, no. 2, pp. 217-224, 1984.

[41] M. Tanaka and T. Katayama, "Robust Kalman filter for linear discrete-time system with Gaussian sum noises," *Int. J. Systems Sci.*, vol. 18, no. 9, pp. 1721-1731, 1987.

[42] G. Kitagawa, "Non-Gaussian seasonal adjustment," *Computers Math. Applic.*, vol. 18, no. 6/7, pp. 503-514, 1989.

[43] C. J. Masreliez, "Approximate non-Gaussian filtering with linear state and observation relations," *IEEE Trans. Automat. Contr.*, vol. AC-20, no. 1, pp. 107-110, Feb. 1975.

[44] C. J. Masreliez and R. D. Martin, "Robust Bayesian estimation for the linear model and robustifying the Kalman filter," *IEEE Trans. Automat. Contr.*, vol. AC-22, no. 3, June 1977.

[45] C. Tsai and L. Kurz, "An adaptive robustizing approach to Kalman filtering," *Automatica*, vol. 19, no. 3, pp. 279-288, 1983.

[46] P. S. Maybeck and R. I. Suizu, "Adaptive tracker field-of-view variation via multiple model filtering," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-21, no. 4, pp. 529-538, July 1985.

[47] D. M. Tobin and P. S. Maybeck, "Enhancements to a multiple model adaptive estimator/image-tracker," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-24, no. 4, pp. 417-426, July 1988.

[48] H. A. P. Blom and Y. Bar-Shalom, "The interacting multiple model algorithm for systems with Markovian switching coefficients," *IEEE Trans. Automat. Contr.*, vol. 33, no. 8, Aug. 1988.

[49] Y. Bar-Shalom, K. C. Chang, and H. A. P. Blom, "Tracking a maneuvering target using input estimation versus the interacting multiple model algorithm," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-25, no. 2, pp. 296-300, Mar. 1989.

[50] E. Emre and J. Seo, "A unifying approach to multitarget tracking," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-25, no. 4, pp. 520-527, July 1989.

[51] R. L. Kirlin and A. Moghaddamjoo, "Robust adaptive Kalman filtering for systems with unknown step inputs and non-Gaussian measurement errors," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 34, no. 2, pp. 252-263, Apr. 1986.

[52] A. Moghaddamjoo and R. L. Kirlin, "Robust adaptive Kalman filtering with unknown inputs," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, no. 8, pp. 1166-1175, Aug. 1989.

[53] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, 3rd ed. New York: McGraw-Hill, Inc., 1991.

[54] J. W. Nilsson, *Electric Circuits*, 3rd ed. New York: Addison-Wesley, Inc., 1990.

[55] J. K. Tugnait, "Comments on 'State estimation for discrete systems with switching parameters'," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-15, no. 3, p. 464, May 1979.

[56] J. Raisch, "Comments on 'A multi-model adaptive predictor for stochastic processes with Markov switching parameters'," *Int. J. Control*, vol. 45, no. 4, pp. 1489-1490, 1987.

[57] R. G. Brown, *Introduction to Random Signal Analysis and Kalman Filtering*, 2nd ed. New York: John Wiley and Sons, Inc., 1991.

[58] J. S. Bendat and A. G. Piersol, *Engineering Applications of Correlation and Spectral Analysis.* New York: John Wiley and Sons, Inc., 1980.

[59] H. F. VanLandingham, *Introduction to Digital Control Systems.* New York: Macmillan Publishing Co., 1985.

[60] J. H. Mendel, *Lessons in Digital Estimation Theory.* Englewood Cliffs, NJ: Prentice-Hall, Inc., 1987.

[61] H. Stark and J. W. Woods, *Probability, Random Processes, and Estimation Theory for Engineers.* Englewood Cliffs, NJ: Prentice-Hall, Inc., 1986.

VITA

Mauro J. Caputi was born in Flushing, New York, on July 15, 1959. He received the B.E.E. degree in electrical engineering from Manhattan College, Riverdale, New York, in 1981. He received the M.S. degree in electrical engineering from Virginia Polytechnic Institute & State University, Blacksburg, Virginia, in 1984. He received the Ph.D. degree in electrical engineering from Virginia Polytechnic Institute & State University, Blacksburg, Virginia, in May, 1991.

Dr. Caputi has spent the last seven years working as an Instructor in electrical engineering at Virginia Polytechnic Institute, and conducting research in the area of applied estimation theory. He is a member of the American Society for Engineering Education, Eta Kappa Nu, and Tau Beta Pi. In 1986 he received the Sporn Award from Virginia Polytechnic Institute in recognition of excellence in teaching of engineering subjects. In 1987 he received a Certificate of Teaching Excellence from Virginia Polytechnic Institute, one of four awarded annually to the College of Engineering.

Mauro J. Caputi