A Fuzzy-Kalman Filtering Strategy for State Estimation

A Thesis Submitted to the College of

Graduate Studies and Research

in Partial Fulfillment of the Requirements

for the Degree of Master of Science

in the Department of Mechanical Engineering

University of Saskatchewan

Saskatoon

By

Lee-Ryeok Han

June 2004

Permission to Use

In presenting this thesis in partial fulfillment of the requirements for a Graduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis.

Requests for permission to copy or to make other use of material in this thesis in whole or part should be addressed to:

Head of the Department of Mechanical Engineering,

University of Saskatchewan,

Saskatoon, Saskatchewan S7N 5A9.

Abstract

This thesis considers the combination of Fuzzy logic and Kalman Filtering that have traditionally been considered to be radically different. The former is considered heuristic and the latter statistical. In this thesis a philosophical justification for their combination is presented. Kalman Filtering is revised to enable the incorporation of fuzzy logic in its formulation. This formulation is subsequently referred to as the Revised-Kalman Filter. Heuristic membership functions are then used in the Revised-Kalman Filter to substitute for the system and measurement covariance matrices to form a fuzzy rendition of the Kalman Filter. The Fuzzy Kalman Filter formulation is further revised according to a concept referred to as the "Parallel Distributed Compensation" to allow for further heuristic adjustment of the corrective gain. This formulation is referred to as the Parallel Distributed Compensated-Fuzzy Kalman Filter.

Simulated implementations of the above filters reveal that a tuned Kalman Filter provides the best performance. However, if conditions change, the Kalman filter's performance degrades and a better performance is obtained from the two versions of the Fuzzy Kalman Filters.

Acknowledgments

The author would like to sincerely thank his supervisor Dr. Saeid Habibi for his valuable guidance and supervision. I cannot avoid saying that his direct but warm advice and comments have always been a significant help for me. It has been my pleasure to work under his supervision.

The author would like to extend his hearty thanks to his sisters and his aunt the late D. H. Han as well. She was a mother and a grandmother to him.

The author also extends his great thanks to Dr. Jeeyup Han, Dr. Seok-Bum Ko, Dr. Oon-Doo Baik, Pastor Steven N. Fish and Pastor Stanley S. Sur for their fully warm and prudent advices and assistances.

The author would like to express his honest thanks to his church members and his many friends in Saskatoon. They have always been willing to help the author.

Dedicated to my sisters and my late aunt.

Table of Contents

Nomenclature

Symbol	Definition	Value/Unit
A	System matrix	
$\widetilde{A}$	Fuzzy input	
B	Input matrix	
$\widetilde{B}$	Fuzzy input	
$\widetilde{B}'$	Fuzzy set	
$\widetilde{B}_{\alpha}$	$\alpha-$cut fuzzy set	
C	Matrix for factoring	
$\widetilde{C}$	Fuzzy input	
E	Expectation	
H	Output matrix	
$\widetilde{H}$	Output matrix beyond the assumed measurement noise covariance	
I	Identity matrix	
K	Kalman filter gain	
$\widetilde{K}$	Revised-Kalman filter gain	
P	The a posteriori error covariance matrix	
P_0	The initial value of the a posteriori error covariance matrix	
P^-	The a priori error covariance matrix	
$\widetilde{P}$	Projectile error covariance matrix	
$\widetilde{P}_0$	The initial projectile error covariance matrix	

$\hat{P}$	The measurement error covariance matrix
Q	The system noise covariance matrix
R	The measurement noise covariance matrix
S	Domain of input
T	Transpose matrix
U	Input
X	State vector
$\hat{X}$	The a posteriori state estimate
$\hat{X}_0$	The initial a posteriori state estimate
$\hat{X}^-$	The a priori state estimate
$\widetilde{X}$	The projectile estimate
$\widetilde{X}_0$	The initial projectile estimate
$\widetilde{X}^-$	The pre-processor estimate
$\widetilde{Z}$	Linguistic variable
e	The a posteriori error vector
e_{x_k}	The projectile error vector
e_{z_k}	The measurement error vector
e^-	The a priori error vector
i	Number of fuzzy rules
j	Number of fuzzy rules for a residual
k	Iteration step
n	Integer indicating number of elements

s	Input element
t	Output element
v	Measurement noise
w	System noise
x	State variable
$\tilde{x}$	Fuzzy input
y	Crisp output
z	Measurement
Δ	Sampling time
Φ	Probability density function of the measurement error
Ω	Fuzzy function of a residual
α	$\alpha-$cut
δ	Linguistic variable
η	Noise in output matrix
κ	Fuzzy-Kalman filter gain
λ	Membership function of submodel
μ	Membership function
ξ	Crisp output
ψ	Crisp output of submodel
ω	Noise in system matrix

List of Tables

List of Figures

Chapter 1

Introduction

Mr. Kim is a well known Korean who has won more than five investment tournaments, trading real stocks since 2001. His success was influenced by an interesting experience he had while staying in Japan as an exchange student. There, he had a part-time job at a pachinko shop (Japanese casino) where he recorded the profitability of each pachinko machine for fun. After a while, he found that some regular rhythms existed in each machine. He quit his job, tried gambling, and won around 36,000 Canadian dollars in 2 weeks. Then, he used a similar strategy for investing in stock markets. Quoting him directly, he used "Scientific analysis and intuition to guarantee success in investment", [1].

Another example was Mr. Jang at Daishin Securities Co. who received an extra bonus of around 3.6 million Canadian dollars from his company. He did not have a university degree. His colleagues attributed his success to his ability to analyze data while having a "special intuition" to interpret them, investing in the future's market, [2]. It is apparent from the above examples that sometimes expert experience can outperform, if not complement, scientific methods, [2].

The aim of fuzzy theory is to combine scientific rigour with expert intuition. Fuzzy control using fuzzy theory has been studied and applied to many industrial problems since Lofti A. Zadeh introduced this concept in the 1960s, [3]. In contrast, there are very few publications pertaining to fuzzy estimation. This may be partly explained by the

success of the Kalman filter and the historical feud between the proponents of these two methods. Quoting from Kalman "...Zadeh's proposal could be severely, ferociously, even brutally criticized from a technical point of view. This would be out of place here. But a blunt question remains: Is Zadeh presenting important ideas or is he indulging in wishful thinking?...", [4], and rebutted by Zadeh "…just waived it off and said the Kalman filter was " too Gaussian." That means it depends too much on a bell curve ", [5].

Kalman filtering has been used in navigation, guidance, estimation and other control-oriented processes, in a discrete or a continuous form. It is a rigorous optimal state estimation strategy that is applied to stochastic signals. In contrast, fuzzy theory uses an intuitive experience based-approach for problems that are too difficult to model mathematically. The question is which is better for state estimation? Would the combined scientific rigour of an optimal stochastic method and the intuitive but nonetheless rigorous fuzzy theory provide a more effective estimation method? Our hypothesis is that substantive benefit is gained when the two approaches are combined. This combination allows a more rigorous capture of a priori information through Fuzzy logic when formulating the optimal stochastic estimation method of Kalman filtering.

1.1 Objective of the Thesis

The objective of the thesis is to develop a Fuzzy-Kalman filter that would combine the benefits of the two concepts. A significant limitation of the Kalman filter pertains to assumptions made concerning to its internal model. These assumptions are circumvented by using Fuzzy logic in the proposed combined Fuzzy-Kalman formulation.

1.2 Competition or Cooperation

In this chapter, the rational for combining the Kalman Filter and the Fuzzy logic is hypothetically discussed and explored. Despite Kalman's and Zadeh's reservations in crediting and acknowledging their respective theories that are based on statistics and intuition, real world problems require both viewpoints and perspectives. Even though initially it may seem that they do not need one another, careful reflection reveals that their combination is beneficial. It is easy to predict snow in winter. But it is difficult to say which day it will occur. That can be forecasted with statistics and information pertaining to humidity, temperature, etc. That information should be obtained prior to forecasting. Although it might be thought that if statistics is used, experience is not needed, application of our proposed hypothesis to the above example assumes that weather is so difficult to model that its forecasting requires both intuitive experience and statistics.

The first successful application of fuzzy theory was developed by E.H. Mamdani [3, 6, 7, 8] and involved the control of a steam engine. Since then, fuzzy theory has been applied to transportation systems such as subway trains, process control, image processing, and appliances, [3]. Hitachi developed a fuzzy control system for the automatic operation of a subway trains. This fuzzy system adopted train drivers' experiences in controlling the velocity, acceleration and braking systems. The rules of logic were derived from interviews conducted with train drivers. The rules were obtained in view of improving the safety, the convenience, the energy consumption, the travel time, and the precision of stoppage at the subway platform. Their implementation

resulted in fewer applications of brakes, which led to lower energy consumption and improved ride quality for passengers, [3].

The concept of vagueness in fuzzy theory should not be confused with probability theory employed by the Kalman filter. Fuzzy theory treats the solution with plausibility and Kalman filtering does it with probability. It can be said that in fuzzy theory the challenge of understanding the uncertain events is made subjectively. Fuzzy theory is subjective because it tries to depict the uncertainty of events by a vague definition. Kalman filtering is objective as it is related to the degree of occurrence of phenomena. John F. Nash, Jr. showed that two differing concepts such as these benefit from cooperation by using his famous 'Game Theory' in mathematics, [9]. It can be shown with Prisoners' dilemma in Table 1.1, [10].

Suspects		A		
	Responses	Silence	Confession	Sentence
B	Silence	1 year 1 year	Released 10 years	A B
	Confession	10 years Released	5 years 5 years	A B

Table 1.1 Prisoners' Dilemma

Assume that two people are charged with theft. If one confessed his fault and the other denied it, the one who confessed would be freed but the other would be jailed for 10 years. If both denied their charges, both would be jailed for 1 year. And if they both confessed, they would be both jailed for 5 years. Note that the best outcome in Table 1.1

occurs when the accused trust one another and remain silent. That is Nash's equilibrium. Game theory shows that mutual cooperation is profitable as competition could sometimes worsen result. In this thesis, elements from the Kalman Filter concept and the Fuzzy theory are combined to "cooperate" in order to improve the quality of state estimation in the presence of modeling uncertainties.

1.3 Inductive Inference and Deductive Inference

Professor Park of the Department of Aerospace Engineering at Inha University, Korea, wrote in his introduction to the book of Prouty's stresses, on the importance of deductive thinking, [11]. Deductive inference logically derives the last proposition from the assumption that previous propositions are true. In other words, the last proposition is deduced by an assumed logical relationship with previous propositions. But the truth of the premises does not guarantee the truth of conclusions.

1.3.1 Deductive Inference and Fuzzy Theory

An example of deductive inference is as follows:

I. Hypothetical proposition:

   Canadian (Molson beer) is Canada's most favourite beer.

II. Categorical proposition:

   I drink Canadian (Molson beer).

III. Conclusion:

   I am a Canadian.

The truths of the assumed conclusion are not examined here. Although the categorical proposition is true, the conclusion could be false. This is non sequitur (illogical conclusion) as the premises may not be relevant, [12]. It means that the import of conclusions cannot be drawn from the premise because the truth of premises cannot confirm whether the conclusion is true or false positively.

The following example illustrates the basis of the Fuzzy logic application to estimation.

- Fuzzy Theory (Deduction)
  I. Hypothetical proposition: System model
  II. Categorical proposition: Rules of logic applying to the model used for estimation
  III. Conclusion: State estimate

In the above example, the hypothetical proposition is assumed to be true. The categorical proposition is always thought to be true as it relates to a measured quantity. But the conclusion can be true or false. This is non sequitur and means the logic of relationships are inadequate. For example, if as a result of the application of rules of logic the system model becomes unobservable, then the conclusion is no longer valid.

1.3.2 Inductive Inference and Kalman Filtering

When a conclusion is derived from a proposition, inductive inference is used with the assumption that the proposition is reasonable to induce the corresponding conclusion. Nonetheless, a proposition is true; the conclusion is probably true but not absolutely. Typically applying inductive inference, a proposition includes the element of a conclusion, [12]. Consider the following inductive inference:

I. Hypothetical proposition: Most Canadians drink Canadian (Molson beer).
II. Categorical proposition: I am a Canadian.
III. Conclusion: I like Canadian (Molson beer).

Recursive inference is inductive inference, [13, 14]. As such, the Kalman filter can be thought of as being based on inductive inference, with a recursive function using feedback correction. The Kalman filter is an iterative methodology that includes data collection, deduction, mathematical construction from a model, and feedback correction to the deduction, [15, 16]. However, Inductive inference starts from premises. Empirical investigation is needed to discriminate the truth of premises. The truth of the last proposition depends on these premises that often involve assumptions and not facts. If the proposition is true, then the conclusion is probably but not absolutely true. This assumption can be nonetheless false under certain circumstances at a low degree of probability as follows:

- Kalman filtering (Induction)

  I. Hypothetical proposition: System model

  II. Categorical proposition: Prediction and feedback correction

  III. Conclusion: State estimate

Hypothetical proposition is true if the model remains within a confined operating region. Otherwise, for a possible lack of observability, the categorical proposition could be false. Therefore the conclusion is true or false. Therefore, probability (categorical proposition) does not guarantee the right output. As a simple example consider the presence of an unmodeled drift in the measurement signal. This drift is not detected or compensated for by the Kalman filter resulting in an inaccurate prediction or a false conclusion. In this case, Gaussian distribution does not express all the natural phenomena present in the environment.

Scientific theories use mathematical constructs to model natural phenomena. However, such physical phenomena are often too complex and need to be represented in terms of mathematics that would include both deductive and inductive reasoning. History has shown that, in far too many cases, reliance on operator experience versus rules of logic has been the right approach. An empirical investigation for constructing the rules of logic is therefore necessary. This is the reason why fuzzy theory needs expert experience to improve the reliability or the truth of conclusions. Nonetheless, the problem is how comprehensive can fuzzy logic rules be and who has the proper expertise to define them.

1.4 Combination of Kalman Filtering and Fuzzy Logic (Inductive Logic)

Kalman filtering and fuzzy theory seem heterogeneous in nature. This heterogeneity might be the root cause of the challenge to have them cooperate. The Kalman filtering and the fuzzy logic have shown their respective superiority in different applications and these need to be identified for their combination. Zadeh developed fuzzy theory to express ambiguity in a quantitative language. Therefore fuzzy theory relates to intelligence in systems. Fuzzy theory exhibits a dominant ability to imitate human being's subjective decision. With fuzzy theory, smart systems can be made to adapt to their environment.

The objective of the fuzzy theory is to provide a more effective representation of natural phenomena. Fuzzy theory maps an abstract universe with membership functions; accordingly, the semantic of a membership function is based on the concept of vagueness. These membership functions are, in effect, simple assumptions that isolate the features of interest in natural phenomena. On the other hand, complex abstractions may be achieved by combining such simple functions with other forms of mathematical knowledge of the system, [16]. It is in this context that the fusion or marriage of the Kalman filtering concept and fuzzy logic can occur. To combine the two theories, the concept of marriage is used as follows.

- Step 1. - Understanding: "I have to understand myself and my bride (or bridegroom) for marriage."
  - Fuzzy theory: deduction must have a resonable relationship between premise (model) and conclusion (estimate).

- Kalman filter: statistical characterization can provide an effective mean of capturing a priori information to provide modeling of an uncertain complex nonlinear system.

- Step 2. - Yield: "I admit that my bride (or bridegroom) has a forte and virtuous personality. I am ready to accept my partner's strengths."
  - Fuzzy theory: The inductive inference of fuzzy can benefit from supplementing it with the probability theory.
  - Kalman filter: The probability is more reliable if related to the perceptual nonlinear relationship of cause and effect.

- Step 3. - Resemblance: "I start to look and behave like my wife."
  - Fuzzy theory: Fuzzy takes after the mathematical form of Kalman filtering.
  - Kalman filter: According to Gödel's incompleteness theorem[*1], [16], relations defined by rules of deduction are not, in general, complete, [17].

This combination of the fuzzy logic and the Kalman filter is considered in this thesis. An outline of the thesis is as follows.

(1 "Gödel's incompleteness theorem makes it clear that mathematics cannot achieve the goal of completeness."

1.5 An Outline of the Thesis

An introduction to fuzzy logic is provided in chapter two. This includes a consideration of the fuzzy set, the membership function, the fuzzy inference, and a critical mapping proposed by Takagi and Sugeno, [6]. Chapter 3 provides an introduction to the Kalman filter. The derivation of this filter and its associated estimation process are discussed. A revision to the Kalman filter that would allow its fusion to fuzzy logic is presented in Chapter 4. The advantages and the disadvantages of this formulation are discussed. A combined Fuzzy-Kalman filter is proposed and presented in Chapter 5. In Chapter 6, the Fuzzy-Kalman filter is applied in simulation to an example system. The results of this simulation are presented and comparatively discussed. The concluding remarks are contained in Chapter 7.

Chapter 2

Introduction to Fuzzy Theory

Fuzzy theory was proposed by Lofti Zadeh, [3, 5, 6, 7, 8]. Since its introduction and further to its application to control problems as demonstrated by Mamdani, [3, 6, 7, 8], it has gained considerable acceptance within the scientific community. Fuzzy theory can be used for constructing nonlinear relationships with heuristic information. As such, it can capture operator experience or knowledge even though that may not be initially in a mathematical form. Fuzzy theory has had a profound impact on control and intelligent systems. In this chapter a brief introduction to fuzzy theory is provided. The text books by Passino and Yurkovich, [6], and Tsoukalas and Uhrig, [7] provide an extensive and more detailed description of this concept.

2.1 Fuzzy and Crisp Sets

In the fuzzy context, crisp sets are defined as having a collection of definite and distinguishable elements.

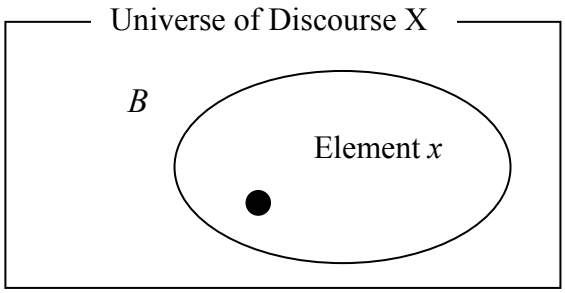


Figure 2.1 Abstract Crisp Set

Consider the crisp set B with elements $B = \{4,5,6\}$. Let the domain of B be X, containing all values that the elements of B can have. In fuzzy terminology, this domain is referred to as the universe of discourse as shown in Figure 2.1. For example, if X is the set of all positive integer numbers in the range (0 to 10), and if the requirement is that the elements of B be integers in the range of 0 and 10, then X is referred as the universe of discourse. Associated with a set are special functions that can be used to establish its elements referred to as membership functions. The crisp membership function $\mu_B(x)$ is used to define membership for elements x of X that are members of B. As such if an element $x \in X$ is also a member of the crisp set B, then the membership function $\mu_B(x)$ is equal to 1. Otherwise, $\mu_B(x)$ is equal to 0. The membership function of a crisp set can only be either '1' or '0'. For example, let the membership of B be restricted to integer numbers in the range of 4 to 6. Then:

$\mu_B(x) = 1$ iff $4 \le x \le 6, x \in X$

$\mu_B(x) = 0$ iff $x < 4$ or $x > 6, x \in X$

If x is more than or equal to 4 and less than or equal to 6, it is considered to be a member of the crisp set B, otherwise not, as shown in Figure 2.2.

$$\mu_B(x) = \begin{cases} 1 \text{ iff } x \in B \\ 0 \text{ iff } x \notin B \end{cases}$$

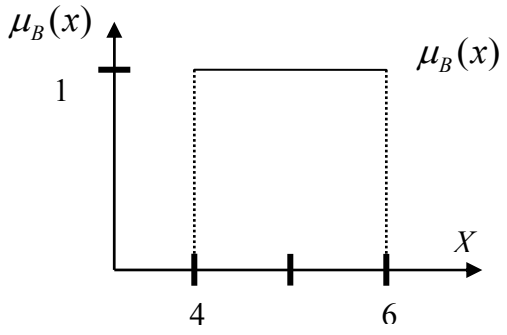


Figure 2.2 Classical Crisp Set, [18]

In fuzzy theory, it is possible to also define sets for which the boundary is vaguely defined as shown in Figure 2.3, [6, 18]. Fuzzy sets are denoted here by using the symbol $\widetilde{B}$. The membership function of a fuzzy set allows values between '0' and '1'. The membership function of a fuzzy set is denoted as $\mu_{\widetilde{B}}(x)$ and provides a value indicating the possibility or the degree of membership.

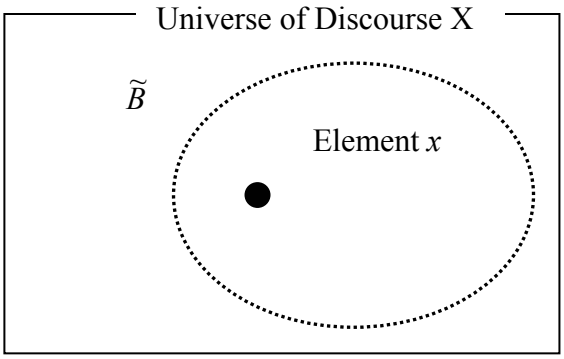


Figure 2.3 Abstract Fuzzy Set

Linguistic variables are the cornerstone of fuzzy logic. They can be used to categorize fuzzy sets. For example, let the fuzzy set $\widetilde{B}$ be used to describe the notion 'Average'. Then associated with the set $\widetilde{B}$ is a membership function. Consider $S = \{1,4,5,7,2\}$ as the score of five pupils. Fuzzy logic can be used to define a set $\widetilde{B}$ associated with the subjective notion 'Average' with each pupil's degree of membership quantified by $\widetilde{B}$'s membership function $\mu_{\widetilde{B}}$. As such to represent a fuzzy set, the use of what is termed as a singleton is required. Singleton is defined as the pair $\widetilde{B} = \{(s, \mu_{\widetilde{B}})\}$ for, $s \in S$, [6, 18].

As an example, consider a fuzzy set $\widetilde{B}$ such that

$$\widetilde{B} = \left\{\left\langle 1, \mu_{\widetilde{B}}(1)\right\rangle, \left\langle 4, \mu_{\widetilde{B}}(4)\right\rangle, \left\langle 5, \mu_{\widetilde{B}}(5)\right\rangle, \left\langle 7, \mu_{\widetilde{B}}(7)\right\rangle, \left\langle 2, \mu_{\widetilde{B}}(2)\right\rangle\right\} \tag{2.1}$$

The membership function of a fuzzy set not only determines, but also qualifies the degree of membership as shown for the notion 'Average' in Figure 2.4. The membership function of a fuzzy set $\mu_{\widetilde{B}}(x)$ can therefore have a value between '0' to '1'. The elements of the set *S* have a degree of membership in $\widetilde{B}$ and as such the boundary of $\widetilde{B}$ is not clearly defined.

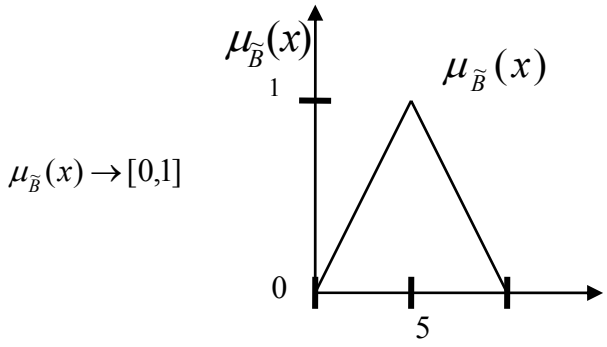


Figure 2.4 Fuzzy Set with the Membership Function for 'Average'

Further to Figure 2.4 and equation (2.1), the fuzzy set $\widetilde{B}$ with evaluated singletons to depict 'Average' is:

$$\widetilde{B} = \{(1,0.2),(4,0.8),(5,1.0),(7,0.6),(2,0.4)\} \tag{2.2}$$

The degree of membership of the crisp set *S* to the fuzzy notion 'Average' can also be quantified by using the membership function $\mu_{\widetilde{B}}(x)$ such that:

$$\widetilde{B}' = \left\{\mu_{\widetilde{B}}(s_1)/x_1 + \mu_{\widetilde{B}}(s_2)/x_2 + \ldots \mu_{\widetilde{B}}(s_n)/x_n\right\} \tag{2.3}$$

If an element is '0', then that element could be ignored. If this mathematical notion is adopted, $\widetilde{B}'$ can be expressed in a simplified form as, [18]:

$$\widetilde{B}' = \sum_{i=1}^{n} \mu_{\widetilde{B}}(s_i) / s_i \tag{2.4}$$

In the previous example where $\widetilde{B}$ is subjectively associated with 'Average' and when evaluated for S is: $\widetilde{B} = \{(1,0.2),(4,0.8),(5,1.0),(7,0.6),(2,0.4)\}$

The degree of 'averageness' of $\widetilde{B}$ can be indicated by the sum of its evaluated membership functions, by scalar cardinality such that, [3]:

$$\left|\widetilde{B}\right| = \sum_{s \in S} \mu_{\widetilde{B}}(s) \tag{2.5}$$

Further to equation (2.5), then the degree of membership of S (e.g. group of pupils' grade) to the notion 'Average' is obtained as:

$$\left|\widetilde{B}\right| = 0.2 + 0.8 + 1.0 + 0.6 + 0.4 = 3 \tag{2.6}$$

It is possible to create multiple membership functions, such as in this example, consider the notions of 'Low', 'Average', and 'High'. Let the corresponding membership functions be $\mu_{\widetilde{C}}(s)$, $\mu_{\widetilde{B}}(s)$, and $\mu_{\widetilde{A}}(s)$ as shown in Figure 2.5.

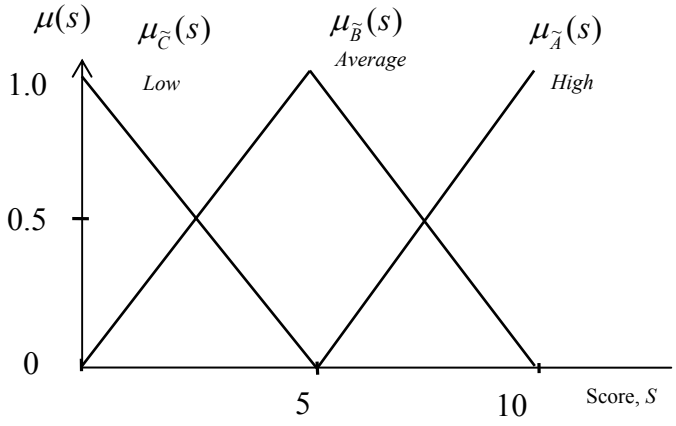


Figure 2.5 Fuzzy Set and Membership Functions

The corresponding degree of membership of function S to 'Low' and 'High' are:

$\left|\widetilde{C}\right| = 0.8 + 0.2 + 0.0 + 0.0 + 0.6 = 1.4$ and $\left|\widetilde{A}\right| = 0.0 + 0.0 + 0.0 + 0.4 + 0.0 = 0.4$

The cardinality associated with 'Low', 'Average' and 'High' are respectively 1.4, 3 and 0.4. The analysis using fuzzy logic indicates that the pupils' performance given by grades on the set S can be mainly categorized as average.

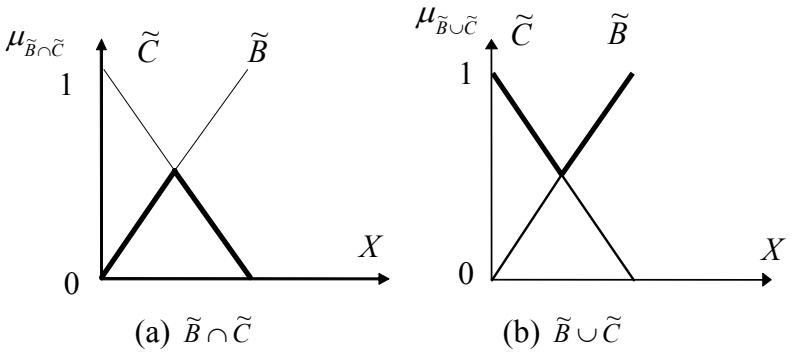
2.2 Fuzzy Operations

Fuzzy logic has a number of common operators that are briefly described in this section. The intersection of fuzzy sets $\widetilde{A}$ and $\widetilde{B}$ is defined as:

$$\mu_{\widetilde{A}\cap\widetilde{B}}(x) = \min\left\{\mu_{\widetilde{A}}(x), \mu_{\widetilde{B}}(x)\right\}, \forall x \in X \tag{2.7}$$

Fuzzy intersection is expressed with $\widetilde{A} \cap \widetilde{B}$ and takes the smallest membership value of the subset of elements that belong to fuzzy sets $\widetilde{A}$ and $\widetilde{B}$. Fuzzy union is denoted by $\widetilde{A} \cup \widetilde{B}$ and takes the largest membership value of the set containing all elements of $\widetilde{A}$ or $\widetilde{B}$. Union of fuzzy sets $\widetilde{A}$ and $\widetilde{B}$ is defined as:

$$\mu_{\widetilde{A}\cup\widetilde{B}}(x) = \max\left\{\mu_{\widetilde{A}}(x), \mu_{\widetilde{B}}(x)\right\}, \forall x \in X \tag{2.8}$$



(a) $\widetilde{B} \cap \widetilde{C}$ (b) $\widetilde{B} \cup \widetilde{C}$

Figure 2.6 Fuzzy Set Operations

In the previous example that used a linguistic approach to describing a student's score as 'Low', 'Average', and 'High', where $S = \{\text{Student's score}\} = \{1,4,5,7,2\}$, then further to the membership functions of Figure 2.6, the intersection $\cap$ and union $\cup$ operators in the fuzzy context are illustrated in Figure 2.6.

2.3 α – Cut and Resolution Principle

To represent the fuzziness of data, it is useful to set a limit on the degree of membership. $\alpha-$cut of a fuzzy set is the set on the universe of discourse where the values of the membership functions are greater than or equal to α . $\alpha-$cut denoted as $\widetilde{B}_\alpha$ is a crisp set and can be used to screen elements. For example, let:

$$\widetilde{B}_\alpha = \left\{ s \in S \middle| \mu_{\widetilde{B}}(s) \geq \alpha \right\}, \quad 0 \leq \alpha \leq 1$$

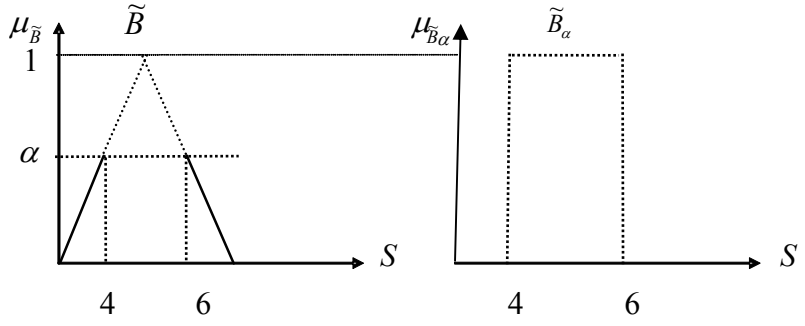


Figure 2.7 A Fuzzy Set $\widetilde{B}$ and $\alpha-$cut

Then graphically $\widetilde{B}_\alpha$ can be represented as in Figure 2.7. In this example, where $S = \{1,4,5,7,2\}$, $\widetilde{B}_\alpha$ is equal to the crisp set $\{4,5\}$. An approximate fuzzy set can be represented by the unions of $\alpha-$cuts, such that:

$$\widetilde{B} = \bigcup_{\alpha\in[0,1]} \alpha\widetilde{B}_{\alpha} \text{ or } \mu_{\widetilde{B}}(S) = \bigcup_{0<\alpha\le 1} \alpha\mu_{\widetilde{B}\alpha}(s) \tag{2.9}$$

Equation (2.9) represents the Resolution Principle which means that fuzzy sets can be decomposed, [8]. In Figure 2.8, $\mu_{\widetilde{B}}(s)$ is shown as a union of $\alpha-$cut functions.

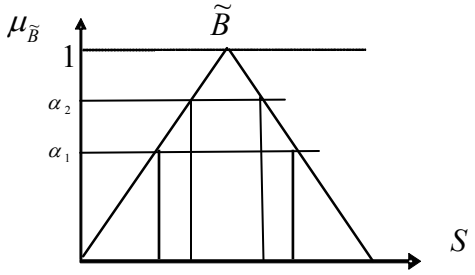
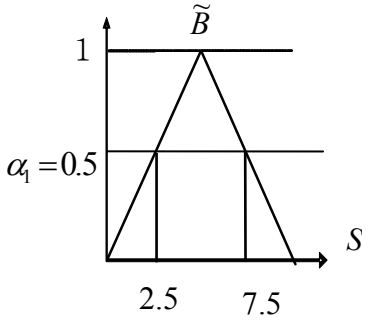
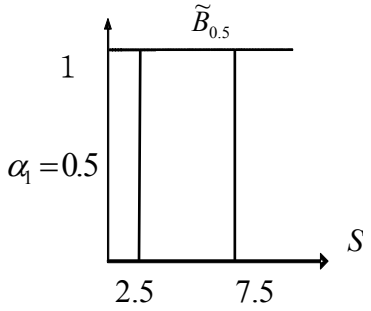


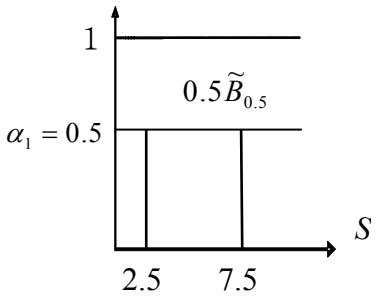
Figure 2.8 A Fuzzy Set $\widetilde{B}$ and $\alpha-$cuts



(a) Membership function $\widetilde{B}$



(b) Membership function $\widetilde{B}_{0.5}$



(c) Function $0.5\widetilde{B}_{0.5}$

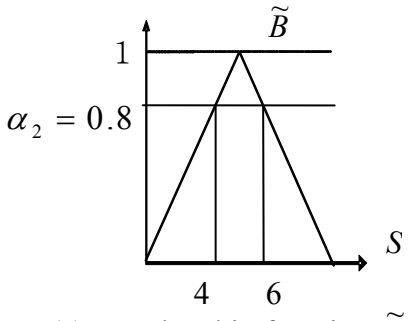
Figure 2.9 A Fuzzy Set $\widetilde{B}$ and 0.5-cut

For example, let Figure 2.8 stand for 'Average'. For $\alpha_1 = 0.5$ and $\alpha_2 = 0.8$, then

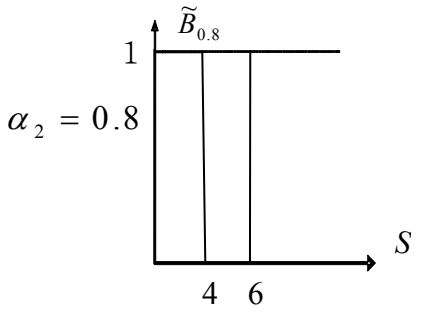
$\widetilde{B}_{0.5} = \{s \in S | \mu_{\widetilde{B}}(s) \geq 0.5\} = \{4,5\}$, and $\widetilde{B}_{0.8} = \{s \in S | \mu_{\widetilde{B}}(s) \geq 0.8\} = \{5\}$.

A crude approximation of the fuzzy set $\widetilde{B}$ can be obtained according to equation (2.9).

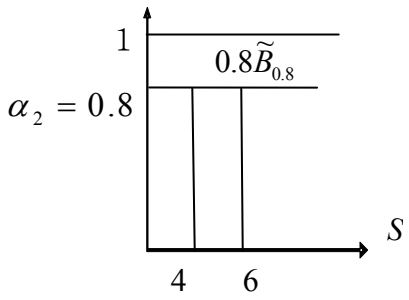
Here, $0.5\widetilde{B}_{0.5} = 0.5 \times \{\langle 4, \mu(4)\rangle, \langle 5, \mu(5)\rangle\}$ and $0.8\widetilde{B}_{0.8} = 0.8 \times \{\langle 5, \mu(5)\rangle\}$, then



(a) Membership function $\widetilde{B}$



(b) Membership function $\widetilde{B}_{0.8}$



(c) Function $0.8\widetilde{B}_{0.8}$

Figure 2.10 A Fuzzy Set $\widetilde{B}$ and 0.8-cut

As such, $\widetilde{B} = \bigcup_{\alpha \in [0.5,0.8]} \alpha \widetilde{B}_\alpha \approx 0.5\widetilde{B}_{0.5} + 0.8\widetilde{B}_{0.8}$ as depicted in Figure 2.11.

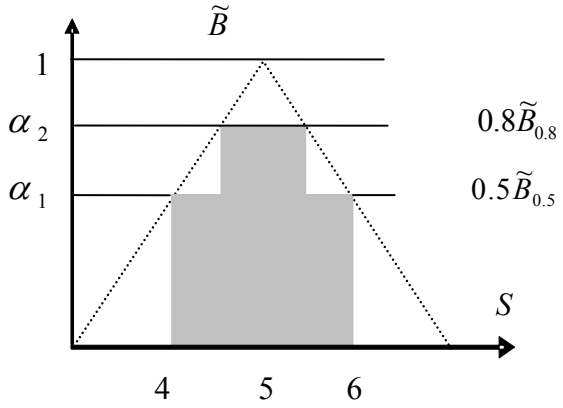


Figure 2.11 Approximate Fuzzy Set $\widetilde{B}$ by $\alpha-$cuts

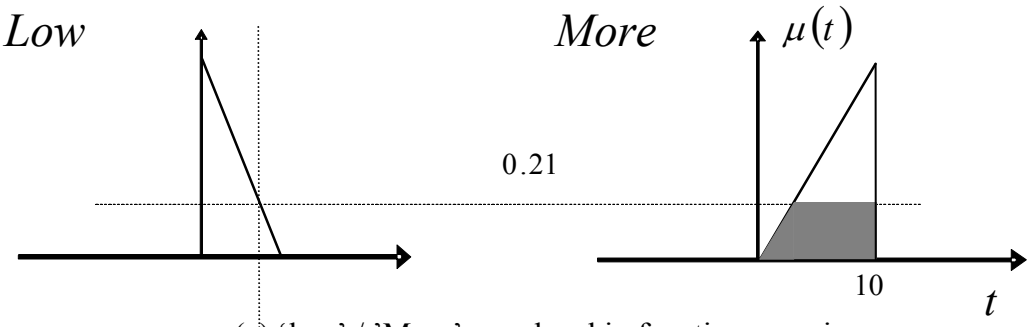
2.4 Fuzzy Inference

Fuzzy inference can be illustrated by using the above example and discussion. Fuzzy inference adopts the If ~Then~ rule. As such, if a condition is "GIVEN", then the conclusion is "INFERRED":

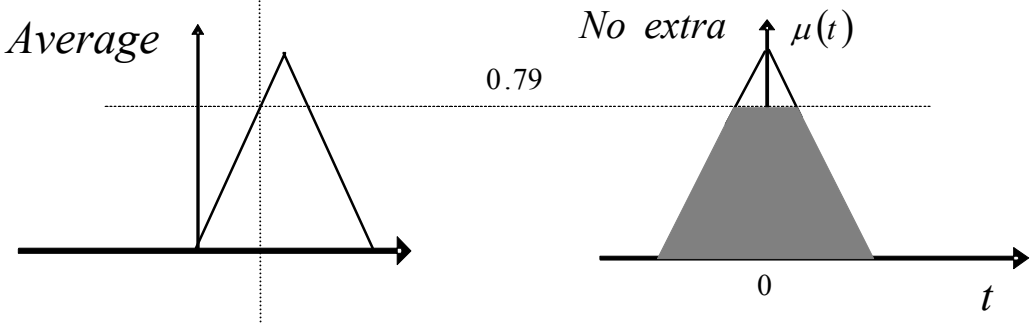
If <fuzzy proposition>, Then <fuzzy proposition>

For example: IF <the overall class performance is low (example of section 2.1)>, THEN <more assignments are given>.

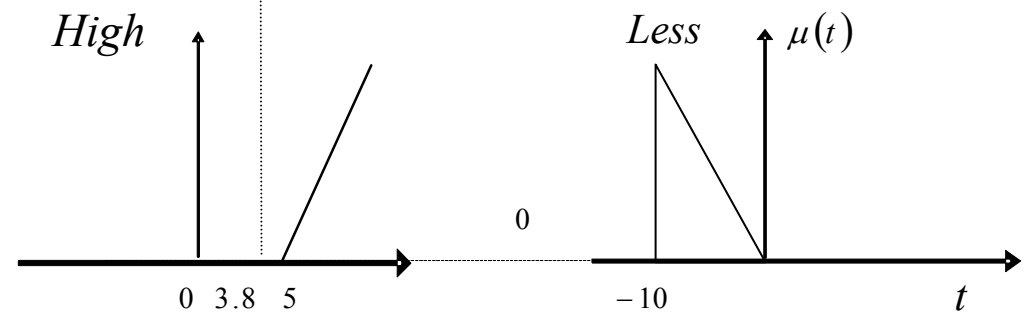
In the example of pupil score, the class performance is 0.4 on the notion 'High', 3 on 'Average' and '1.4' on 'Low'. Fuzzy logic has no problem with this inconsistency. It is now possible to use this information under a second fuzzy premise to establish a corrective defuzzified action. Further to this example, if the fuzzy rules for the corrective action pertain to the number of assignments that are given to the class in relation to their performance ("e.g. mid-term exam grade") then a simple set of fuzzy rules may be as depicted in Figure 2.12.



(a) ‘low’ / ’More’ membership function mapping

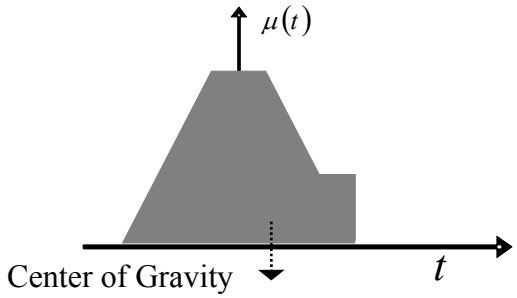


(b) ‘Average’ / ’No extra’ membership function mapping



(c) ‘High’ / ’Less’ membership function mapping

Additional assignment, T, $t \in T$



(d) Union of active defuzzification membership functions

Figure 2.12 Inferences

The associated fuzzy rules for our corrective action could be as given in Table 2.1.

PROPOSITION	CONCLUSION
Score is low	More assignment
Score is average	No extra assignment
Score is high	Reduce the number of assignment

Table 2.1 Fuzzy Rules for Assignment Scheduling

Applying these rules to our problem, the mapping between the 1st and 2nd precepts can be summarized as given in Table 2.2 that identifies membership functions associated with fuzzification and defuzzification.

SCORE		NUMBER OF ASSIGNMENTS
$\widetilde{A}$: 'High'	→	'Less'
$\widetilde{B}$: 'Average'	→	'No extra'
$\widetilde{C}$: 'Low'	→	'More'

Table 2.2 Correspondence of Fuzzy Membership Functions

Table 2.2 effectively provides the mapping relationship between membership functions used to evaluate scores (given in Figure 2.5) and the membership functions that relate to the resulting action (Figure 2.12). This mapping is illustrated in Figure 2.12 where T is the set pertaining to extra assignments. Evaluation of the corrective action from Figure 2.12 based on the performance of pupils as follows. Consider the pupil grade example with $S = \{1,4,5,7,2\}$. Let N be the number of pupils and $\widetilde{S}$ be the relative cardinality of S which can be indicated by the sum of division by N, such that:

$$\widetilde{S} = \sum_{s \in S} \frac{|s|}{N} \tag{2.10}$$

Further to equation (2.10), then the degree of membership of $\widetilde{S}$ (e.g. group of pupils' grade) is obtained as:

$$\widetilde{S} = \frac{1}{5} + \frac{4}{5} + \frac{5}{5} + \frac{7}{5} + \frac{2}{5} = 3.8 \tag{2.11}$$

From section 2.1, the overall class performance has membership associations with the notions 'Low', 'Average' and 'High' that are respectively 1.4 with 3 and 0.4.

Furthermore, the degree of membership of $\widetilde{S}$ is considered to the notion of 'Average'.

Further to Table 2.2, the fuzzy rules that relate the pupils score to the corrective action or number of assignments may be stated as:

IF $<\widetilde{S}$ is 'Low'> THEN $<T$ is 'More assignments would be given'>.

ALSO IF $<\widetilde{S}$ is 'Average'> THEN $<T$ is 'No extra assignment'>.

ALSO IF $<\widetilde{S}$ is 'High'> THEN $<T$ is Less assignments would be given'>.

From Table 2.2, the 'Low', 'Average' and 'High' correspond to 'More', 'No extra' and 'Less' respectively in additional assignment. For this example from Figure 2.5, 3.8 maps to 'More assignments' with a '0.21' membership function. Furthermore, this 'Average' corresponds to 'No extra assignments' with '0.79', and '0' of 'Less' assignments'. These relationships can be used for deciding the number of additional assignment by using "defuzzification".

A commonly used defuzzification method was proposed by Mamdani and involves calculating the union of areas of membership functions corresponding to the

required corrective action, with the overlap regions being multiply counted, [7]. This resulting shape is the shaded area in Figure 2.12 (d). The final defuzzified crisp corrective action "ξ" is obtained by calculating the center of gravity of the shaded area. A similar but simpler strategy to Mamdani's is to calculate the averaging scheme referred specifically as the Center of Gravity, [6]. Where *t* is the identifier for the membership function associated with the corrective action, the Center of Gravity method is specified as:

$$\xi = \frac{\sum_{i=1}^{3} t_i \times \mu(t_i)}{\sum_{i=1}^{3} \mu(t_i)} \tag{2.12}$$

For the example pertaining to student assignments and the relative cardinality, ξ is obtained as:

$$\xi = \frac{10 \times 0.21}{0 + 0.21 + 0.79} + \frac{0 \times 0.79}{0 + 0.21 + 0.79} + \frac{-10 \times 0}{0 + 0.21 + 0.79} = 2.1 + 0 + 0 = 2.1 \tag{2.13}$$

where *t* in this example is the number of additional assignments inferred from students' score [7]. In this example, two extra assignments would be given to the class.

As Fuzzy logic is considered to be deductive inference, [12, 18], in deductive inference, if the proposition is true, then the conclusion should be true. IF ~THEN~ rule can be explained with modus ponens meaning that the mode affirms. Modus ponens means that the 1st premise is a hypothetical proposition and the 2nd one is a categorical proposition. The categorical proposition affirms the priori of the hypothetical proposition. The conclusion is derived by affirming the posteriori of hypothetical proposition as shown in Figure 2.13.

1st premise (hypothetical proposition):

If score is 2 THEN more assignment will be given.

2nd premise (categorical proposition):

P : Score is 1

Conclusion:

Q : Two extra assignments are given.

Figure 2.13 Fuzzy Inferences

At first, the hypothetical proposition is given such as "IF score is 2 THEN more assignment will be given". The categorical proposition (score is 1) recognizes that the hypothetical proposition is true, and based on that, it expands its conclusion to "Two extra assignments are given".

2.5 Fuzzy Logic Control (FLC) and Takagi-Sugeno Fuzzy System

As shown in the previous example, fuzzy control involves the application of fuzzy inference. The rules of inference are linguistic and correspond to fuzzy sets. As such the inputs should be fuzzified and inference rules are then applied to these to produce fuzzy outputs. Fuzzy outputs are then transformed into crisp forms to allow connectivity to physical systems. This process is referred to as defuzzification and was demonstrated in

section 2.4. Experts' experiences and knowledge are essential to the construction of fuzzy rules and membership functions pertaining to fuzzification and defuzzification. The general concept for fuzzy system implementation is illustrated in Figure 2.14.

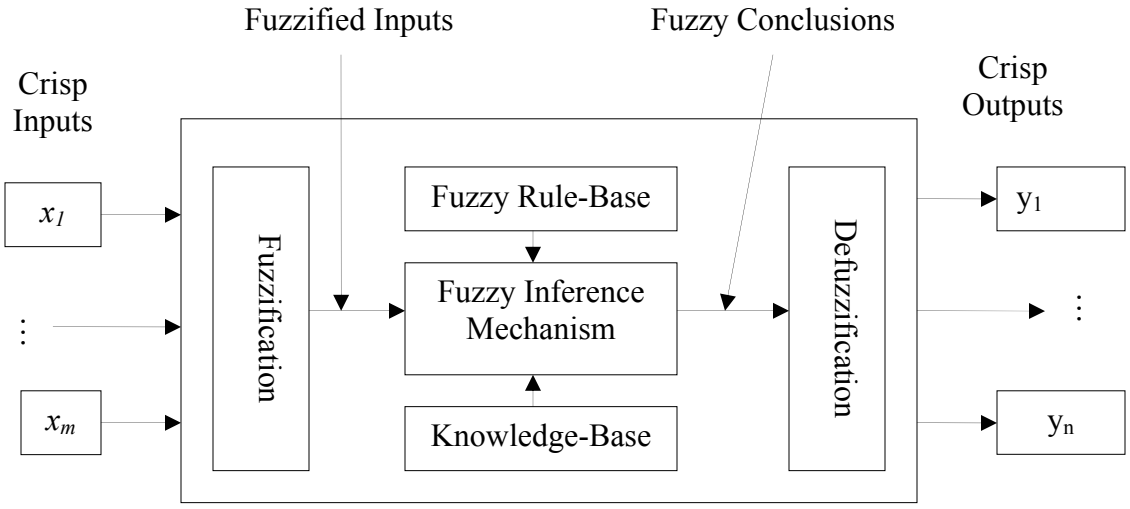


Figure 2.14 Fuzzy Controller Systems, [6]

The rules of inference are made from an expert's subjectivity and one's experience. There are, in general, a number of these rules of inference that are evaluated in parallel. Furthermore, the rules should be considered and applied to all of the data.

Fuzzy rules may also incorporate complex function evaluations**.** A mathematical function may be used to establish a relationship between inputs and outputs to show the conclusion from given inputs. For example:

$$\text{If } e_z \text{ is } \widetilde{Z}\text{, Then } y = f(e_z). \tag{2.14}$$

where $f(x)$ can be a complex dynamic function such as a model of a physical system. $\widetilde{Z}$ is a linguistic variable for fuzzy inputs and e_z is the set of fuzzy inputs.

The use of complex dynamic functions in fuzzy rules was first proposed by Takagi and Sugeno, [6], with the name of Takagi-Sugeno Fuzzy Logic Controller (FLC).

Takagi-Sugeno's methodology was proposed to put a controller in the right-hand side of a rule. The method also can allow changing of parameters in a controller or a model and, lead to the construction of special membership functions such as a Gaussian distribution function. Further to Takagi-Sugeno's methodology, a fuzzy rule pertaining to model update may be written as:

If e_{z_1} is $\widetilde{Z}_1$ and e_{z_2} is $\widetilde{Z}_2 \cdots$, e_{z_j} is $\widetilde{Z}_j$, then $X_{k+1} = A_i \times X_k + B_i \times U_k$ (2.15)

where X_k is a state vector, A_i is a system matrix, B_i is an input matrix, U_k is an input signal, e_z be fuzzy inputs that can be selected by designer's subjective decision, and $\widetilde{Z}_j$ are fuzzy sets. For instance, e_z can be a measured output that could be used to identify regions of piecewise linearity in a nonlinear system. Accordingly fuzzy sets $\widetilde{Z}_j$ quantify the meaning of a fuzzy input or in this example identify the region of piecewise linearity. For example, let there be two fuzzy inputs e_{z_1} and e_{z_2} to stand for position and velocity estimates in a hypothetical actuation system. Let $\widetilde{Z}_1$ and $\widetilde{Z}_2$ be operating regions linguistically defined as "Micro region" and "Macro region". In accordance to these fuzzy inputs and corresponding $\widetilde{Z}_j$, the fuzzy output is inferred by fuzzy membership functions $\mu_i(e_z)$, and a function according to Takagi-Sugeno method, which are determined by a designer. The inferred crisp output may be obtained by defuzzification of fuzzy output and the Center of Gravity method described in section 2.4 as, [6, 7]:

$$X_{k+1} = \frac{\sum_{i=1}^{S} (A_i \times X_k + B_i \times U_k) \times \mu_i(e_z)}{\sum_{i=1}^{S} \mu_i(e_z)} \tag{2.16}$$

where S is the number of fuzzy rules to be inferred.

Let ξ be a crisp variable specified as:

$$\xi_i = \frac{1}{\sum_{i=1}^{S} \mu_i(e_z)} \times \mu_i(e_z) \tag{2.17}$$

Then equation (2.16) can be depicted with this crisp output to show corrective action such as:

$$X_{k+1} = \frac{\sum_{i=1}^{S} (A_i \times X_k + B_i \times U_k) \times \mu_i(e_z)}{\sum_{i=1}^{S} \mu_i(e_z)} = \frac{\sum_{i=1}^{S} A_i \times X_k \times \mu_i(e_z)}{\sum_{i=1}^{S} \mu_i(e_z)} + \frac{\sum_{i=1}^{S} B_i \times U_k \times \mu_i(e_z)}{\sum_{i=1}^{S} \mu_i(e_z)}$$

$$= \left\{ \sum_{i=1}^{S} A_i \times \xi_i(e_z) \right\} \times X_k + \left\{ \sum_{i=1}^{S} B_i \times \xi_i(e_z) \right\} \times U_k \tag{2.18}$$

where, it is assumed that the above equation is in a matrix form.

μ_i can be a special membership function such as a Gaussian distribution function.

2.6 Parallel Distribution Compensation

This methodology is used in conjunction with the Takagi-Sugeno method for applications that involve corrective state feedback, [6]. Here, the input signal U_k is considered to be a defuzzified crisp signal that is a function of a compensation gain and system states. For a gain κ_j subjected to adjustment by defuzzification functions ψ_j, then

$$\left\{ \sum_{j=1}^{S} \kappa_j \times \psi_j(e_z) \right\} \times X_k = U_k \tag{2.19}$$

where similarly to equation (2.17), for a set of a membership functions λ_j and ψ_j, the crisp output is defined as:

$$\psi_j = \frac{\lambda_j(e_z)}{\sum_{j=1}^{S} \lambda_j(e_z)} \tag{2.20}$$

This method allows feedback gains to smoothly readjust in the presence of uncertainties, such as nonlinearities and time-varying dynamics. Substituting equation (2.19) into (2.18), the state equation is obtained as:

$$X_{k+1} = \left\{ \sum_{i=1}^{S} A_i \times \xi_i(e_z) \right\} \times X_k + \left\{ \sum_{i=1}^{S} B_i \times \xi_i(e_z) \right\} \times \left\{ \sum_{j=1}^{S} \kappa_j \times \psi_j(e_z) \right\} \times X_k \tag{2.21}$$

Chapter 3

Kalman Filter

The motion or operation of dynamic systems is described by variables referred to as states. For example, an aircraft's states may be defined as its position, velocity and acceleration. Availability of these states can greatly improve control of the aircraft. For example, knowing the position, velocity and acceleration of an aircraft can greatly aid an air traffic controller or a collision avoidance system in charting a safe course in a busy flight space. It is not always possible to measure all of the states associated with a system. In such circumstances, the states need to be estimated from the limited number of measurements that are available. To compound the problem, measurements are often corrupted by noise, thus adversely affecting the quality of the estimation process. An optimal strategy that is commonly used is the Kalman filter. This strategy is model-based and was first introduced by Kalman in the 1960s, [19, 20]. The Kalman filter uses an internal model to predict the initial or a priori estimate of the states. An optimal correction that is a function of the error in the predicted output and actual output of the system is applied to the a priori state estimate to obtain a refined or a posteriori state vector. In the Kalman filter formulation, both the system (w) and measurement (v) noise are taken into account. Noise is considered as being white, meaning that it is random and has a mean of zero. In the following sections an overview of the Kalman filter concept is provided.

3.1 Derivation of the Error Covariance

Consider a system described in a discrete form by the following state space equations:

$$X_{k+1} = A \times X_k + B \times U_k + w_k \tag{3.1}$$

$$z_k = H \times X_k + v_k \tag{3.2}$$

where the state vector is $X_k = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$, the input is $U_k = \begin{bmatrix} u_1 \\ \vdots \\ u_m \end{bmatrix}$, A is the system matrix, B is the input matrix, H is the output matrix, w_k is the system noise and v_k is the measurement noise.

Let $\hat{X}_k^-$ and $\hat{X}_k$ be defined as the a priori and the a posteriori estimates. The model of equations (3.1) and (3.2) are used for obtaining the a priori estimate such that:

$$\hat{X}_{k+1}^- = A \times \hat{X}_k + B \times U_k \tag{3.3}$$

$\hat{X}_k^-$ is referred to as the a priori estimate and the associate a priori error vector is:

$$e_k^- = X_k - \hat{X}_k^- \tag{3.4}$$

The a priori error covariance matrix is obtained as:

$$P_k^- = E\left[e_k^- \times e_k^{-T}\right] \tag{3.5}$$

Assuming that w_k (system noise), and v_k (measurement noise) are zero mean value, their covariance matrices are denoted as Q_k and R_k defined as:

$$Q_k = E\left[w_k \times w_k^T\right] \tag{3.6}$$

$$R_k = E\left[v_k \times v_k^T\right] \tag{3.7}$$

To obtain the optimal correction that is applied to the a priori estimate of equation (3.3), mathematical expressions need to be obtained for the a priori and the a posteriori error covariance matrices as follows.

3.1.1 Derivation of the A Priori Error Covariance

The Kalman filter is a recursive method. After predicting the a priori estimates using the model of the system as given by equation (3.3), the filter applies an optimal correction to the a priori estimate. This optimal correction is derived by examining and expanding the a priori error equation as follows:

$$\begin{aligned} e_{k+1}^- &= X_{k+1} - \hat{X}_{k+1}^- = A \times X_k + B \times U_k + w_k - \left(A \times \hat{X}_k + B \times U_k\right) \\ &= A \times X_k - A \times \hat{X}_k + B \times U_k - B \times U_k + w_k = A \times \left(X_k - \hat{X}_k\right) + w_k \end{aligned} \tag{3.8}$$

The corresponding a priori error covariance matrix expression is obtained by substituting equation (3.8) in equation (3.5):

$$\begin{aligned} P_{k+1}^- &= E\left[e_{k+1}^- \times e_{k+1}^{-\,T}\right] = E\left[\left\{A \times \left(X_k - \hat{X}_k\right) + w_k\right\} \times \left\{A \times \left(X_k - \hat{X}_k\right) + w_k\right\}^T\right] \\ &= E\left[A \times \left(X_k - \hat{X}_k\right) \times \left(X_k - \hat{X}_k\right)^T \times A^T\right] + E\left[w_k \times \left(X_k - \hat{X}_k\right)^T \times A^T\right] \\ &\quad + E\left[A \times \left(X_k - \hat{X}_k\right) \times w_k^T\right] + E\left[w_k \times w_k^T\right] \end{aligned} \tag{3.9}$$

Since w_k is white and uncorrelated to the vector $A \times \left(X_k - \hat{X}_k\right)$, the terms,

$E\left[w_k \times \left(X_k - \hat{X}_k\right)^T \times A_k^T\right]$ and $E\left[A_k \times \left(X_k - \hat{X}_k\right) \times w_k^T\right]$ are equal to zero.

Furthermore;

$$E\left[A\times\left(X_k-\hat{X}_k\right)\times\left(X_k-\hat{X}_k\right)^T\times A^T\right]=A\times E\left[\left(X_k-\hat{X}_k\right)\times\left(X_k-\hat{X}_k\right)^T\right]\times A^T$$

Equation (3.9) can then be simplified to:

$$P_{k+1}^{-}=A_k\times E\left[\left(X_k-\hat{X}_k\right)\times\left(X_k-\hat{X}_k\right)^T\right]\times A_k^T+E\left[w_k\times w_k^T\right] \tag{3.10}$$

Let e_k be the a posteriori error and defined as:

$$e_k=X_k-\hat{X}_k \tag{3.11}$$

Then the corresponding a posteriori error covariance is obtained as:

$$P_k=E\left[e_k\times e_k^T\right] \tag{3.12}$$

Substituting equations (3.6), (3.11) and (3.12) into (3.10), the a posteriori error covariance is obtained as:

$$P_{k+1}^{-}=A_k\times P_k\times A_k^T+Q_k \tag{3.13}$$

Equation (3.13) can therefore be used to project the a priori error covariance matrix one step ahead by using the a posteriori error covariance, P_k, and the system noise covariance matrix, Q_k.

3.1.2 Derivation of the A Posteriori Error Covariance and of the Gain

The optimal correction of Kalman filter to the a priori state estimate $\hat{X}_k^-$ consists of the Kalman gain times the error in the a priori output estimate, such that:

$$\hat{X}_k=\hat{X}_k^-+K_k\times\left(z_k-H\times\hat{X}_k^-\right) \tag{3.14}$$

The Kalman gain K_k is derived to minimize the a posteriori error covariance matrix, P_k and hence optimize the a posteriori estimate $\hat{X}_k$. The a posteriori error is written as:

$$e_k = X_k - \hat{X}_k = X_k - \left\{\hat{X}_k^- + K_k \times \left(z_k - H \times \hat{X}_k^-\right)\right\} \tag{3.15}$$

Substituting from equations (3.2), the a posteriori error is obtained as:

$$\begin{aligned} e_k &= X_k - \hat{X}_k^- - K_k \times \left(H \times X_k + v_k - H \times X_k^-\right) \\ &= X_k - \hat{X}_k^- - K_k \times H \times X_k - K_k \times v_k + K_k \times H \times X_k^- \\ &= \left(I - K_k \times H\right) \times X_k - \left(I - K_k \times H\right) \times \hat{X}_k^- - K_k \times v_k \\ &= \left(I - K_k \times H\right) \times \left(X_k - \hat{X}_k^-\right) - K_k \times v_k \end{aligned} \tag{3.16}$$

From equations (3.12) and (3.16), the a posteriori error covariance matrix is obtained as:

$$\begin{aligned} P_k &= E\left[\left\{\left(I - K_k \times H\right) \times \left(X_k - \hat{X}_k^-\right) - K_k \times v_k\right\} \times \left\{\left(I - K_k \times H\right) \times \left(X_k - \hat{X}_k^-\right) - K_k \times v_k\right\}^T\right] \\ &= E\Big[\left\{\left(I - K_k \times H\right) \times \left(X_k - \hat{X}_k^-\right) \times \left(X_k - \hat{X}_k^-\right)^T \times \left(I - K_k \times H\right)^T\right\} \\ &\quad - \left(I - K_k \times H\right) \times \left(X_k - \hat{X}_k^-\right) \times v_k^T \times K_k^T \\ &\quad - K_k \times v_k \times \left(X_k - \hat{X}_k^-\right)^T \times \left(I - K_k \times H\right)^T + K_k \times v_k \times v_k^T \times K_k^T\Big] \end{aligned} \tag{3.17}$$

As v_k is random and uncorrelated to the vector of $\left(I - K_k \times H\right) \times \left(X_k - \hat{X}_k^-\right)$, then $E\left[\left(I - K_k \times H\right) \times \left(X_k - \hat{X}_k^-\right) \times v_k^T \times K_k^T\right]$ and $E\left[K_k \times v_k \times \left(X_k - \hat{X}_k^-\right)^T \times \left(I - K_k \times H\right)^T\right]$ are equal to zero. For furthermore,

$E\left[\left\{(I-K_k\times H)\times\left(X_k-\hat{X}_k^-\right)\times\left(X_k-\hat{X}_k^-\right)^T\times(I-K_k\times H)^T\right\}\right]$ can be rewritten as

$(I-K_k\times H)\times E\left[\left(X_k-\hat{X}_k^-\right)\times\left(X_k-\hat{X}_k^-\right)^T\right]\times(I-K_k\times H)^T$ and $E\left[K_k\times v_k\times v_k^T\times K_k^T\right]$ changes to $K_k\times E\left[v_k\times v_k^T\right]\times K_k^T$.

Then the a posteriori error covariance matrix is obtained as:

$$P_k=(I-K_k\times H)\times E\left[\left(X_k-\hat{X}_k^-\right)\times\left(X_k-\hat{X}_k^-\right)^T\right]\times(I-K_k\times H)^T+K_k\times R_k\times K_k^T \qquad (3.18)$$

Note that the measurement noise covariance matrix is $R_k=E\left[v_k\times v_k^T\right]$ from equation (3.7) and that the a priori error covariance matrix is equal to $E\left[\left(X_k-\hat{X}_k^-\right)\times\left(X_k-\hat{X}_k^-\right)^T\right]$. Substituting these in equation (3.18), the a posteriori error covariance is now simplified to:

$$P_k=(I-K_k\times H)\times P_k^-\times(I-K_k\times H)^T+K_k\times R_k\times K_k^T \qquad (3.19)$$

The a posteriori error covariance can be rewritten as, [19]:

$$\begin{aligned}P_k&=(I-K_k\times H)\times P_k^-\times\left\{I^T-\left(H^T\times K_k^T\right)\right\}+K_k\times R_k\times K_k^T\\&=(I-K_k\times H)\times P_k^--(I-K_k\times H)\times P_k^-\times\left(H^T\times K_k^T\right)+K_k\times R_k\times K_k^T\\&=(I-K_k\times H)\times P_k^--P_k^-\times H^T\times K_k^T+K_k\times H\times P_k^-\times H^T\times K_k^T+K_k\times R_k\times K_k^T\end{aligned} \qquad (3.20)$$

In the quadratic form, $H\times P_k^-\times H^T+R_k$ is assumed to be symmetric and positive definite. To minimize P_k, the derivative of equation (3.20) is obtained with respect to K_k and equated to zero, such that:

$$-P_k^-\times H^T+K_k\times H\times P_k^-\times H^T+K_k\times R_k=0$$

The optimal Kalman gain K_k is obtained by rearranging the above equation such that, [19]:

$$K_k = P_k^- \times H^T \times \left(H \times P_k^- \times H^T + R_k\right)^{-1} \tag{3.21}$$

Substituting equation (3.21) into equation (3.20), the a posteriori error covariance matrix can be obtained by:

$$P_k = P_k^- - K_k \times H \times P_k^- = \left(I - K_k \times H\right) \times P_k^- \tag{3.22}$$

3.2 The Kalman Filter Estimation Process and Algorithm

The objective of the Kalman filter is to minimize the covariance matrix of e_k in order to obtain an optimal estimate $\hat{X}_k$ of the stochastic signal X_k. The Kalman filter is a predictor corrector method and its process is depicted in Figure 3.1. The estimation process involves the computation of the Kalman gain K_k (Step 1). This gain is then used in conjunction with the error in the prediction of the output, for correcting the estimate of the state vector into its a posteriori form (Step 2). The error covariance matrix is then computed for the updated or a posteriori state vector and a projection of it is obtained (Step 3 and 4). An a priori estimate of the state vector is predicted for the next iterative cycle using the model of the system (Step 5). The above steps are repeated iteratively as shown in Figure 3.1.

Step 1 Calculation of Kalman gain by using the a priori projection or initial condition: $K_k = P_k^- \times H^T \times \left(H \times P_k^- \times H^T + R_k\right)^{-1}$

$$\Downarrow$$

Step 2 Correction of the a priori estimate in its a posteriori form:

$$\hat{X}_k = \hat{X}_k^- + K_k \times \left(z_k - H \times \hat{X}_k^-\right)$$

$$\Downarrow$$

Step 3 Calculation of the a posteriori error covariance matrix:

$$P_k = \left(I - K_k \times H\right) \times P_k^-$$

$$\Downarrow$$

Step 4 Projection of the error covariance matrix for obtaining the a priori estimate for the next iteration cycle: $P_{k+1}^- = A \times P_k \times A^T + Q_k$

Step 5 Calculation of the a priori state estimate for the next iteration cycle:

$$\hat{X}_{k+1}^- = A \times \hat{X}_k + B \times U_k$$

Iteration

Figure 3.1 The Kalman Filter Process

Chapter 4

The Revised Kalman Filter

A revised state estimation error is presented here to enable the integration of the fuzzy logic with the Kalman filter concept. Here, the state estimates are denoted by using the symbol "~" in order to highlight the distinction of the Kalman filter from its revised formulation.

4.1 The Pre-Processor and the Projectile Estimates

The physical context associated with the output from the predictor and the corrector stages of the Revised-Kalman Filter do not lend themselves to the terminologies and the natural interpretations corresponding to those of the Kalman filter. Again to highlight the difference, the output of the prediction part of this revised form is referred to as the pre-processor estimate $\widetilde{X}^{-}$ and the corrector output is referred as the projectile estimate $\widetilde{X}$. The corresponding equations for the pre-processor and the projectile estimates are the following:

$$\widetilde{X}_{k+1}^{-} = A \times \widetilde{X}_k + \widetilde{K}_k \times \left(z_k - H \times \widetilde{X}_k \right) \tag{4.1}$$

$$\widetilde{X}_{k+1} = \widetilde{X}_{k+1}^{-} + B \times U_k \tag{4.2}$$

$\widetilde{X}_{k+1}^{-}$ is a function of the previous measurement error, the previous projectile estimate $\widetilde{X}_k$, and the gain $\widetilde{K}_k$. Substituting equation (4.1) into (4.2), the following governing equation is obtained:

$$\widetilde{X}_{k+1} = A \times \widetilde{X}_k + B \times U_k + \widetilde{K}_k \times \left(z_k - H \times \widetilde{X}_k \right) \tag{4.3}$$

This differs from the governing equation of the Kalman filter obtained from equations (3.3) and (3.14) that simplify to:

$$\hat{X}_{k+1} = \hat{X}_{k+1}^{-} + K_{k+1} \times \left(z_{k+1} - H \times \hat{X}_{k+1}^{-} \right) = A \times \hat{X}_k + B \times U_k + K_{k+1} \times \left(z_{k+1} - H \times \hat{X}_{k+1}^{-} \right) \tag{4.4}$$

In the revised method, $\widetilde{X}_k$ is corrected by a gain and a measurement error vector that lag by one time step. The fundamental difference in approach is that the current estimate is optimized in the case of the revised formulation with the previous a posteriori output error rather than the current a priori output estimate. With this modification, the initial measurement error can be assumed to be zero. Therefore, recursive computation can be started with possibly a smaller initial error that is transmitted into the next step. Furthermore, the revised formulation is not strictly a predictor-corrector methodology and can be considered as being an optimal filter.

4.2 Derivation of the Projectile Estimate Error Covariance Matrix $\widetilde{P}_k$

Let the error between the actual state and the projectile estimate be defined as:

$$e_{x_k} = X_k - \widetilde{X}_k \tag{4.5}$$

Then the projectile error e_{x_k} covariance is denoted by $\widetilde{P}_k$ and is defined as:

$$\widetilde{P}_k = E\left[e_{x_k} \times e_{x_k}^T\right] \tag{4.6}$$

Substituting equation (4.2) into (4.5) then:

$$e_{x_{k+1}} = X_{k+1} - \widetilde{X}_{k+1} = X_{k+1} - \left(\widetilde{X}_{k+1}^- + B \times U_k\right) \tag{4.7}$$

Substituting equations (3.1) and (4.1) into equation (4.7), then:

$$e_{x_{k+1}} = A \times X_k + B \times U_k + w_k - \left\{A \times \widetilde{X}_k + \widetilde{K}_k \times \left(z_k - H \times \widetilde{X}_k\right) + B \times U_k\right\} \tag{4.8}$$

Substituting equation (3.2) into equation (4.8):

$$e_{x_{k+1}} = A \times X_k + B \times U_k + w_k - \left\{A \times \widetilde{X}_k + \widetilde{K}_k \times \left(H \times X_k + v_k - H \times \widetilde{X}_k\right) + B \times U_k\right\} \tag{4.9}$$

The projectile error equation can be re-arranged as:

$$e_{x_{k+1}} = \left(A - \widetilde{K}_k \times H\right) \times \left(X_k - \widetilde{X}_k\right) + w_k - \widetilde{K}_k \times v_k \tag{4.10}$$

Where, the projectile covariance matrix is specified as:

$$\widetilde{P}_{k+1} = E\left[e_{x_{k+1}} \times e_{x_{k+1}}^T\right] \tag{4.11}$$

Substituting from equation (4.10) into equation (4.11) and rearranging:

$$\begin{aligned}
\widetilde{P}_{k+1} &= E\Big[\Big\{\left(A - \widetilde{K}_k \times H\right) \times \left(X_k - \widetilde{X}_k\right) + w_k - \widetilde{K}_k \times v_k\Big\} \\
&\quad \times \Big\{\left(A - \widetilde{K}_k \times H\right) \times \left(X_k - \widetilde{X}_k\right) + w_k - \widetilde{K}_k \times v_k\Big\}^T\Big] \\
&= E\Big[\Big\{\left(A - \widetilde{K}_k \times H\right) \times \left(X_k - \widetilde{X}_k\right) + w_k - \widetilde{K}_k \times v_k\Big\} \\
&\quad \times \Big\{\left(X_k - \widetilde{X}_k\right)^T \times \left(A - \widetilde{K}_k \times H\right)^T + w_k^T - v_k^T \times \widetilde{K}_k^T\Big\}\Big] \\
&= E\Big[\left(A - \widetilde{K}_k \times H\right) \times \left(X_k - \widetilde{X}_k\right) \times \left(X_k - \widetilde{X}_k\right)^T \times \left(A - \widetilde{K}_k \times H\right)^T \\
&\quad + \left(A - \widetilde{K}_k \times H\right) \times \left(X_k - \widetilde{X}_k\right) \times w_k^T - \left(A - \widetilde{K}_k \times H\right) \times \left(X_k - \widetilde{X}_k\right) \times v_k^T \times \widetilde{K}_k^T
\end{aligned}$$

$$+ w_k \times \left(X_k - \widetilde{X}_k\right)^T \times \left(A - \widetilde{K}_k \times H\right)^T + w_k \times w_k^T - w_k \times v_k^T \times \widetilde{K}_k^T$$

$$- \widetilde{K}_k \times v_k \times \left(X_k - \widetilde{X}_k\right)^T \times \left(A - \widetilde{K}_k \times H\right)^T - \widetilde{K}_k \times v_k \times w_k^T + \widetilde{K}_k \times v_k \times v_k^T \times \widetilde{K}_k^T \Big] \quad (4.12)$$

The vectors w_k and v_k are white and uncorrelated, and hence $\widetilde{P}_{k+1}$ simplifies to the following.

$$\widetilde{P}_{k+1} = E\Big[\left(A - \widetilde{K}_k \times H\right) \times \left(X_k - \widetilde{X}_k\right) \times \left(X_k - \widetilde{X}_k\right)^T \times \left(A - \widetilde{K}_k \times H\right)^T + w_k \times w_k^T$$

$$+ \widetilde{K}_k \times v_k \times v_k^T \times \widetilde{K}_k^T \Big] \quad (4.13)$$

The projectile covariance matrix can be further simplified to:

$$\widetilde{P}_{k+1} = \left(A - \widetilde{K}_k \times H\right) \times E\Big[\left(X_k - \widetilde{X}_k\right) \times \left(X_k - \widetilde{X}_k\right)^T\Big] \times \left(A - \widetilde{K}_k \times H\right)^T + E\left[w_k \times w_k^T\right]$$

$$+ \widetilde{K}_k \times E\left[v_k \times v_k^T\right] \times \widetilde{K}_k^T \quad (4.14)$$

For the system noise covariance matrix defined as $Q_k = E\left[w_k \times w_k^T\right]$, the measurement noise covariance defined as $R_k = E\left[v_k \times v_k^T\right]$ and where $\widetilde{P}_k = E\Big[\left(X_k - \widetilde{X}_k\right) \times \left(X_k - \widetilde{X}_k\right)^T\Big]$ then from equation (4.14) the projectile error covariance is obtained as:

$$\widetilde{P}_{k+1} = \left(A - \widetilde{K}_k \times H\right) \times \widetilde{P}_k \times \left(A - \widetilde{K}_k \times H\right)^T + Q_k + \widetilde{K}_k \times R_k \times \widetilde{K}_k^T \quad (4.15)$$

In an expanded form, the projectile error covariance can be expressed as:

$$\widetilde{P}_{k+1} = \left(A - \widetilde{K}_k \times H\right) \times \widetilde{P}_k \times \left(A^T - H^T \times \widetilde{K}_k^T\right) + Q_k + \widetilde{K}_k \times R_k \times \widetilde{K}_k^T$$

$$= A \times \widetilde{P}_k \times A^T - A \times \widetilde{P}_k \times H^T \times \widetilde{K}_k^T - \widetilde{K}_k \times H \times \widetilde{P}_k \times A^T + \widetilde{K}_k \times H \times \widetilde{P}_k \times H^T \times \widetilde{K}_k^T + Q_k$$

$$+ \widetilde{K}_k \times R_k \times \widetilde{K}_k^T \quad (4.16)$$

4.3 Derivation of the Gain $\widetilde{K}_k$

Equation (4.16) can be expanded such that:

$$\widetilde{P}_{k+1} = A\times\widetilde{P}_k\times A^T - A\times\widetilde{P}_k\times H^T\times\widetilde{K}_k^T - \widetilde{K}_k\times H\times\widetilde{P}_k\times A^T + Q_k$$

$$+\widetilde{K}_k\times\left(H\times\widetilde{P}_k\times H^T + R_k\right)\times\widetilde{K}_k^T \tag{4.17}$$

Similarly to the Kalman filter, $\widetilde{K}_k\times\left(H\times\widetilde{P}_k\times H^T + R_k\right)\times\widetilde{K}_k^T$ is considered as symmetric and positive definite. As such, it can be factored as, [20]:

$$S_k\times S_k^* = H\times\widetilde{P}_k\times H^T + R_k \tag{4.18}$$

Where S_k^* is symmetric and the transpose complex conjugate of the matrix S_k. Assuming S_k to be real, then, equation (4.16) can be expressed in terms of S_k such that:

$$\widetilde{P}_{k+1} = A\times\widetilde{P}_k\times A^T - A\times\widetilde{P}_k\times H^T\times\widetilde{K}_k^T - \widetilde{K}_k\times H\times\widetilde{P}_k\times A^T + Q_k + \widetilde{K}_k\times S_k\times S_k^T\times\widetilde{K}_k^T \tag{4.19}$$

After completing the square, the projectile estimate covariance is formed as:

$$\widetilde{P}_{k+1} = A\times\widetilde{P}_k\times A^T + \left(\widetilde{K}_k\times S_k - C\right)\times\left(\widetilde{K}_k\times S_k - C\right)^T + Q_k - C\times C^T \tag{4.20}$$

where: $C = A\times\widetilde{P}_k\times H^T\times\left(S_k^T\right)^{-1}$. Note that C does not involve $\widetilde{K}_k$.

Let the term $\left(\widetilde{K}_k\times S_k - C\right)\times\left(\widetilde{K}_k\times S_k - C\right)^T$ in equation (4.20) be forced to be zero by letting:

$$\widetilde{K}_k = C\times S_k^{-1} \tag{4.21}$$

This minimizes $\widetilde{P}_{k+1}$ since it would result in $\frac{\partial\widetilde{P}_{k+1}}{\partial\widetilde{K}_k} = 0$. If the matrix S_k is invertible, then from the equation (4.21), the Revised-Kalman gain is obtained as:

$$\widetilde{K}_k = C \times S_k^{-1} = A \times \widetilde{P}_k \times H^T \times \left(S_k^T\right)^{-1} \times S_k^{-1} = A \times \widetilde{P}_k \times H^T \times \left(S_k \times S_k^T\right)^{-1}$$

$$= A \times \widetilde{P}_k \times H^T \times \left(H \times \widetilde{P}_k \times H^T + R_k\right)^{-1} \tag{4.22}$$

Substituting equation (4.22) in (4.17), $\widetilde{P}_{k+1}$ is rewritten as:

$$\widetilde{P}_{k+1} = A \times \widetilde{P}_k \times A^T - A \times \widetilde{P}_k \times H^T \times \widetilde{K}_k^T - \widetilde{K}_k \times H \times \widetilde{P}_k \times A^T + Q_k$$

$$+ A \times \widetilde{P}_k \times H^T \times \left(H \times \widetilde{P}_k \times H^T + R_k\right)^{-1} \times \left(H \times \widetilde{P}_k \times H^T + R_k\right) \times \widetilde{K}_k^T \tag{4.23}$$

Rearranging equation (4.23), $\widetilde{P}_{k+1}$ simplifies to:

$$\widetilde{P}_{k+1} = A \times \widetilde{P}_k \times A^T - \widetilde{K}_k \times H \times \widetilde{P}_k \times A^T + Q_k \tag{4.24}$$

The projectile error covariance matrix is then obtained as:

$$\widetilde{P}_{k+1} = \left(A - \widetilde{K}_k \times H\right) \times \widetilde{P}_k \times A^T + Q_k \tag{4.25}$$

The process for the Revised-Kalman filter can now be summarized as given in Figure 4.1.

Step 1 Calculation of the Revised-Kalman gain:

$$\widetilde{K}_k = A \times \widetilde{P}_k \times H^T \times \left(H \times \widetilde{P}_k \times H^T + R_k\right)^{-1}$$

$$\Downarrow$$

Step 2 Correction of the projectile estimate into its a priori pre-processor form

$$\widetilde{X}_{k+1}^{-} = A \times \widetilde{X}_k + \widetilde{K}_k \times \left(z_k - H \times \widetilde{X}_k\right)$$

$$\Downarrow$$

Step 3 Calculation of the projectile state estimate for the next iteration cycle

$$\widetilde{X}_{k+1} = \widetilde{X}_{k+1}^{-} + B \times U_k$$

$$\Downarrow$$

Step 4 Calculation of the projectile error covariance matrix for obtaining the projectile estimate for the next iteration cycle:

$$\widetilde{P}_{k+1} = \left(A - \widetilde{K}_k \times H\right) \times \widetilde{P}_k \times A^T + Q_k$$

Iteration

Figure 4.1 The Revised-Kalman Filter Process

Chapter 5

The Fuzzy Kalman Filter

Fuzzy logic has been considered for state estimation in [21, 22, 23]. The objective of this research is to propose a Fuzzy-Kalman state estimation strategy. The Fuzzy control concept employs a linguistic approach in that the instantaneous value of controlled variables depends on the inference derived by the IF-THEN-ELSE type rules. The rules are generally drawn from state space relationships as might apply to piecewise linear systems. Accordingly, the fuzzy logic approach can be applied to uncertain linear or nonlinear systems. Furthermore fuzzy membership function can be used for improving the characterization of system and measurement noise.

Conventional logic is binary, that is, something is true or false, positive or negative, on or off. Most real world situations, however, do not easily conform to such rigid rules and are therefore difficult to control or be modeled using conventional logic. Fuzzy logic is a reasoning method that can deal with the uncertainty of the real world. Based on fuzzy set theory, fuzzy logic describes inputs in terms of their membership or relevance to a description. In this way, Fuzzy logic can provide an effective treatment for uncertainty description as it is linguistically based.

Fuzzy logic systems rely on two distinct and crisp set of variables. These are the system's inputs and outputs that require numerical values and physical units. Using fuzzy logic theory, such physical variables can be corresponded to several linguistic variables

through membership functions as explained in Chapter 2. A membership function defines the degree to which a certain physical variable can be associated with a linguistic variable.

It is also noted that the linguistic variables can be presented using stochastic descriptions. In the search for an easy, efficient, cost-effective control design to satisfy increasing demands in terms of accuracy, maneuverability, and stability, a stochastic form of rule-construction in fuzzy logic seems to provide a method of reducing the complexity of systems and rules while increasing control performance. In this chapter, the merging of the stochastic Revised-Kalman filter with fuzzy logic is considered. The ability to model problems in a simple and human-oriented way and to produce smooth control actions around the set points makes fuzzy logic an especially suitable candidate for use in control applications.

5.1 The Fuzzy-Kalman Filter

Let $\widetilde{X}_k$ be the Fuzzy-Kalman projectile estimate instead of the Revised-Kalman estimate. And let e_{z_k} be the measurement error defined as:

$$e_{z_k} = z_k - H \times \widetilde{X}_k \tag{5.1}$$

As such, e_{z_k} is used as the input to a membership function with the corresponding element for center of gravity crisp output being defined as:

$$\xi_i = \frac{\mu_i\left(e_{z_k}\right)}{\sum_{i=1}^{S} \mu_i\left(e_{z_k}\right)} \tag{5.2}$$

Further to the pre-processor equation of the Revised-Kalman filter in equation (4.1), the corresponding Takagi-Sugeno fuzzy system is obtained as:

$$\widetilde{X}_{k+1}^{-} = \sum_{i=1}^{S} A \times \xi_i\left(e_{z_k}\right) \times \widetilde{X}_k + \sum_{i=1}^{S} \xi_i\left(e_{z_k}\right) \times \widetilde{K}_k \times \left(z_k - H \times \widetilde{X}_k\right) \tag{5.3}$$

While adopting the concept of the "parallel distributed compensators", discussed in Chapter 2, a residual $\widetilde{K}_k \times \left(z_k - H \times \widetilde{X}_k\right)$ can be expressed with a compensator by a projectile estimate $\widetilde{X}_k$ as, [6]:

$$\text{IF } e_{z_k} \text{ is } \delta_k \text{ THEN } \left\{\widetilde{K}_k \times \left(z_k - H \times \widetilde{X}_k\right)\right\} = \kappa_k \times \widetilde{X}_k \tag{5.4}$$

κ_k is a gain that is derived later further to considerations for optimality the fuzzy rule of equation (5.4) can be realized by using corresponding membership and defuzzification functions, such that from equation (2.20), the defuzzified crisp output ψ of the above realization is obtained as:

$$\psi_j = \frac{\lambda_j\left(e_{z_k}\right)}{\sum_{j=1}^{S} \lambda_j\left(e_{z_k}\right)} \tag{5.5}$$

Further to equations (5.4) and (5.5), then:

$$\widetilde{K}_k \times \left(z_k - H \times \widetilde{X}_k\right) = \sum_{j=1}^{S} \kappa_{k_j} \times \psi_j\left(e_{z_k}\right) \times \widetilde{X}_k \tag{5.6}$$

Note the state feedback is now scaled according to a membership function that is a function of the estimate error. It is effectively a gain that is adjusted according to the region of the error as well as its magnitude. Further to this strategy, such regions may be defined not only in terms of error magnitude but also nonlinearity of the system. The nonlinearities are considered in the definition of fuzzy rules as given in equation (5.4)

and allow for the smooth interpolation of the gain according to the operation of a nonlinear system, [6].

Equation (5.6) effectively provides a corrective gain that is a function of an error between the predictive and actual output of the system, scaled according to a membership function that is adjusted to the operating region. Following a similar strategy as the Revised-Kalman filter and substituting equation (5.6) into equation (5.3), the projectile estimate becomes:

$$\widetilde{X}_{k+1}^{-} = \sum_{i=1}^{S} A \times \xi_i\left(e_{z_k}\right) \times \widetilde{X}_k + \sum_{i=1}^{S} \xi_i\left(e_{z_k}\right) \times \sum_{j=1}^{S} \kappa_{k_j} \times \psi_j\left(e_{z_k}\right) \times \widetilde{X}_k$$

$$= \left\{ \sum_{i=1}^{S} A \times \xi_i\left(e_{z_k}\right) + \sum_{i=1}^{S} \xi_i\left(e_{z_k}\right) \times \sum_{j=1}^{S} \kappa_{k_j} \times \psi_j\left(e_{z_k}\right) \right\} \times \widetilde{X}_k \tag{5.7}$$

It is assumed that the system is piece-wise linear and that ξ_i can be used for the selection of the operating region according to the error e_z or the measured output z. If there is only one operating region such that *S*=1, then the system is linear and, there is one fuzzy rule for fuzzy inference (i.e. *S*=1). Then the Fuzzy-Kalman pre-processor estimate $\widetilde{X}_{k+1}^{-}$ becomes similar to that of the Revised-Kalman filter. Considering the case of linear systems and 1 membership function such that *S*=1, then ξ_i is of dimension 1×1. Furthermore if only 1 error region is considered, then *S*=1 and equation (5.7) simplifies to:

$$\widetilde{X}_{k+1}^{-} = \left\{ A + \kappa_k \times \psi\left(e_{z_k}\right) \right\} \times \xi\left(e_{z_k}\right) \times \widetilde{X}_k \tag{5.8}$$

5.2 The Fuzzy Residual Representation

By experience and repeated experimentation, a set of membership functions may be obtained. This experimental heuristic based approach for defining membership functions is the advantage of fuzzy logic. The general strategy followed in fuzzy control to achieve this is as follows:

- Experimentation, data collection, and observation
- Generation of membership functions for fuzzification
- Rule based inference
- Defuzzification

Membership functions can be intuitively specified according to the designer's preferences; however, the choice of defuzzification strategy is more limited and the center of gravity is predominantly used. In the fuzzy theory application to filtering, a rule table is needed to derive a proper fuzzy interpolator from data associations between varying noise characteristics and estimates at each operating condition. If proper compensation is obtained, this interpolator can be used to coordinate for smooth switching i.e. in this context from one operating model to the next. In the context of Fuzzy Kalman filter, a mechanism for this interpolation is provided through the use of membership function ψ and ξ. In the remainder of this chapter, the derivation of the function ψ and ξ are considered based on a heuristic discussion of the measurement error covariance matrix, the Kalman Filter, and the Fuzzy-Kalman filter.

5.3 Discussion on the Functionality of the Kalman Filter

To briefly examine how the Kalman gain works according to the a priori estimation error covariance matrix P_k^- and the measurement noise covariance R_k, assume that the output matrix H is an identity matrix such that:

$$H = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{5.9}$$

Then the a posteriori estimate of the Kalman filter is obtained as:

$$\hat{X}_k = \hat{X}_k^- + K_k \times \left(X_k + v_k - \hat{X}_k^- \right) \tag{5.10}$$

In equation (3.21), the Kalman gain is stated as:

$$K_k = P_k^- \times H^T \times \left(H \times P_k^- \times H^T + R_k \right)^{-1} \tag{3.21}$$

where R_k and P_k^- are the measurement noise and the a priori error covariance matrices. If the output matrix H is assumed to be an identity matrix, then the Kalman gain simplifies to:

$$K_k = \frac{P_k^-}{P_k^- + R_k} \tag{5.11}$$

If P_k^- is very large and R_k is very small, then the Kalman gain K_k approximates to:

$$K_k = \frac{P_k^-}{P_k^- + R_k} \approx \frac{P_k^-}{P_k^-} = I \tag{5.12}$$

Then, the a posteriori estimate is obtained as:

$$\lim_{\substack{R_k \to 0, \\ V_k \to 0}} \left(\hat{X}_k \right) = \hat{X}_k^- + I \times \left(X_k - \hat{X}_k^- \right) = \hat{X}_k^- + X_k - \hat{X}_k^- = X_k$$

Therefore, if the measurement noise is very small in amplitude, then the Kalman filter estimate gets close to the actual state. If P_k^- is very small compared to R_k, then K_k can be approximated to:

$$\lim_{P_k^- \to 0}(K_k) = \lim_{P_k^- \to 0}\left(\frac{P_k^-}{P_k^- + R_k}\right) \approx \frac{0}{0 + R_k} = 0 \tag{5.13}$$

Then, the a posteriori estimate is obtained as:

$$\hat{X}_k = \hat{X}_k^- + K_k \times \left(X_k + v_k - \hat{X}_k^-\right) \approx \hat{X}_k^- + 0 \times \left(X_k + v_k - \hat{X}_k^-\right) = \hat{X}_k^-$$

When the estimation error is small even though the measurement error is big, it can be said that the only estimate is the a priori one.

5.4 Discussion of the functionality of the Revised Kalman Filter

Assuming an Identity output matrix, the pre-processor of the Revised-Kalman filter in equation (4.1) changes to:

$$\widetilde{X}_{k+1}^- = A \times \widetilde{X}_k + \widetilde{K}_k \times \left(X_k + v_k - \widetilde{X}_k\right) \tag{5.14}$$

The Revised-Kalman gain then simplifies to:

$$\widetilde{K}_k = \frac{A \times \widetilde{P}_k}{\widetilde{P}_k + R_k} \tag{5.15}$$

If $\widetilde{P}_k$ is very large compared to R_k, the Revised-Kalman gain $\widetilde{K}_k$ can be obtained as:

$$\lim_{R_k \to 0}\left(\widetilde{K}_k\right) = \lim_{R_k \to 0}\left(\frac{A \times \widetilde{P}_k}{\widetilde{P}_k + R_k}\right) \approx \frac{A \times \widetilde{P}_k}{\widetilde{P}_k + 0} = A \tag{5.16}$$

Then, the pre-processor estimate is obtained as:

$$\widetilde{X}_{k+1}^- = A \times \widetilde{X}_k + \widetilde{K}_k \times \left(X_k + v_k - \widetilde{X}_k\right) \approx A \times \widetilde{X}_k + A \times \left(X_k - \widetilde{X}_k\right) = A \times X_k \tag{5.17}$$

If the measurement error is very small, then the pre-processor estimate gets close to the actual state vector. If $\widetilde{P}_k$ is very small compared to R_k, then $\widetilde{K}_k$ approaches zero:

$$\lim_{P_k \to 0}\left(\widetilde{K}_k\right) = \lim_{P_k \to 0}\left(\frac{A \times \widetilde{P}_k}{\widetilde{P}_k + R_k}\right) \approx \frac{A \times 0}{0 + R_k} = 0 \tag{5.18}$$

Then, the pre-processor estimate is obtained as:

$$\widetilde{X}_{k+1}^{-} = A \times \widetilde{X}_k + \widetilde{K}_k \times \left(X_k + v_k - \widetilde{X}_k\right) \approx A \times \widetilde{X}_k + 0 \times \left(X_k + v_k - \widetilde{X}_k\right) = A \times \widetilde{X}_k \tag{5.19}$$

Hence similarly to the Kalman filter, when the amplitude of the measurement noise is high, the final estimate is due to the a priori stage of the filter. Therefore, under that condition, the Kalman gain or the Revised-Kalman filter gain do not have meaningful effects on the estimation error. The higher the error, the more significant is the role of the gains.

5.5 The Measurement Error Covariance Matrix

The projectile error covariance of the Revised-Kalman filter equation (4.14) can be used for constructing the membership function of ψ, which acts for the representation of the residual in equation (5.6). These two equations are restated here and are follows:

$$\widetilde{P}_{k+1} = \left(A - \widetilde{K}_k \times H\right) \times E\left[\left(X_k - \widetilde{X}_k\right) \times \left(X_k - \widetilde{X}_k\right)^T\right] \times \left(A - \widetilde{K}_k \times H\right)^T + E\left[w_k \times w_k^T\right] + \widetilde{K}_k \times E\left[v_k \times v_k^T\right] \times \widetilde{K}_k^T \tag{4.14}$$

$$\widetilde{K}_k \times \left(z_k - H \times \widetilde{X}_k\right) = \sum_{j=1}^{S} \kappa_{k_j} \times \psi_j\left(e_{z_k}\right) \times \widetilde{X}_k \tag{5.6}$$

In the Revised-Kalman filter, the measurement error covariance matrix can be derived from the measurement error vector of $z_k - H \times \widetilde{X}_k$. Let $\hat{P}_k$ be the measurement error covariance matrix obtained as:

$$\hat{P}_k = E\left[e_{z_k} \times e_{z_k}^T\right] \tag{5.20}$$

The state estimation error e_{x_k} is defined as: $e_{x_k} = X_k - \widetilde{X}_k$, then the measurement error e_{z_k} is obtained as:

$$e_{z_k} = H \times \left(X_k - \widetilde{X}_k\right) + v_k = H \times e_{x_k} + v_k \tag{5.21}$$

Substituting equation (5.21) in equation (5.20), then the measurement error covariance equation is as follows:

$$\begin{aligned}
\hat{P}_k &= E\left[\left(H \times e_{x_k} + v_k\right) \times \left(H \times e_{x_k} + v_k\right)^T\right] \\
&= E\left[H \times e_{x_k} \times \left(H \times e_{x_k}\right)^T + v_k \times \left(H \times e_{x_k}\right)^T + \left(H \times e_{x_k}\right) \times v_k^T + v_k \times v_k^T\right] \\
&= E\left[H \times e_{x_k} \times \left(H \times e_{x_k}\right)^T\right] + E\left[v_k \times \left(H \times e_{x_k}\right)^T\right] + E\left[\left(H \times e_{x_k}\right) \times v_k^T\right] \\
&\quad + E\left[v_k \times v_k^T\right]
\end{aligned} \tag{5.22}$$

Since the vector v_k is white and uncorrelated to other terms, then the terms $E\left[v_k \times \left(H \times e_{x_k}\right)^T\right]$ and $E\left[\left(H \times e_{x_k}\right) \times v_k^T\right]$ are equal to zero. Furthermore, $E\left[H \times e_{x_k} \times \left(H \times e_{x_k}\right)^T\right]$, $E\left[e_{x_k} \times e_{x_k}^T\right]$ and $E\left[v_k \times v_k^T\right]$ are equal to $H \times E\left[e_{x_k} \times e_{x_k}^T\right] \times H^T$, $\widetilde{P}_k$ and R_k respectively.

Then, the measurement error covariance is obtained from the projectile estimation error as:

$$\hat{P}_k = H \times \widetilde{P}_k \times H^T + R_k \tag{5.23}$$

5.6 Heuristic Adaptation of the Measurement Noise Covariance Matrix

The measurement error can be calculated as: $z_k - H \times \widetilde{X}_k$ in state estimation. Let the probability distribution of this error function be approximated by the probability density function specified as:

$$\Phi_k = e^{-\left\{\left(z_k - H \times \widetilde{X}_k\right) \times \left(z_k - H \times \widetilde{X}_k\right)^T\right\}} \tag{5.24}$$

Φ_k is a function of the measurement error and like R_k, reflects the presence and magnitude of the measurement noise in the Revised-Kalman filtering of Chapter 4. Normally for linear systems a fixed R_k is used. Here to allow for the adjustment of R_k using a heuristic assumption of a noise amplitude that is normally distributed, R_k is substituted by Φ_k for its bounded variation according to the magnitude of the error that can be viewed as a function of measurement noise.

Let $\widetilde{P}_k$ be the Fuzzy-Kalman projectile estimation error covariance matrix. Then, the fuzzy crisp gain κ_k in equation (5.6) is redefined in terms of Φ_k instead of R as:

$$\kappa_k = A \times \widetilde{P}_k \times H^T \times \left(H \times \widetilde{P}_k \times H^T + \Phi_k\right)^{-1} \tag{5.25}$$

To see briefly how this fuzzy crisp gain works according to the projectile estimate error covariance $\widetilde{P}_k$ and the function Φ_k, assume that the output matrix H is the identity matrix such that: $H = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

Then, the fuzzy crisp gain is simplified to:

$$\kappa_k = \frac{A \times \widetilde{P}_k}{\widetilde{P}_k + \Phi_k} \tag{5.26}$$

If $\widetilde{P}_k$ is very small, then regardless of Φ_k, κ_k approaches zero, as illustrated in equation (5.27):

$$\lim_{\widetilde{P}_k \to 0}(\kappa_k) = \lim_{\widetilde{P}_k \to 0}\left(\frac{A \times \widetilde{P}_k}{\widetilde{P}_k + \Phi_k}\right) = \frac{A \times 0}{0 + \Phi_k} = 0 \tag{5.27}$$

This is conceptually similar to the case of the Kalman Filter and the Revised-Kalman Filter as given by equations (5.13) and (5.18). If $\widetilde{P}_k$ and the measurement error are very large, then Φ_k becomes very small, such that:

$$\kappa_k = \frac{A \times \widetilde{P}_k}{\widetilde{P}_k + \Phi_k} \approx \frac{A \times \widetilde{P}_k}{\widetilde{P}_k + 0} = A \tag{5.28}$$

This is similar to equation (5.16) pertaining to the Revised-Kalman filter where limiting values for *R* lead to the same results. Hence the functionality of the Revised-Kalman filter at the limits of *R* is preserved by replacing a fixed *R* by Φ_k that adapts to the magnitude of the error. Note that when the measurement error is small, then:

$$\Phi_k = e^{-\left\{\left(z_k - H \times \widetilde{X}_k\right) \times \left(z_k - H \times \widetilde{X}_k\right)^T\right\}} \approx e^0 = I \tag{5.29}$$

This results in a fuzzy crisp gain κ_k such that:

$$\kappa_k = \frac{A \times \widetilde{P}_k}{\widetilde{P}_k + \Phi_k} \approx \frac{A \times \widetilde{P}_k}{\widetilde{P}_k + I} \tag{5.30}$$

This feature maintains a corrective action while $\widetilde{P}_k$ and the measurement error is not zero despite the small amplitude of the error. Only when $\widetilde{P}_k$ is equal to zero and the measurement error is equal to zero, then κ_k approaches zero such that:

$$\lim_{\widetilde{P}_k \to 0}(\kappa_k) = \lim_{\widetilde{P}_k \to 0}\left(\frac{A \times \widetilde{P}_k}{\widetilde{P}_k + \Phi_k}\right) = \frac{A \times 0}{0 + I} = 0 \tag{5.31}$$

The heuristic definition of Φ_k given in equation (5.29) therefore presents an advantage over a fixed *R*. Its use in multiples allows the gain to be scaled according to operating "region" of the error function.

5.7 Heuristic Adaptation of the System Noise Covariance Matrix

The adaptation of the system noise covariance matrix *Q* can be handled in a similar fashion. In equation (4.25), Q_k is used in $\widetilde{P}_{k+1}$ to reflect the presence and magnitude of uncertainties and system noise in the Revised-Kalman filter. In equations (3.1) and (4.3), state estimation is obtained by using the residual $\kappa_k \times \left(z_k - H \times \widetilde{X}_k\right)$. This partly incorporates system noise (w_k) and can thus be used for adapting Q_k.

Let Ω_k be defined as:

$$\Omega_k = \kappa_k \times \hat{P}_k \times \kappa_k^T \tag{5.32}$$

This function not only reflects the estimation error but also the corrective gain κ_k that are both related to modeling uncertainty and noise as reflected by *Q*. Let $\widetilde{P}_k$ be the Fuzzy-Kalman projectile estimation error covariance (not the Revised-Kalman's). Let Ω_k and κ_k be substituted into equation (4.25) instead of Q_k and $\widetilde{K}_k$ respectively. Then the Fuzzy-Kalman projectile estimation error covariance $\widetilde{P}_{k+1}$ is obtained as:

$$\widetilde{P}_{k+1} = \left(A - \kappa_k \times H\right) \times \widetilde{P}_k \times A^T + \Omega_k \tag{5.33}$$

The essence of what is being proposed here is the use of functions that are obtained heuristically to adjust the optimal filter according to varying uncertainty levels and noise.

Let equation (5.32) be substituted into equation (5.33), then:

$$\widetilde{P}_{k+1} = \left(A - \kappa_k \times H\right) \times \widetilde{P}_k \times A^T + \kappa_k \times \hat{P}_k \times \kappa_k^T \tag{5.34}$$

where further to equation (5.23), $\hat{P}_k$ is the Fuzzy-Kalman Filter measurement error covariance defined in terms of Φ_k as:

$$\hat{P}_k = H \times \widetilde{P}_k \times H^T + \Phi_k \tag{5.35}$$

with $\widetilde{P}_k$ being the Fuzzy-Kalman projectile error covariance matrix as defined in equation (5.33).

When equation (5.35) is substituted in equation (5.34), then:

$$\widetilde{P}_{k+1} = \left(A - \kappa_k \times H\right) \times \widetilde{P}_k \times A^T + \kappa_k \times \left(H \times \widetilde{P}_k \times H^T + \Phi_k\right) \times \kappa_k^T \tag{5.36}$$

Further to the heuristic selection of Φ_k and Ω_k as given by equations (5.24) and (5.32), the optimality of the state estimation strategy is maintained with respect to the future error such that $\frac{\partial \widetilde{P}_{k+1}}{\partial \kappa_k} = 0$. From equation (5.36), for optimality.

$$\frac{\partial \widetilde{P}_{k+1}}{\partial \kappa_k} = -H \times \widetilde{P}_k \times A^T + \left(H \times \widetilde{P}_k \times H^T + \Phi_k\right) \times \kappa_k^T = 0 \tag{5.37}$$

Rearranging equation (5.37), then:

$$\kappa_k = \left(H \times \widetilde{P}_k \times A^T\right)^T \times \left\{\left(H \times \widetilde{P}_k \times H^T + \Phi_k\right)^{-1}\right\}^T = A \times \widetilde{P}_k \times H^T \times \left(H \times \widetilde{P}_k \times H^T + \Phi_k\right)^{-1} \tag{5.38}$$

Equation (5.38) is identical to equation (5.25), thus confirming the optimality of the gain κ_k with respect to $\widetilde{P}_{k+1}$. The revised estimation process with heuristic and fuzzy membership functions Φ_k and Ω_k is provided in Figure 5.1.

Step 1 Calculation of the Fuzzy-Kalman gain by using the projectile or initial condition

$$\kappa_k = A \times \widetilde{P}_k \times H^T \times \left(H \times \widetilde{P}_k \times H^T + \Phi_k\right)^{-1}$$

$$\Downarrow$$

Step 2 Correction of the projectile estimate in its a priori or a pre-processor form

$$\widetilde{X}_{k+1}^{-} = A \times \widetilde{X}_k + \kappa_k \times \left(z_k - H \times \widetilde{X}_k\right)$$

$$\Downarrow$$

Step 3 Calculation of the projectile state estimate for the next iteration cycle

$$\widetilde{X}_{k+1} = \widetilde{X}_{k+1}^{-} + B \times U_k$$

$$\Downarrow$$

Step 4 Calculation of the projectile error covariance matrix for obtaining the projectile estimate for the next iteration cycle

$$\Omega_k = \kappa_k \times \hat{P}_k \times \kappa_k^T$$

$$\widetilde{P}_{k+1} = \left(A - \kappa_k \times H\right) \times \widetilde{P}_k \times A^T + \Omega_k$$

$$\Downarrow$$

Step 5 Calculation of the measurement error covariance using the projectile estimation error covariance

$$\hat{P}_{k+1} = H \times \widetilde{P}_{k+1} \times H^T + \Phi_{k+1}$$

Iteration

Figure 5.1 Fuzzy-Kalman Filter Process

5.8 Fuzzy Kalman Filter Implementation in a "Parallel Distributed Compensator" Form (PDC-Fuzzy Kalman Filter)

Further use of fuzzy concept in state estimation is possible through the use of the "parallel distributed compensator" that is a heuristic method employed by Passino, [6]. The structure of the corrective gain is changed as a result to that of equation (5.6). Further to equation (5.6), the function ψ_k is used to relate and scale the corrective action κ_k of the Fuzzy Kalman gain according to the error $\left(z_k - H \times \widetilde{X}_k\right)$. As such, let ψ_k be heuristically defined in terms of the measurement error covariance matrix as:

$$\psi_k = \sqrt{\hat{P}_k} \times H \tag{5.39}$$

where $\hat{P}_k$ is the Fuzzy-Kalman measurement error covariance matrix defined in equation (5.35). Substituting equation (5.39) in equation (5.6), then:

$$\widetilde{K}_k \times \left(z_k - H \times \widetilde{X}_k\right) = \kappa_k \times \sqrt{\hat{P}_k} \times H \times \widetilde{X}_k \tag{5.40}$$

The Fuzzy-Kalman pre-processor estimate is obtained by substituting equation (5.39) in equation (5.8) such that:

$$\widetilde{X}_{k+1}^{-} = \left\{A + \kappa_k \times \psi\left(e_{z_k}\right)\right\} \times \xi\left(e_{z_k}\right) \times \widetilde{X}_k = \left\{A + \kappa_k \times \sqrt{\hat{P}_k} \times H\right\} \times \xi\left(e_{z_k}\right) \times \widetilde{X}_k \tag{5.41}$$

In a nonlinear application ξ can be a membership function that would allow for model adjustment according to piecewise linear regions or error regions. Here a linear system is considered with $\xi = 1$. The process for the Parallel Distributed Compensator form of the Fuzzy-Kalman filter (PDC-Fuzzy-Kalman Filter) can now be summarized as given in Figure 5.2.

Step 1 Calculation of the Fuzzy-Kalman gain by using the projectile or initial condition

$$\kappa_k = A \times \widetilde{P}_k \times H^T \times \left(H \times \widetilde{P}_k \times H^T + \Phi_k \right)^{-1}$$

$$\Downarrow$$

Step 2 Correction of the projectile estimate in its a priori or a pre-processor form

$$\widetilde{X}_{k+1}^{-} = \left(A + \kappa_k \times \sqrt{\hat{P}_k} \times H \right) \times \widetilde{X}_k$$

$$\Downarrow$$

Step 3 Calculation of the projectile state estimate for the next iteration cycle

$$\widetilde{X}_{k+1} = \widetilde{X}_{k+1}^{-} + B \times U_k$$

$$\Downarrow$$

Step 4 Calculation of the projectile error covariance matrix for obtaining the projectile estimate for the next iteration cycle

$$\Omega_k = \kappa_k \times \hat{P}_k \times \kappa_k^T$$

$$\widetilde{P}_{k+1} = \left(A - \kappa_k \times H \right) \times \widetilde{P}_k \times A^T + \Omega_k$$

$$\Downarrow$$

Step 5 Calculation of the measurement error covariance using the projectile estimation error covariance

$$\hat{P}_{k+1} = H \times \widetilde{P}_{k+1} \times H^T + \Phi_{k+1}$$

Iteration

Figure 5.2 The PDC-Fuzzy-Kalman Filter Process

Chapter 6

Simulation Results and Discussion

In this chapter, the performance of the Kalman filter, the Revised-Kalman filter and the Fuzzy-Kalman filter are compared by using computer simulation. These methodologies are applied to the model of an Electro hydraulic actuator, [24].

6.1 The Model System: The Electro Hydraulic Actuator

An ElectroHydraulic Actuator has been described and extensively studied in [24]. The mathematical model of this actuator is used in a linearized form in this study.

Let the state space equation of this actuator be specified as:

$$X_{k+1} = A \times X_k + B \times U_k + w_k$$

The elements of the state vector correspond to the actuator's position (cm), velocity (cm/s), and acceleration (cm/s^2). The initial value of the state vector is assumed as:

$$X_0 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \tag{6.1}$$

The system matrix A is obtained as, [24]:

$$A = \begin{bmatrix} 1 & 0.001 & 0 \\ 0 & 1 & 0.001 \\ -522.02 & -28.616 & 0.9418 \end{bmatrix} \tag{6.2}$$

B is the input matrix and specified as:

$$B = \begin{bmatrix} 0 \\ 0 \\ 542.02 \end{bmatrix} \tag{6.3}$$

The system noise is assumed white with a maximum amplitude of 0.001, such that:

$\max(|w_k|) = \begin{bmatrix} 0.001 \\ 0.001 \\ 0.001 \end{bmatrix}$. Random signal with a maximum amplitude of the value of 1 is used as an input, U. The output equation of the actuator is specified as:

$$z_k = H \times X_k + v_k \tag{6.4}$$

where the output matrix H is pseudo-diagonal such that:

$$H = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \tag{6.5}$$

Therefore, the output z_k has elements corresponding to the measured position and velocity. Let the sampling time Δ be 0.001 second. The measurement noise is assumed white with an upper amplitude bound of 0.1:

$$\max(|v_k|) = \begin{bmatrix} 0.1 \\ 0.1 \end{bmatrix}$$

6.2 The Kalman Filter

The Kalman Filter estimation process is summarized in Figure 3.1 of Chapter 3. The initial condition of the state vector is specified as:

$$X_0 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

For this process, the measurement error covariance matrix is specified as:

$$R_k = \begin{bmatrix} 0.1 \times 0.1 & 0 \\ 0 & 0.1 \times 0.1 \end{bmatrix}$$

Similarly the system noise covariance matrix is specified as:

$$Q_k = \begin{bmatrix} 0.001 \times 0.001 & 0 & 0 \\ 0 & 0.001 \times 0.001 & 0 \\ 0 & 0 & 0.001 \times 0.001 \end{bmatrix}$$

The remaining initial conditions for the filter were specified as follows:

$$\hat{X}_0 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \text{ and } P_0^- = \begin{bmatrix} 10 & 0 & 0 \\ 0 & 10 & 0 \\ 0 & 0 & 10 \end{bmatrix}$$

The P_0^-, Q and R matrices were set by trial and error, a tuning method commonly used in the Kalman filter concept.

6.3 The Revised-Kalman Filter

The Revised-Kalman filter is summarized in Figure 4.1 of Chapter 4. It is applied to the system of section 6.1 with the following initial conditions:

$$R_k = \begin{bmatrix} 0.1 \times 0.1 & 0 \\ 0 & 0.1 \times 0.1 \end{bmatrix}, Q_k = \begin{bmatrix} 0.001 \times 0.001 & 0 & 0 \\ 0 & 0.001 \times 0.001 & 0 \\ 0 & 0 & 0.001 \times 0.001 \end{bmatrix}, \tilde{X}_0 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

and $\tilde{P}_0 = \begin{bmatrix} 10 & 0 & 0 \\ 0 & 10 & 0 \\ 0 & 0 & 10 \end{bmatrix}$

6.4 The Fuzzy-Kalman Filter and the PDC Fuzzy-Kalman Filter

The respective process of Fuzzy-Kalman Filter and the PDC Fuzzy-Kalman Filter are summarized in Figures 5.1 and 5.2 of Chapter 5. The initial projectile estimate, projectile estimate error covariance matrix, and the initial measurement error covariance matrix for these filters are specified as:

$$\widetilde{X}_0 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},\ \widetilde{P}_0 = \begin{bmatrix} 0.001\times 0.001 & 0 & 0 \\ 0 & 0.001\times 0.001 & 0 \\ 0 & 0 & 0.001\times 0.001 \end{bmatrix},\ \Phi_0 = \begin{bmatrix} 0.1\times 0.1 & 0 \\ 0 & 0.1\times 0.1 \end{bmatrix}$$

and $\hat{P}_0 = \begin{bmatrix} 0.1\times 0.1 & 0 \\ 0 & 0.1\times 0.1 \end{bmatrix}$. The initial condition pertaining to ψ of the PDC Fuzzy-Kalman Filter is set to:

$$\psi_0 = \sqrt{\hat{P}_0} \times H = \sqrt{\begin{bmatrix} 0.1\times 0.1 & 0 \\ 0 & 0.1\times 0.1 \end{bmatrix}} \times \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 0.1 & 0 & 0 \\ 0 & 0.1 & 0 \end{bmatrix}.$$ The initial κ_0 is calculated by substituting $\widetilde{P}_0$ and Φ_0 in equation (5.25).

6.5 Simulation Results and Comparative Discussions

The simulation results from the implementation of the three filters are provided in this section. The first simulation is conducted by setting the covariance matrices to correctly represent the system noise and the measurement error as given in equations (5.32) and (5.24). Figure 6.1 depicts the input and the output signals. Through Figures 6.2 to 6.4, the simulated actual and estimated states pertaining to the Kalman filter, the Revised-Kalman filter and the Fuzzy-Kalman filter are presented. In Figures 6.5 to 6.7, the results are compared by expanded views. By simulation, it is shown that the state

estimates from the Kalman filter are marginally more accurate than the Fuzzy-Kalman filter. Similar results are presented for the PDC-Fuzzy-Kalman Filter in figures 6.8 to 6.12. The system noise and the measurement error are assumed to be within their assumed covariance level. In Figures 6.9 to 6.11, the simulated actual and estimated states pertaining to the Kalman filter, the Revised-Kalman filter and the PDC-Fuzzy-Kalman filter are compared by their expanded views. Again, it can be observed that state estimates from the Kalman filter are marginally more accurate than the PDC-Fuzzy-Kalman filter. The Fuzzy-Kalman filter and the PDC-Fuzzy-Kalman filter are compared by an expanded view of simulation results in Figure 6.12. This figure shows that Fuzzy-Kalman filter and the PDC-Fuzzy-Kalman filter performances are comparable.

The second simulation is conducted by increasing the measurement noise level to beyond its assumed covariance level as:

$$z_k = \widetilde{H} \times X_k, \tag{6.6}$$

where $\widetilde{H} = \begin{bmatrix} 1+\eta & 0 & 0 \\ 0 & 1+\eta & 0 \end{bmatrix}$. Equation (6.6) instead of equation (3.2) is used for output measurement. η is a random signal noise with an upper amplitude bound of 0.1. All four filters are applied to this new condition without further retuning. This simulation is conducted to reflect the assumption that real noise can go beyond its assumed level. The state estimation results from the filters are shown in Figures 6.13 to 6.21. It is observed that the estimates from the Fuzzy-Kalman filtering and the PDC-Fuzzy-Kalman filtering are more correct than the Kalman filter. From the simulation results, it can be observed that the accuracy of the Kalman filter and the Revised-Kalman filter are comparable.

The third simulation is conducted by changing the condition of measurement error as:

$$X_{k+1} = \widetilde{A} \times X_k + B \times U_k \tag{6.7}$$

$$z_k = \widetilde{H} \times X_k \tag{6.6}$$

where $\widetilde{A} = \begin{bmatrix} 1+\omega & 0.001 & 0 \\ 0 & 1+\omega & 0.001 \\ -522.02+\omega & -28.616+\omega & 0.9418+\omega \end{bmatrix}$. Equation (6.7) is used instead of equation (3.1) to calculate the actual states. ω is a random signal noise with an upper amplitude bound of 0.001. And equation (6.6) is used for its measurements. Figure 6.22 depicts the input and the output signals. Through Figures 6.23 to 6.36, the simulated states and the state estimation results pertaining to the Kalman filter, the Revised-Kalman filter, the Fuzzy-Kalman Filter, and the PDC-Fuzzy-Kalman filter are presented. By simulation, it is observed that state estimates from the PDC-Fuzzy-Kalman filter are more accurate than the Kalman filter.

For Comparing the Kalman Filters and the Fuzzy-Kalman Filter Estimations

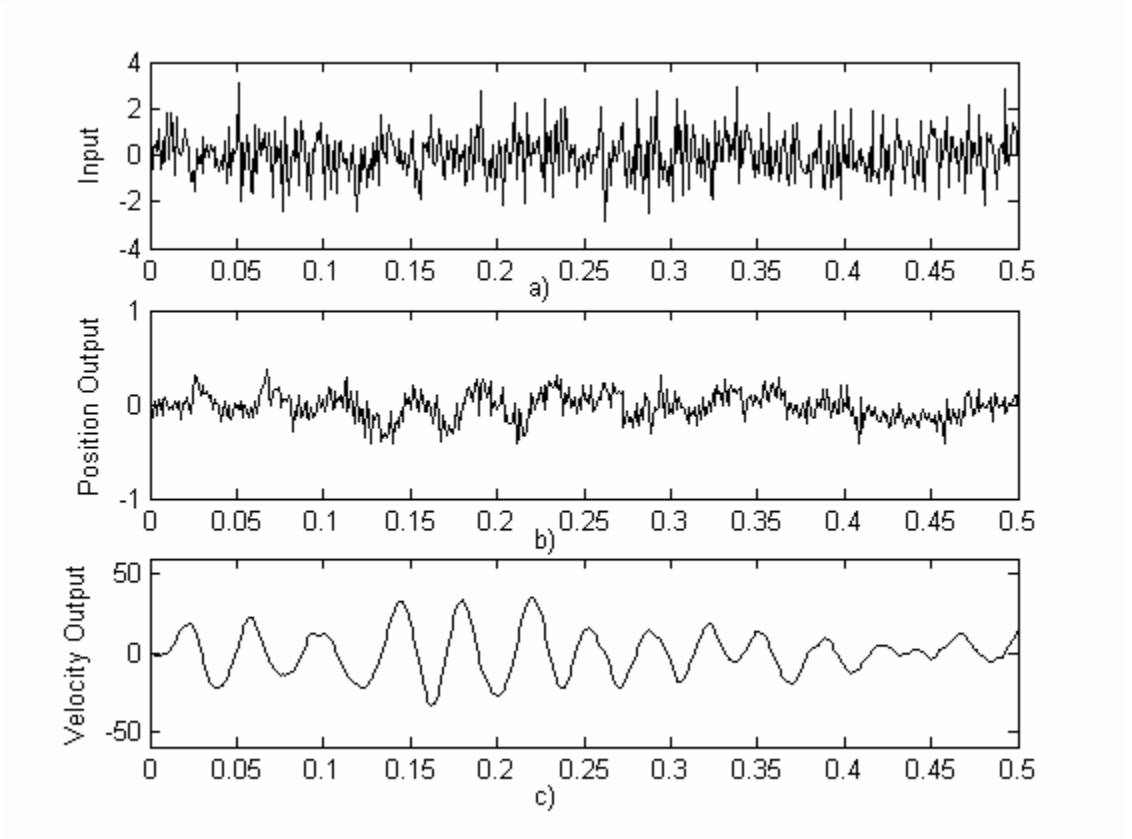


Figure 6.1 Input and Output Signals within the Assumed Noise Covariance

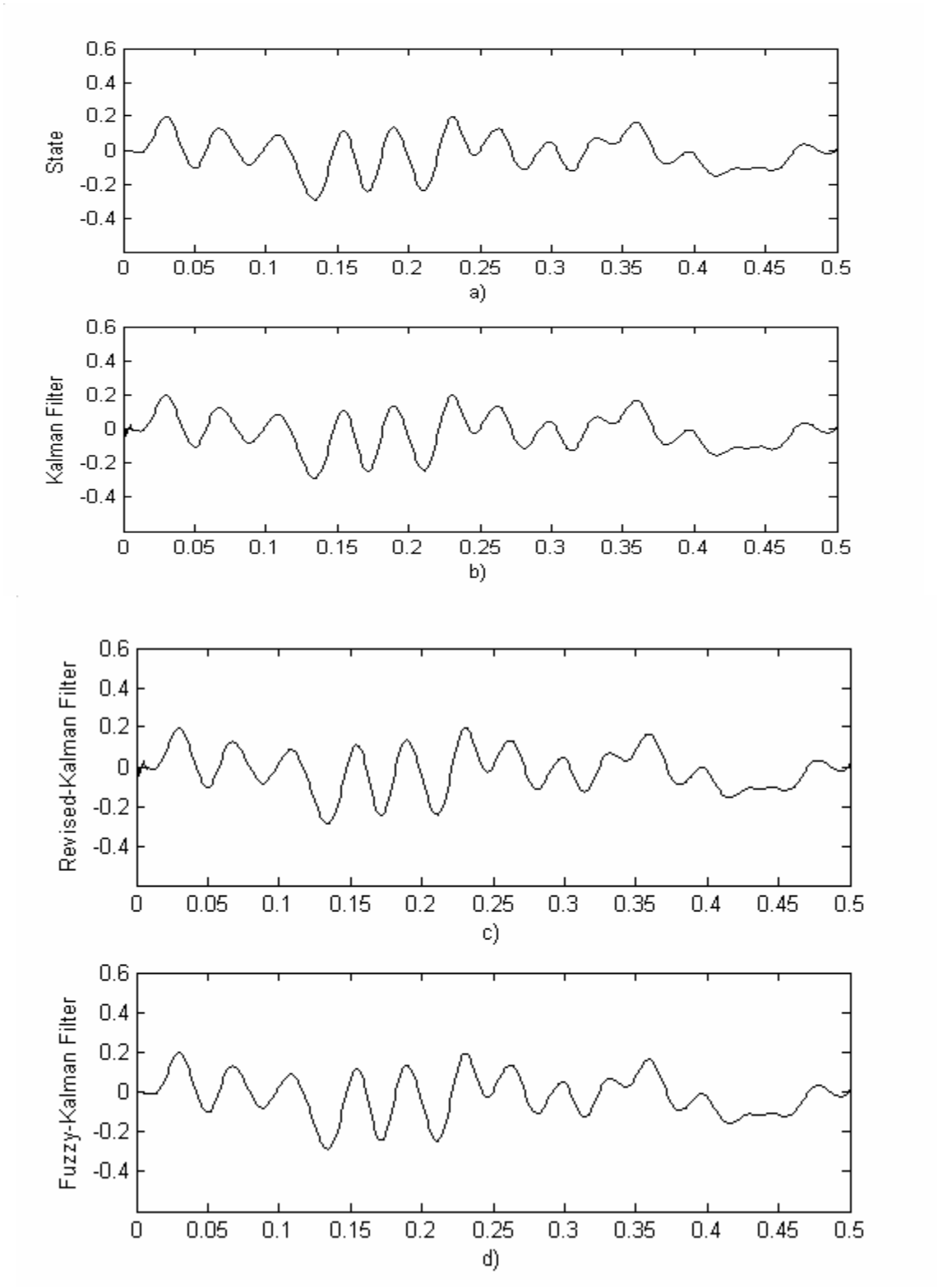


Figure 6.2 Position Estimate within the Assumed Noise Covariance (cm)

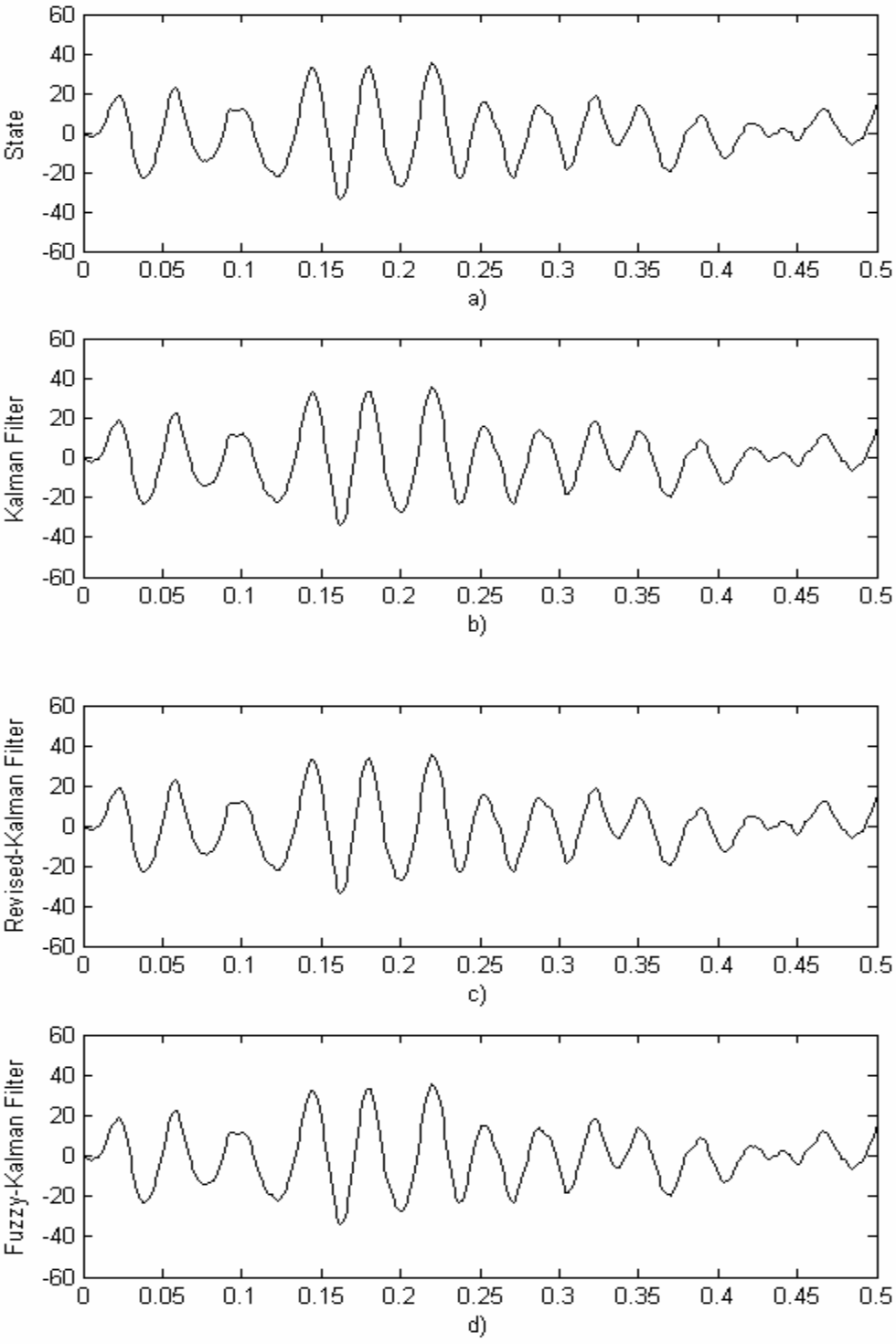


Figure 6.3 Velocity Estimate within the Assumed Noise Covariance (cm/s)

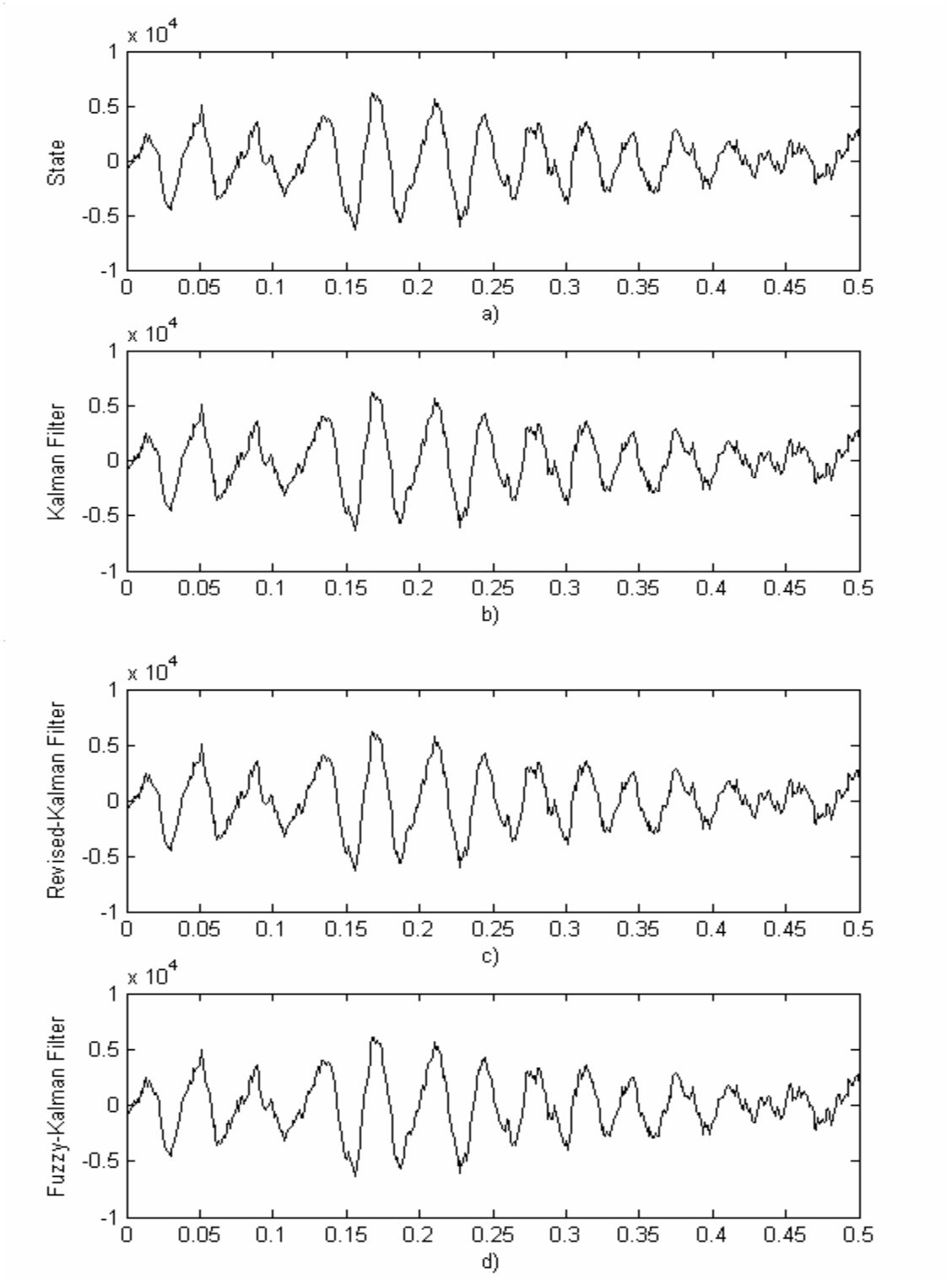


Figure 6.4 Acceleration Estimate within the Assumed Noise Covariance (cm/s^2)

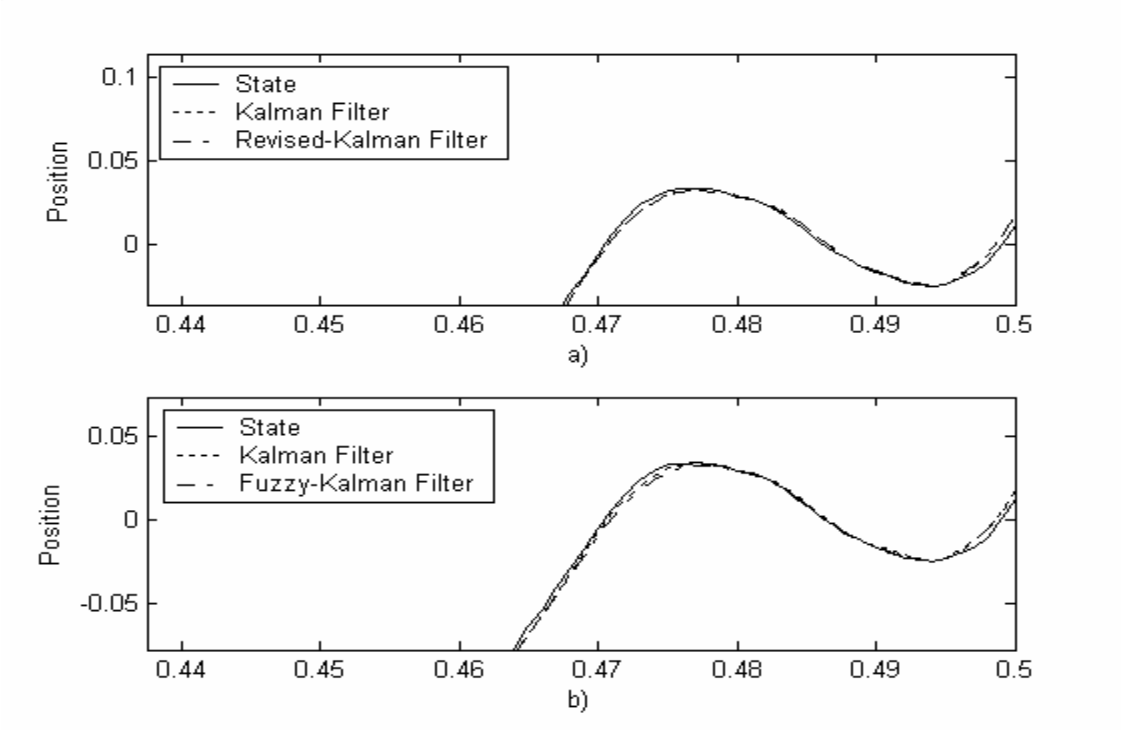


Figure 6.5 Expanded View for Comparing Position Estimations (cm)

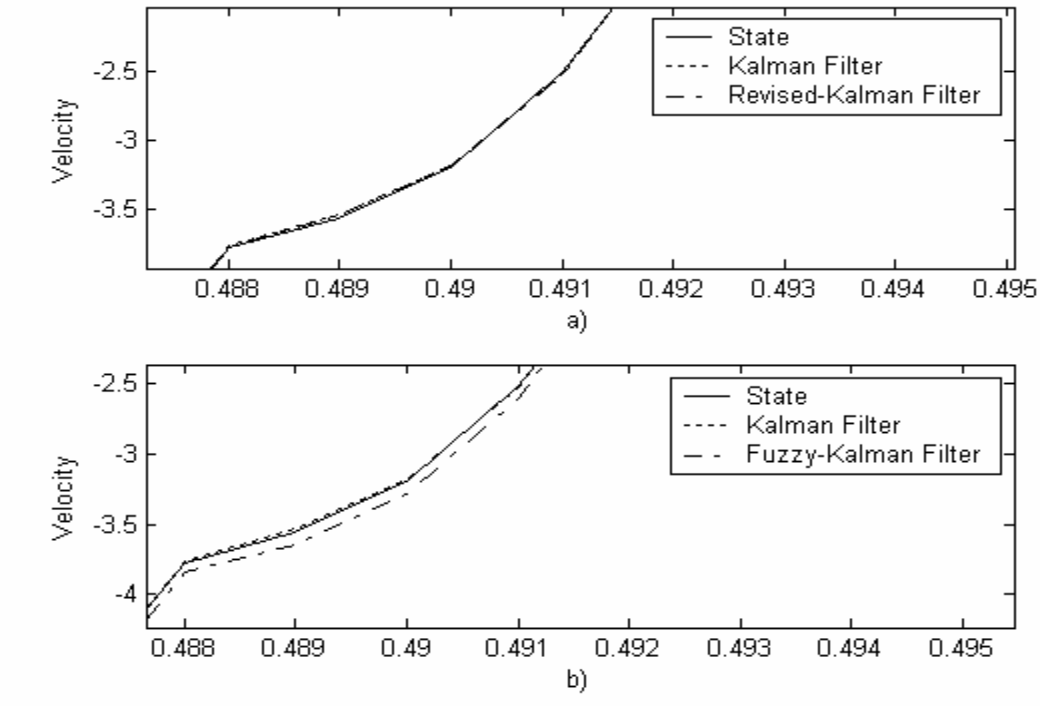


Figure 6.6 Expanded View for Comparing Velocity Estimations (cm/s)

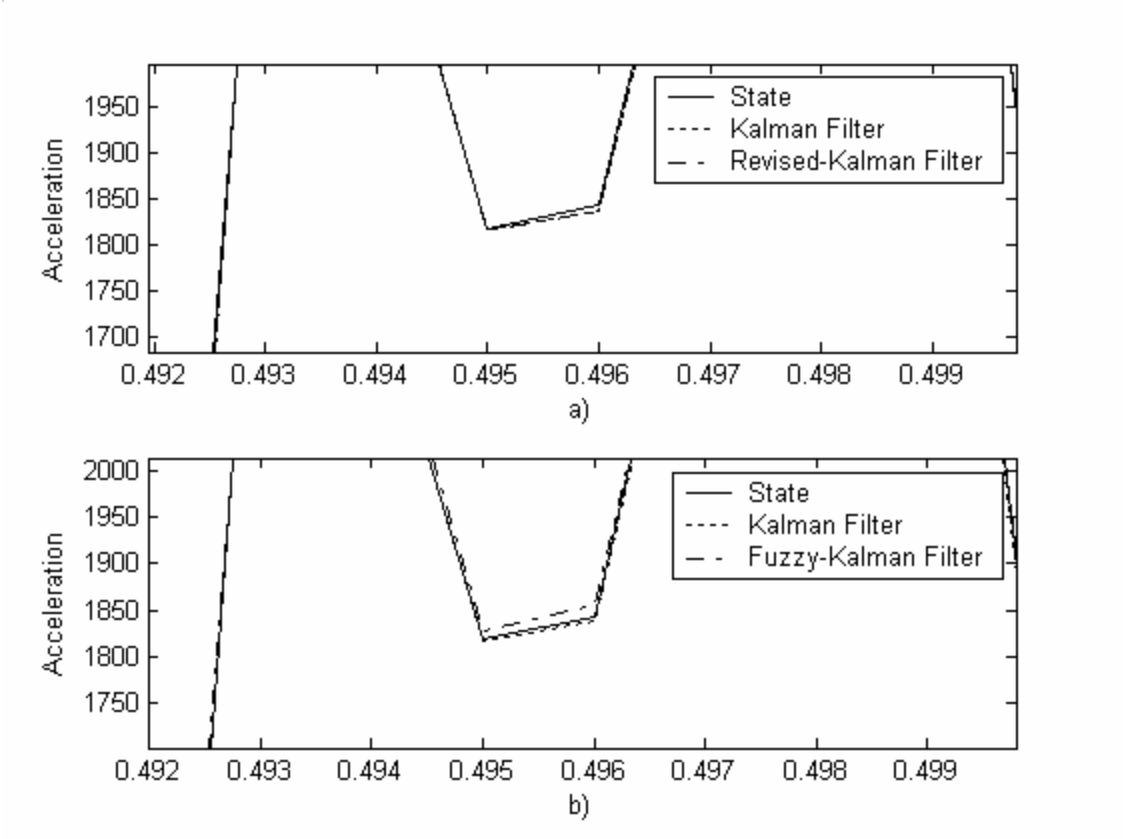


Figure 6.7 Expanded View for Comparing Acceleration Estimations (cm/s^2)

For Comparing the Kalman Filters and the PDC-Fuzzy-Kalman Filter Estimations

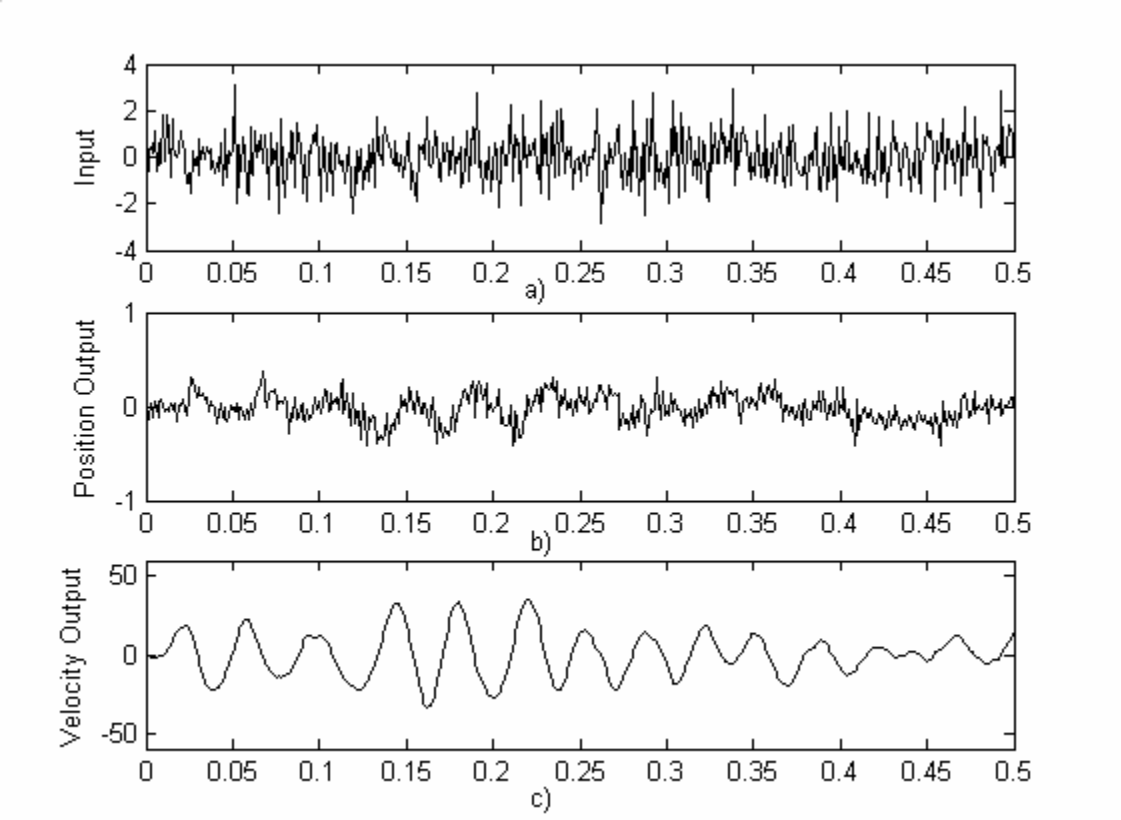


Figure 6.8 Input and Output Signals within the Assumed Noise Covariance

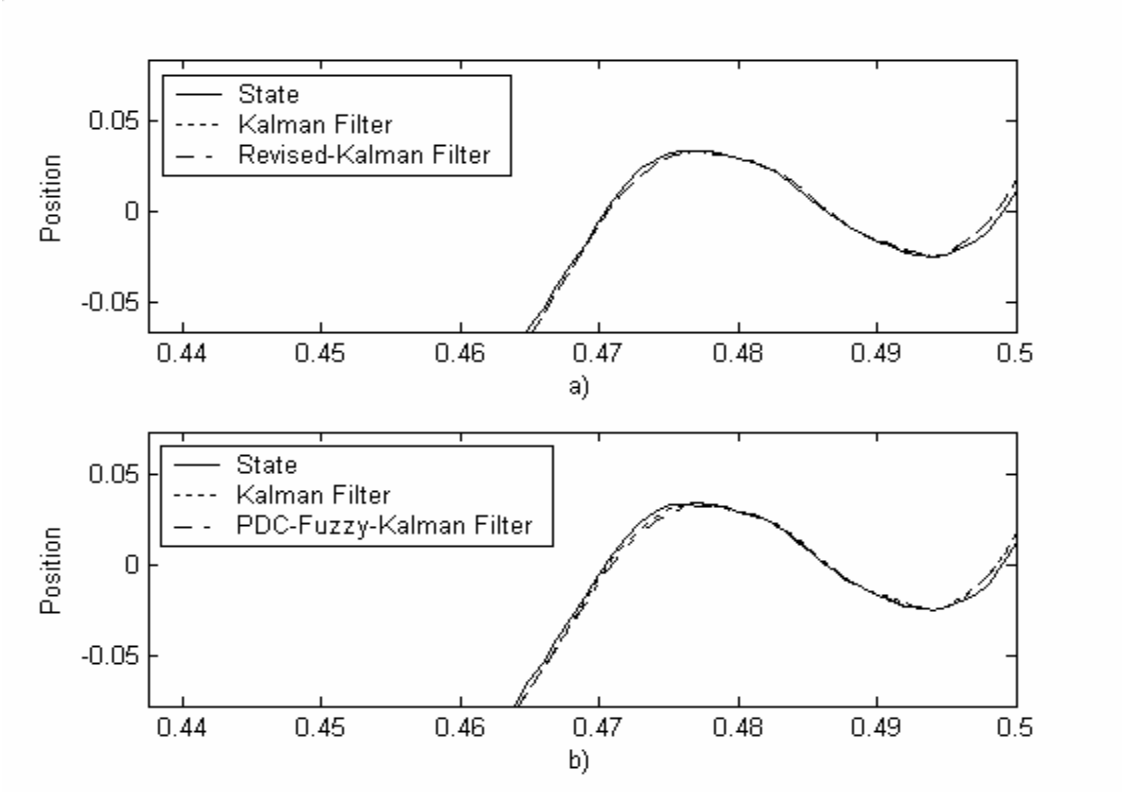


Figure 6.9 Expanded View for Comparing Position Estimations (cm)

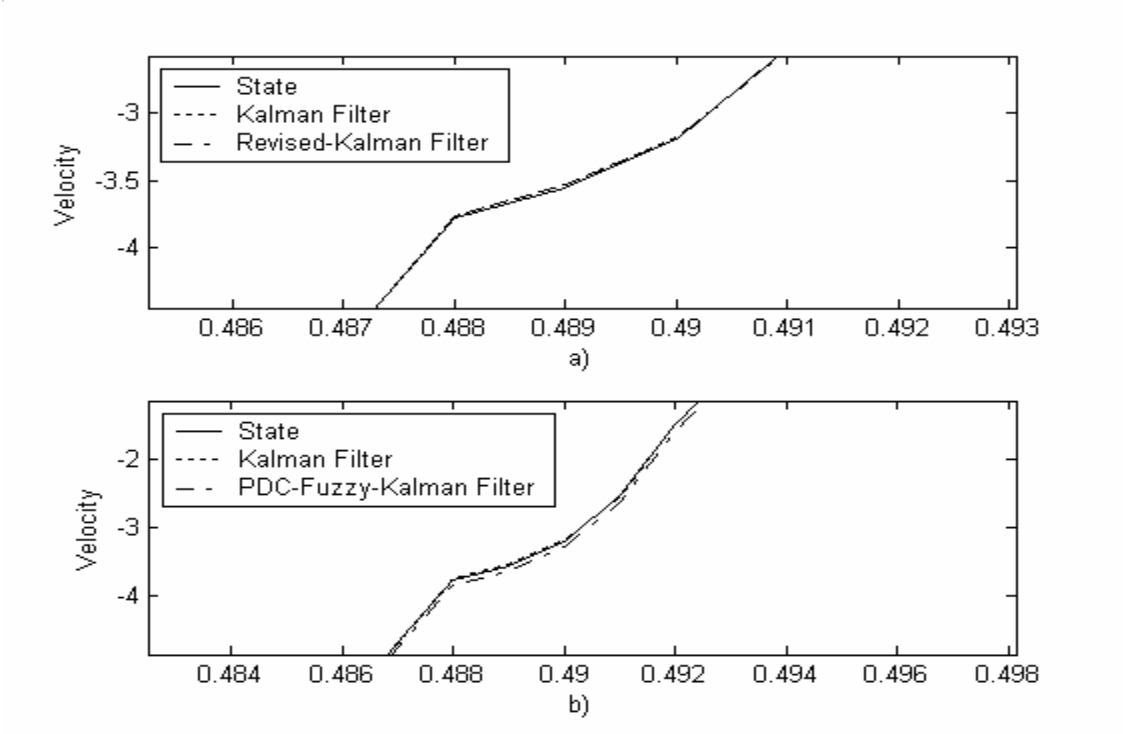


Figure 6.10 Expanded View for Comparing Velocity Estimations (cm/s)

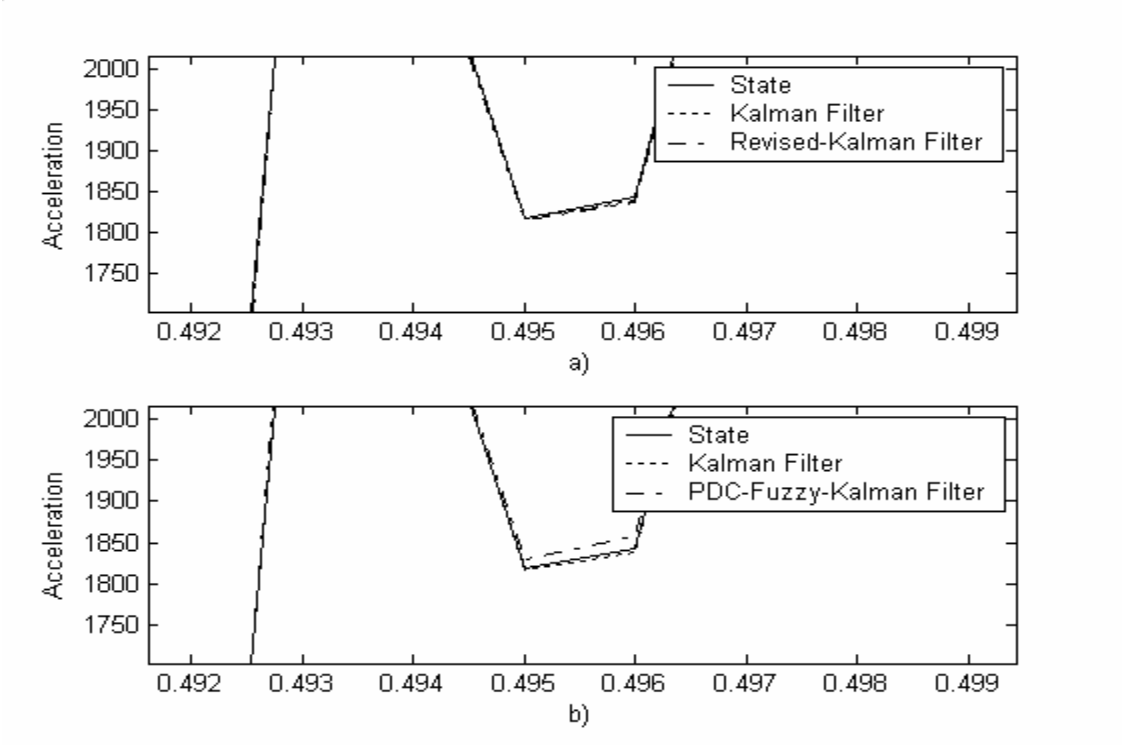


Figure 6.11 Expanded View for Comparing Acceleration Estimations (cm/s^2)

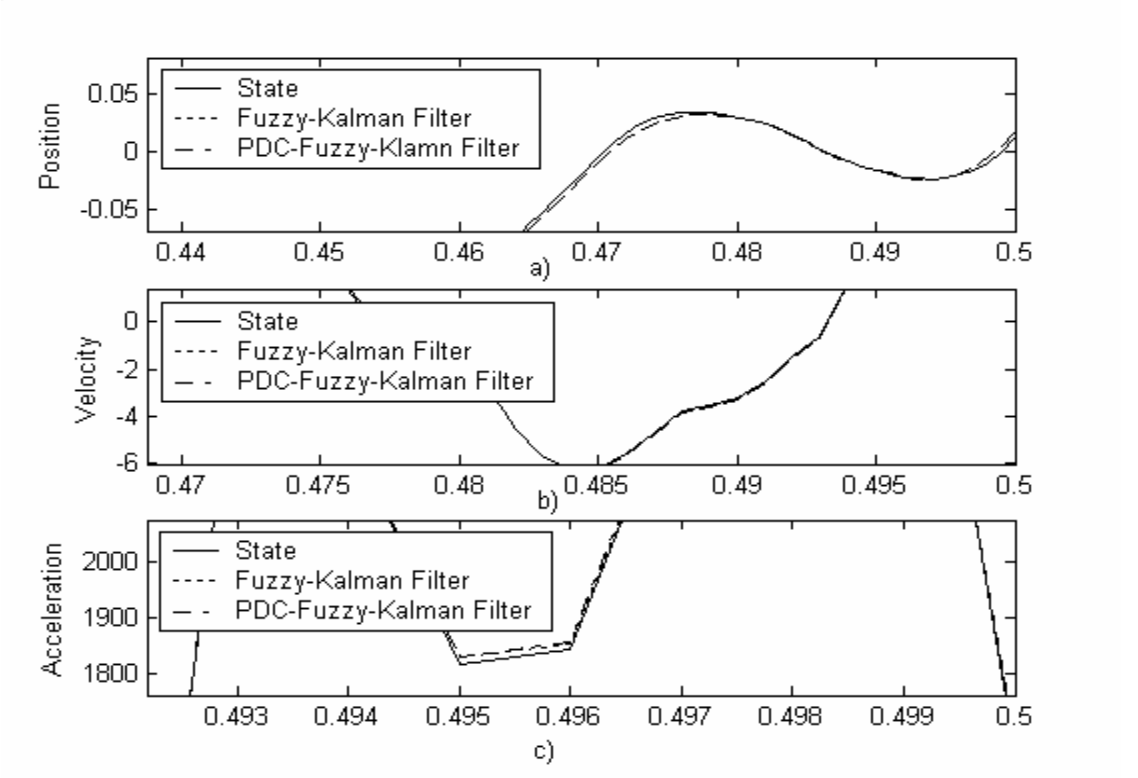


Figure 6.12 Expanded View for Comparing Two Fuzzy Kalman Filters

For Comparing the Kalman Filters and the Fuzzy-Kalman Filter Estimations

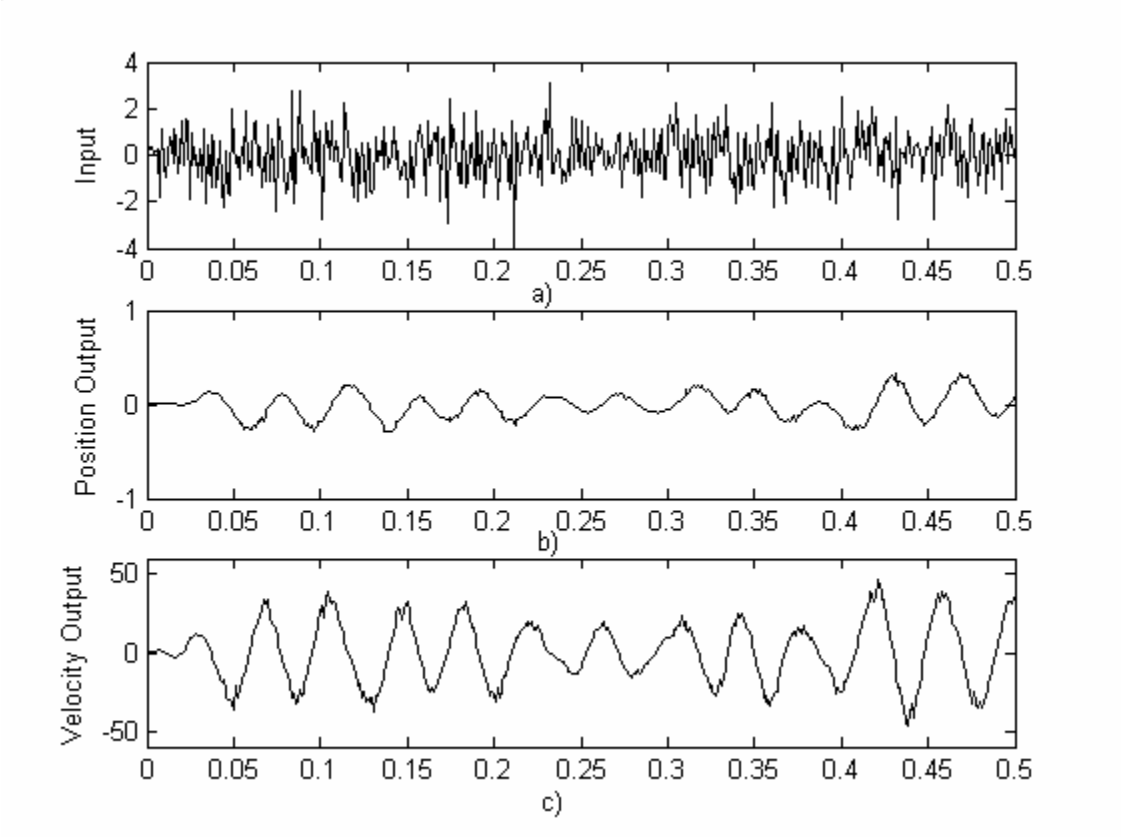


Figure 6.13 Input and Outputs beyond the Assumed Measurement Noise Covariance

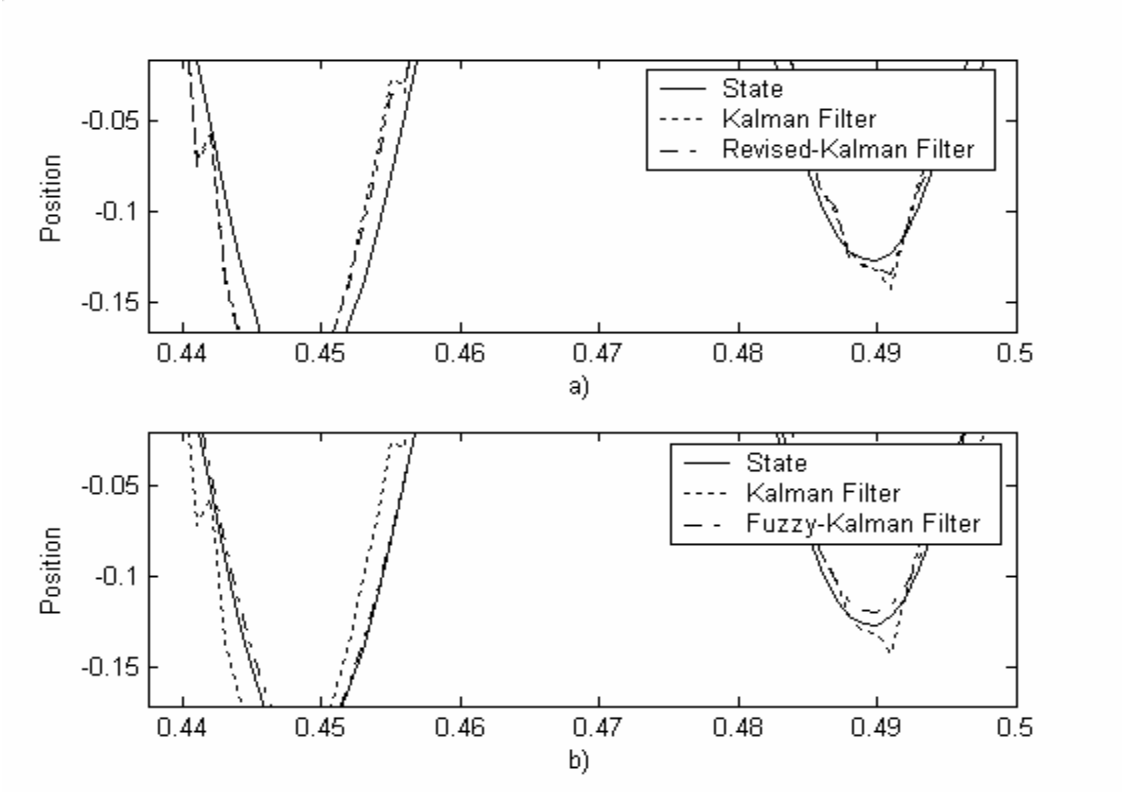


Figure 6.14 Expanded View for Comparing Position Estimations (cm)

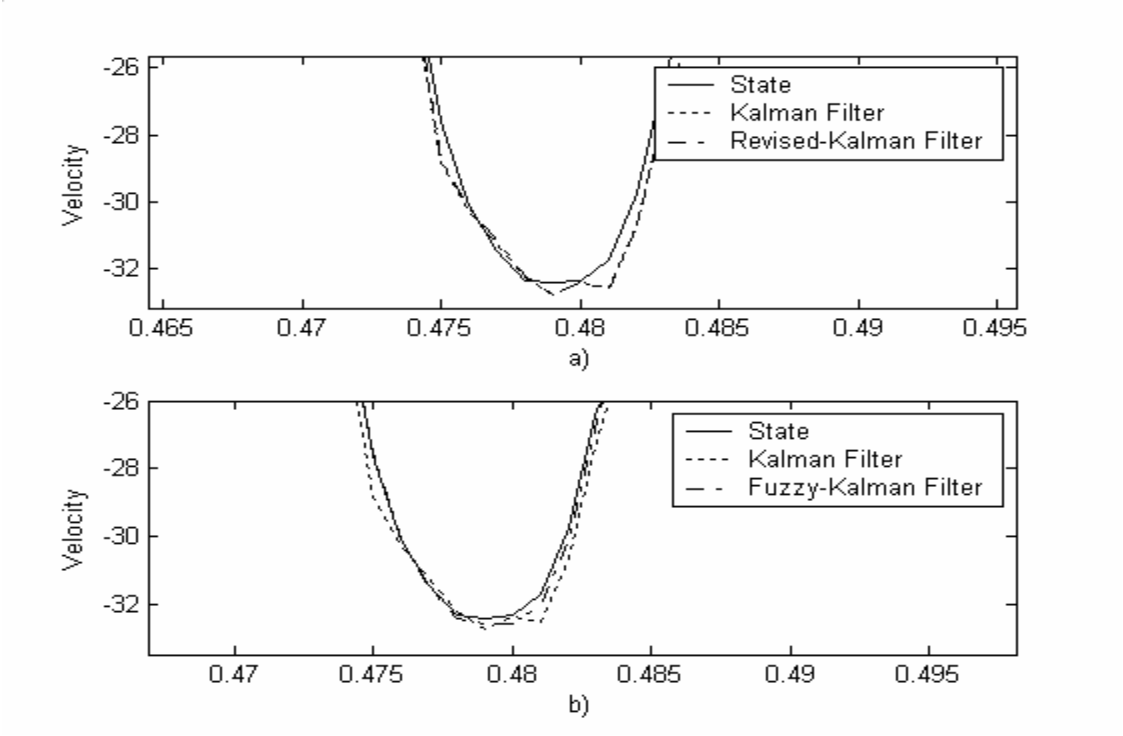


Figure 6.15 Expanded View for Comparing Velocity Estimations (cm/s)

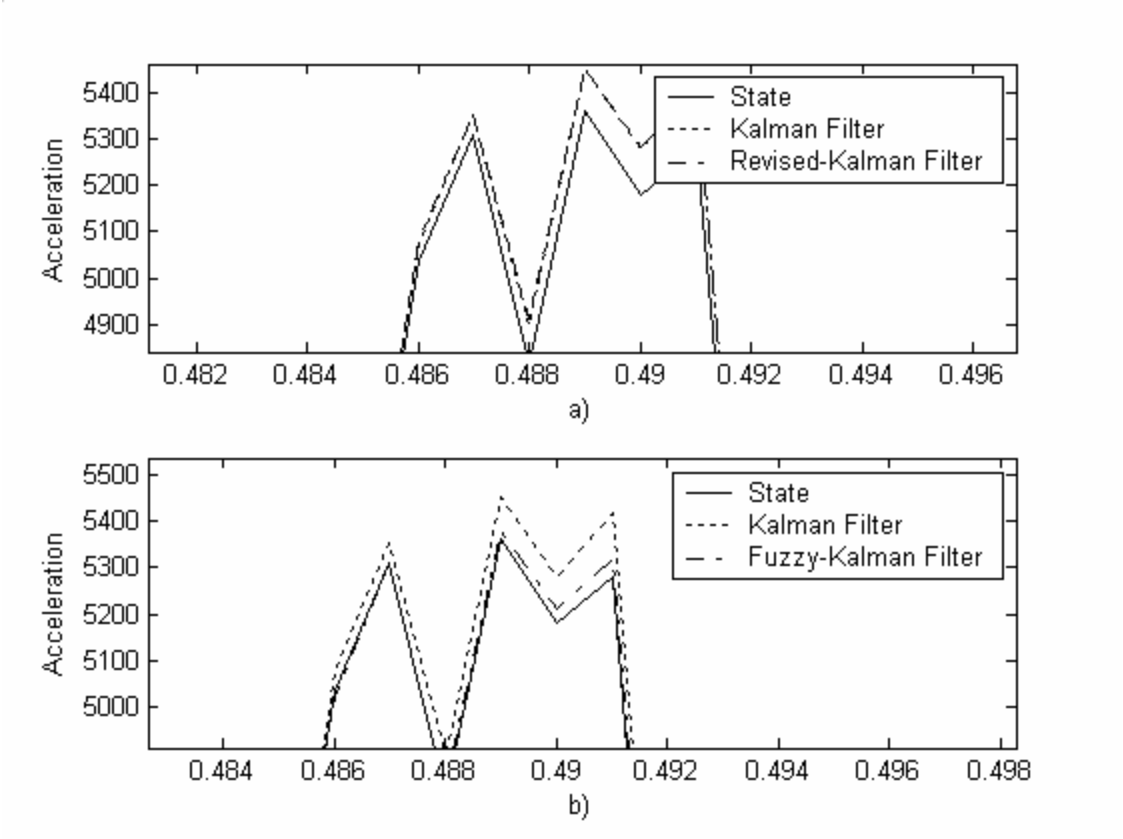


Figure 6.16 Expanded View for Comparing Acceleration Estimations (cm/s^2)

For Comparing the Kalman Filters and the PDC-Fuzzy-Kalman Filter Estimations

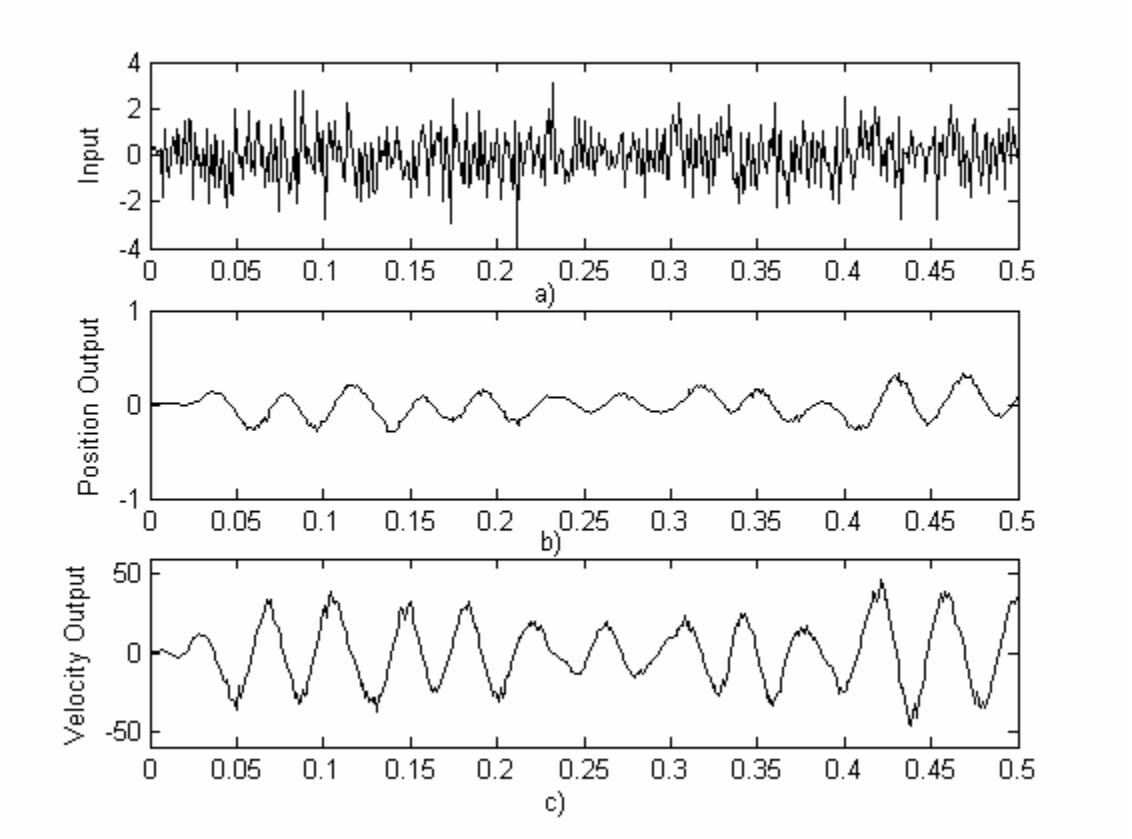


Figure 6.17 Input and Outputs beyond the Assumed Measurement Noise Covariance

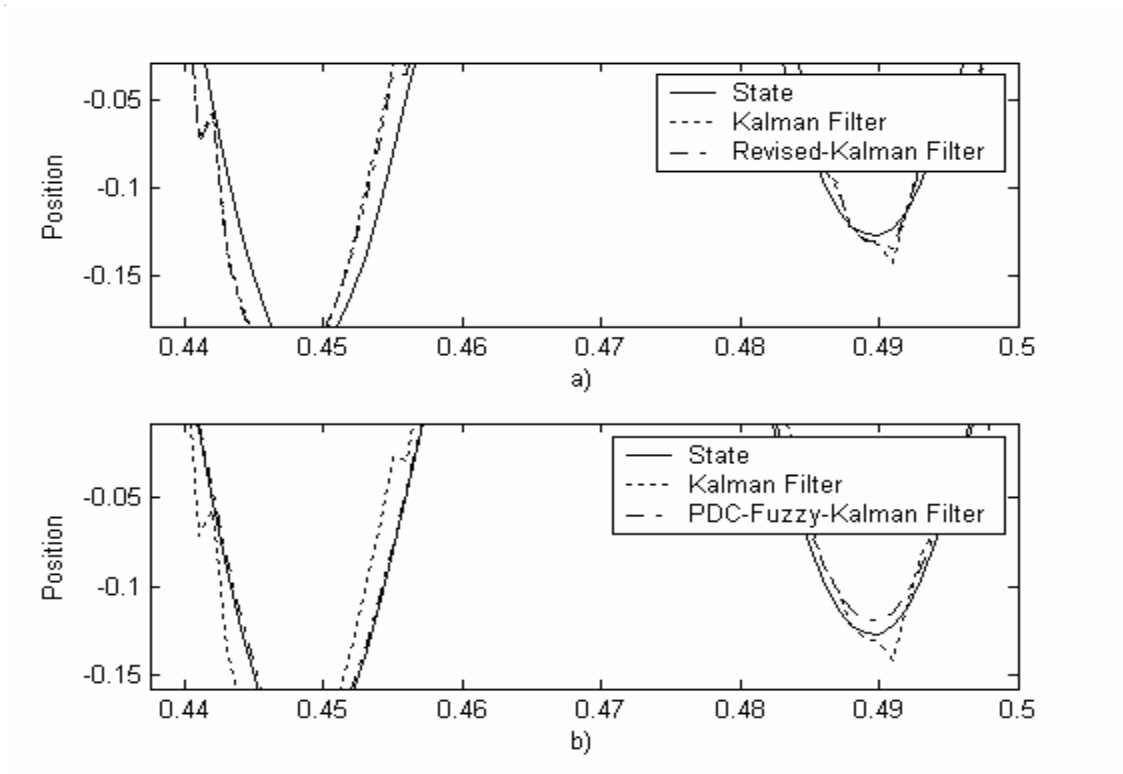


Figure 6.18 Expanded View for Comparing Position Estimations (cm)

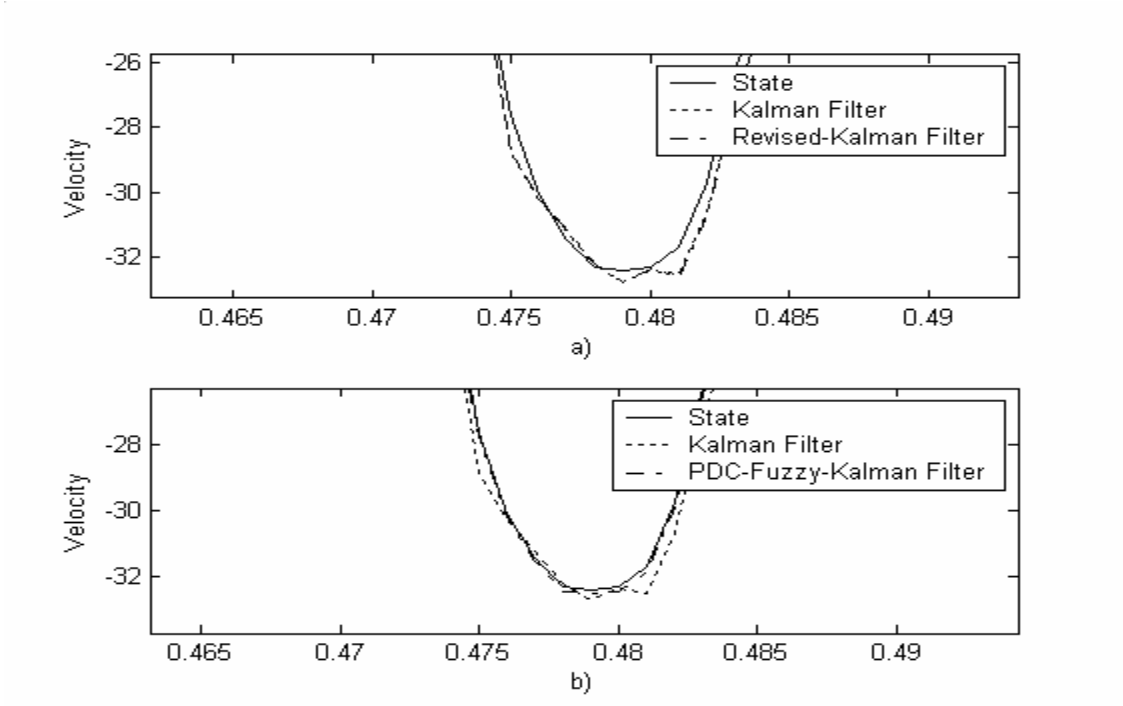


Figure 6.19 Expanded View for Comparing Velocity Estimations (cm/s)

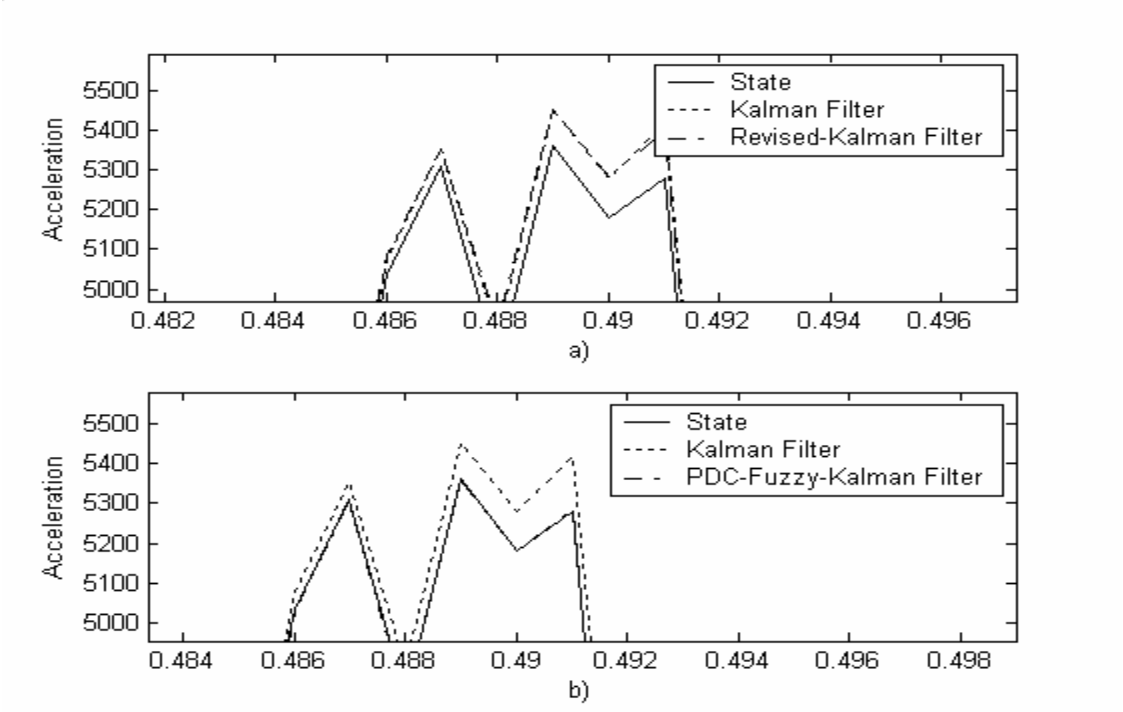


Figure 6.20 Expanded View for Comparing Acceleration Estimations (cm/s^2)

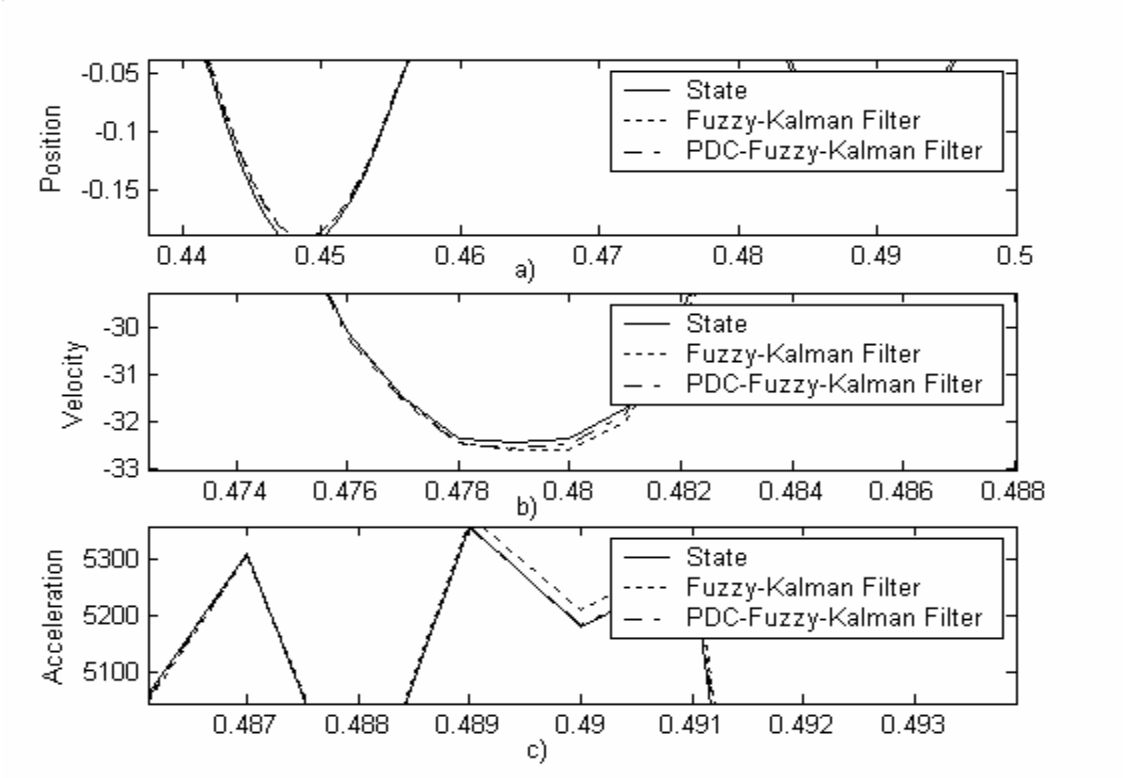


Figure 6.21 Expanded View for Comparing Two Fuzzy Kalman Filters

For Comparing the Kalman Filters and the Fuzzy-Kalman Filter Estimations

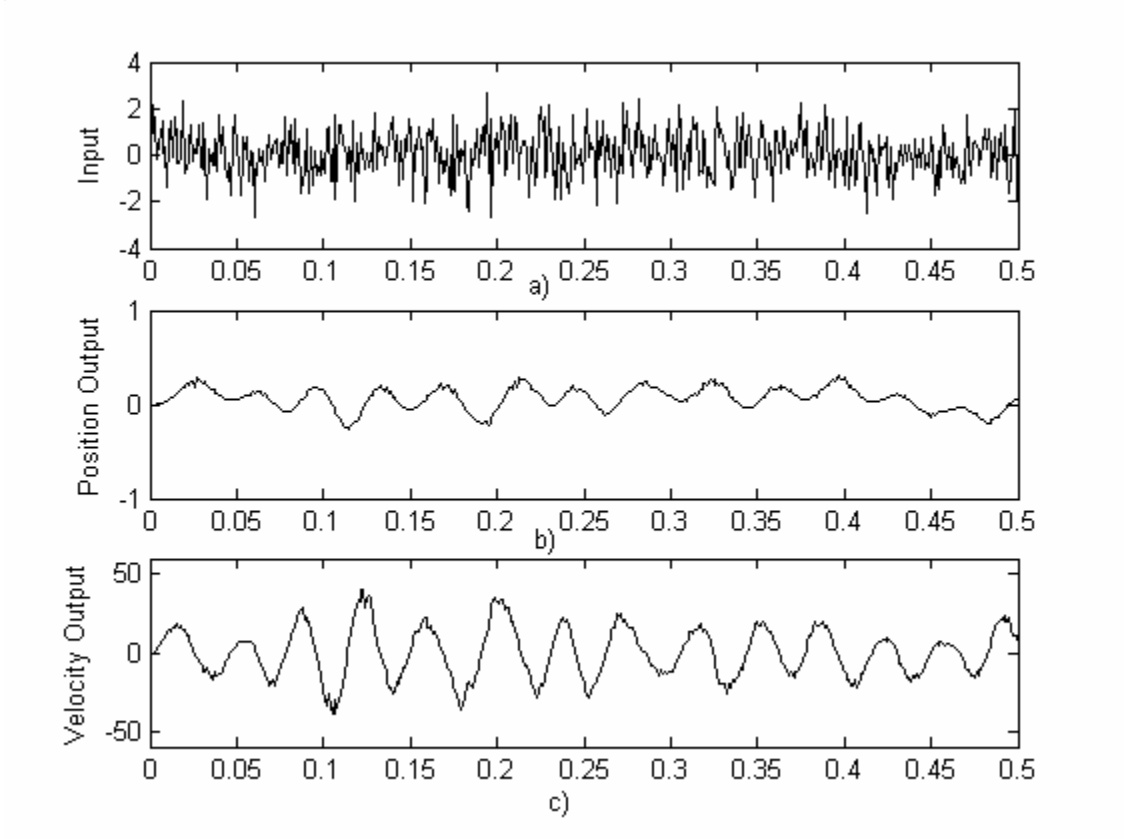


Figure 6.22 Input and Output Signals beyond the Assumed Noise Covariance

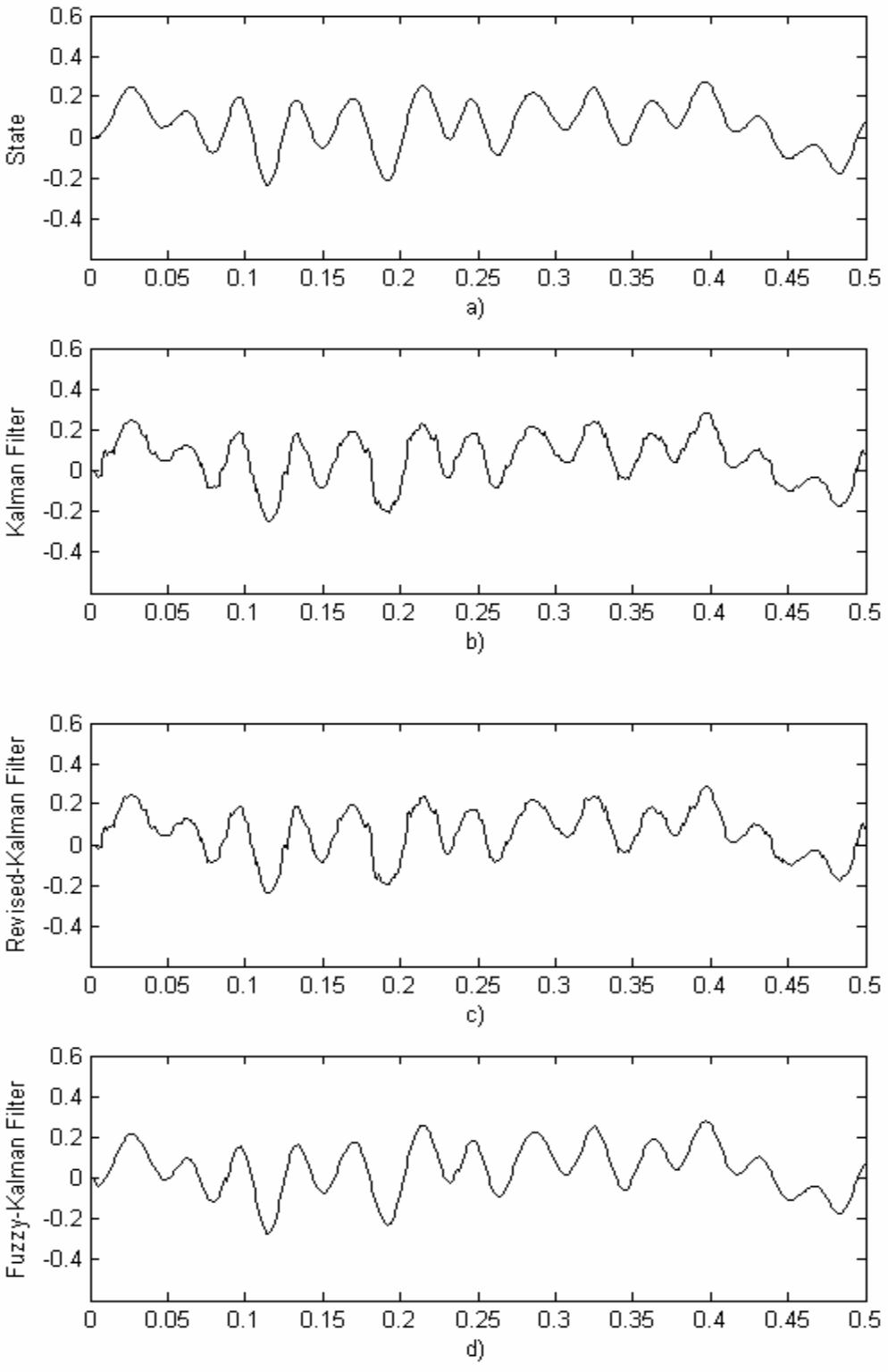


Figure 6.23 Position Estimate beyond the Assumed Noise Covariance (cm)

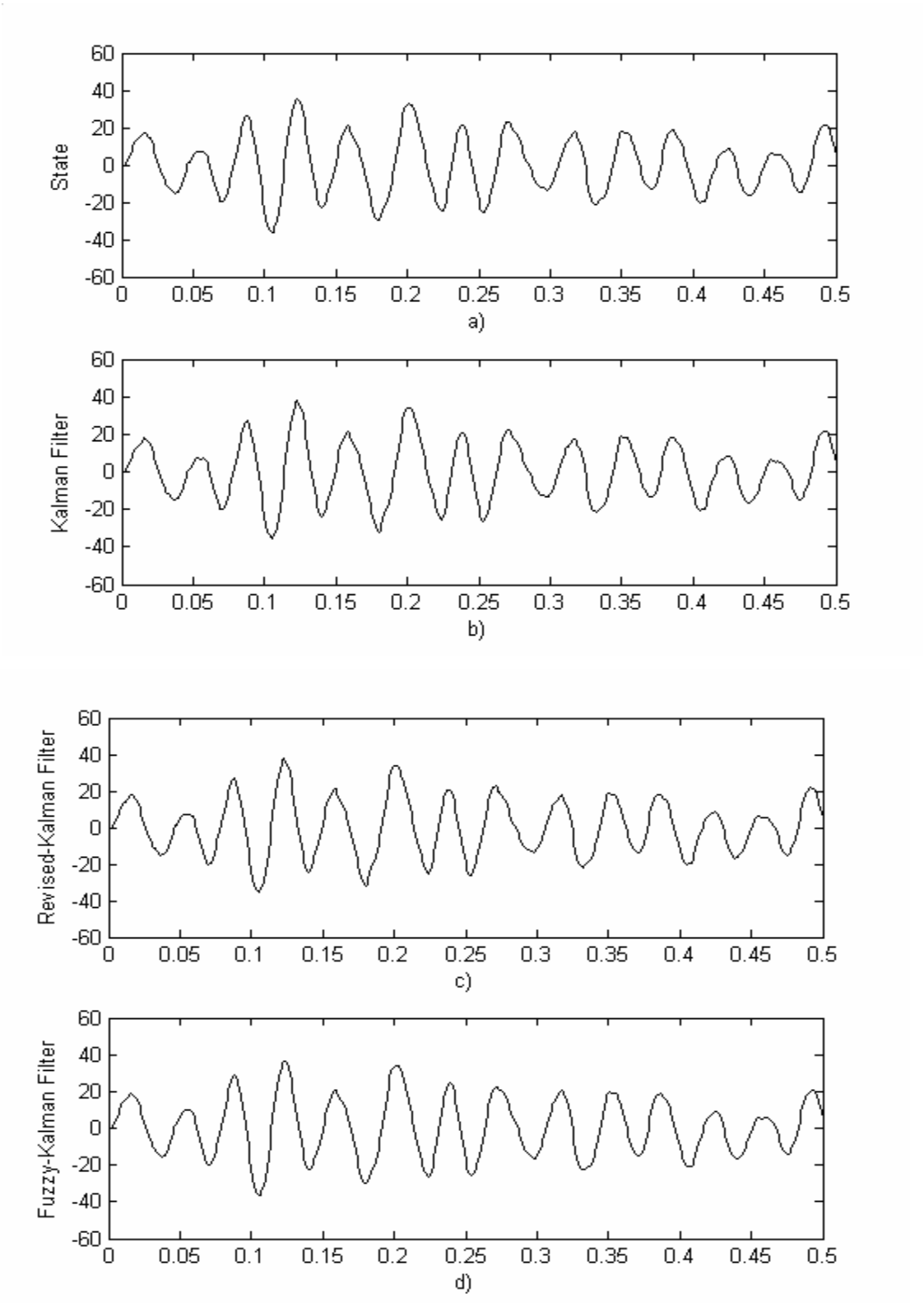


Figure 6.24 Velocity Estimate beyond the Assumed Noise Covariance (cm/s)

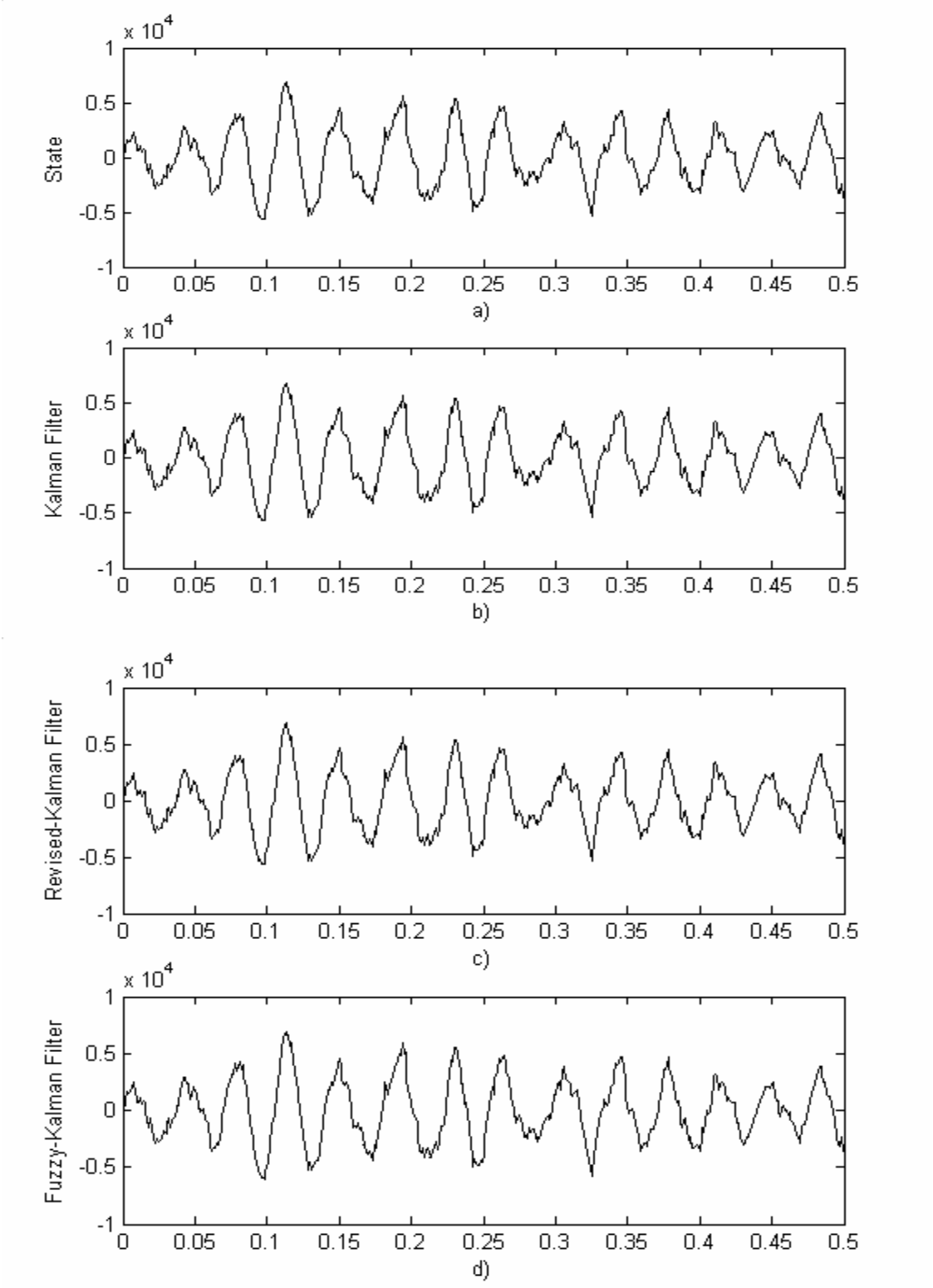
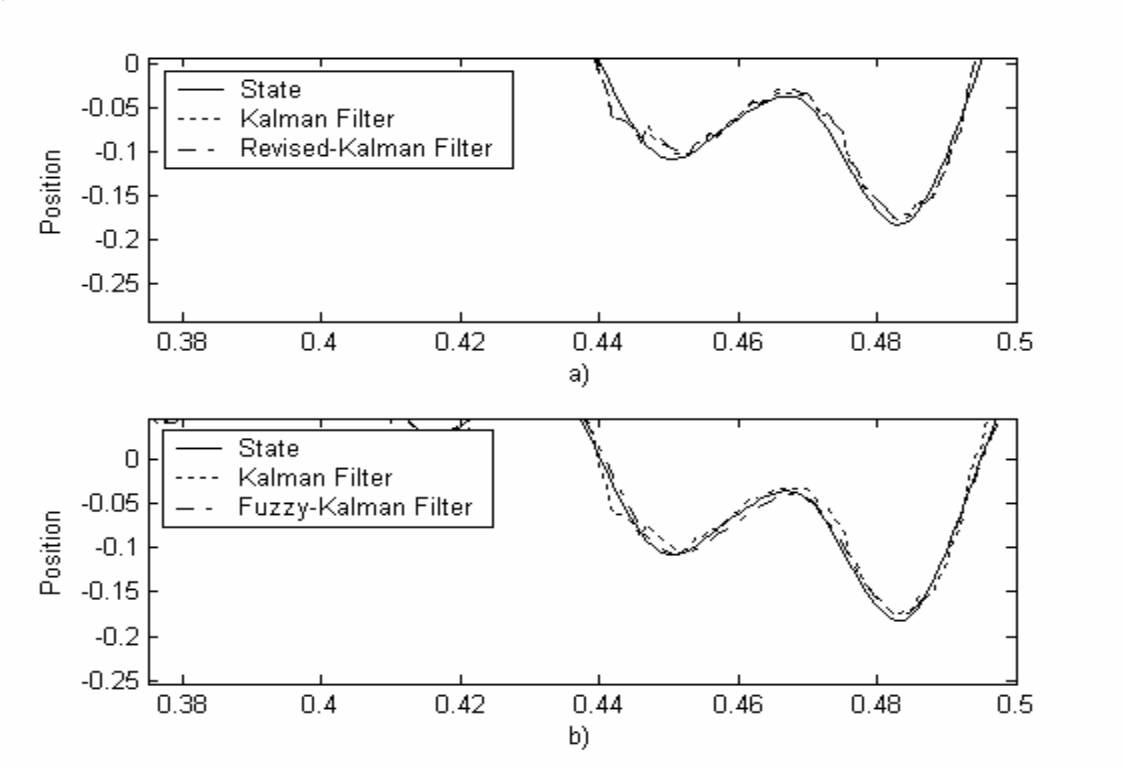
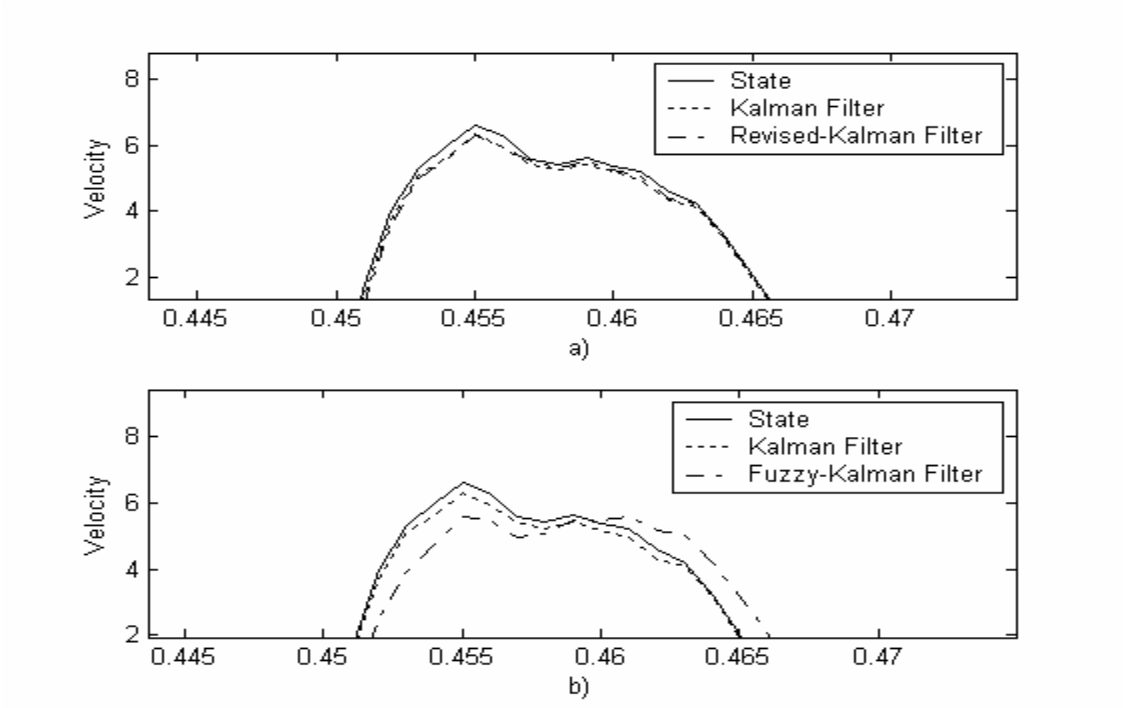


Figure 6.25 Acceleration Estimate beyond the Assumed Noise Covariance (cm/s^2)



Time (Sec)

Figure 6.26 Expanded View for Comparing Position Estimations (cm)



Time (Sec)

Figure 6.27 Expanded View for Comparing Velocity Estimations (cm/s)

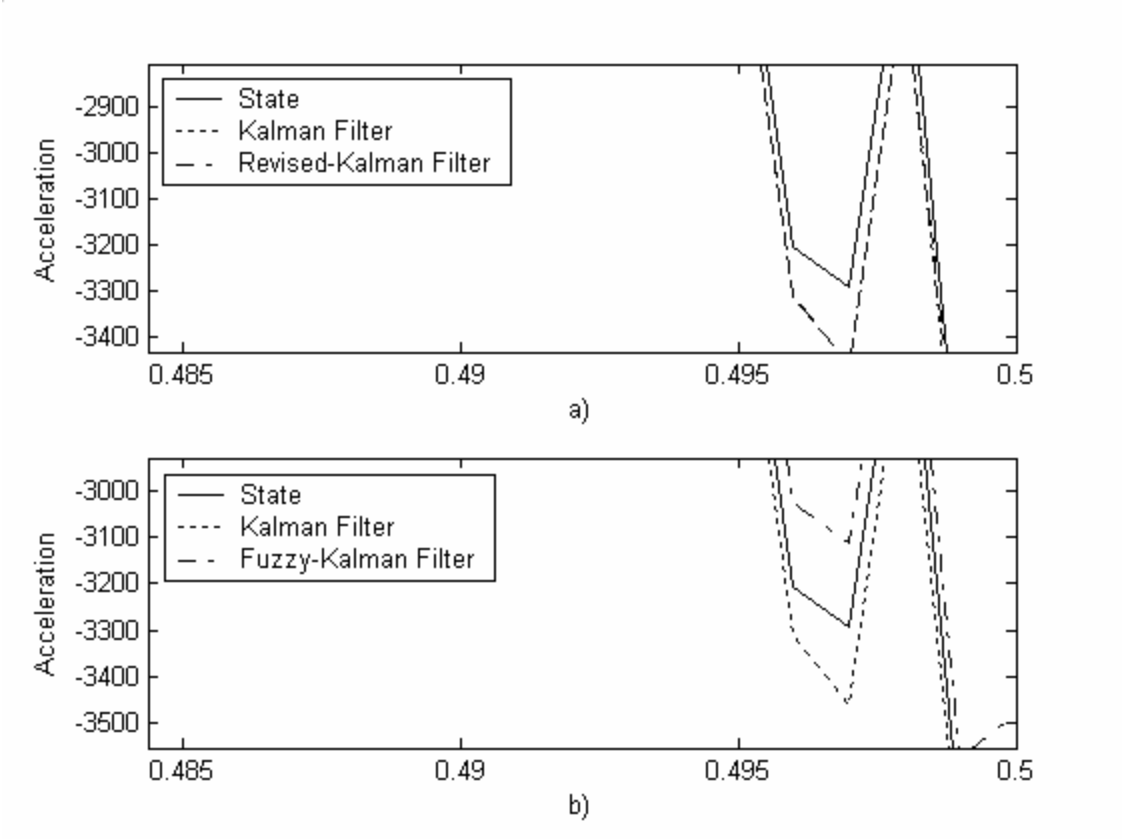


Figure 6.28 Expanded View for Comparing Acceleration Estimations (cm/s^2)

For Comparing the Kalman Filters and the PDC-Fuzzy-Kalman Filter Estimations

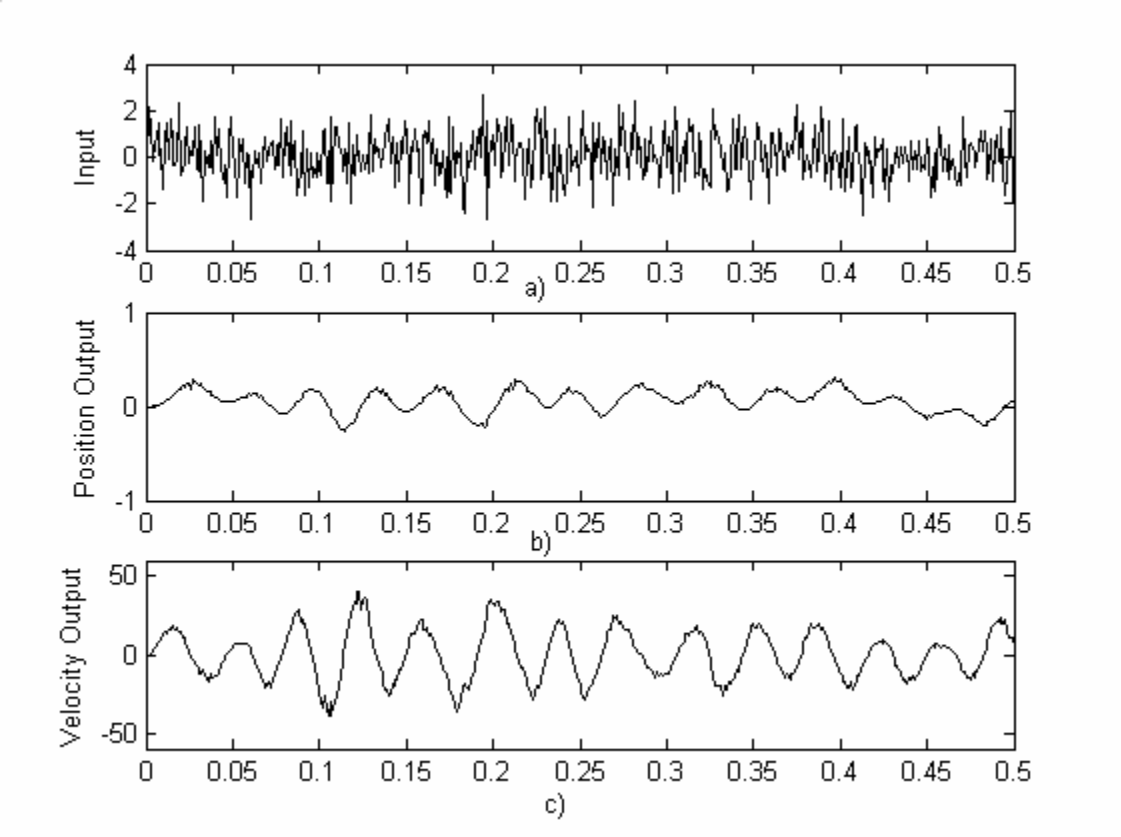


Figure 6.29 Input and Output Signals beyond the Assumed Noise Covariance

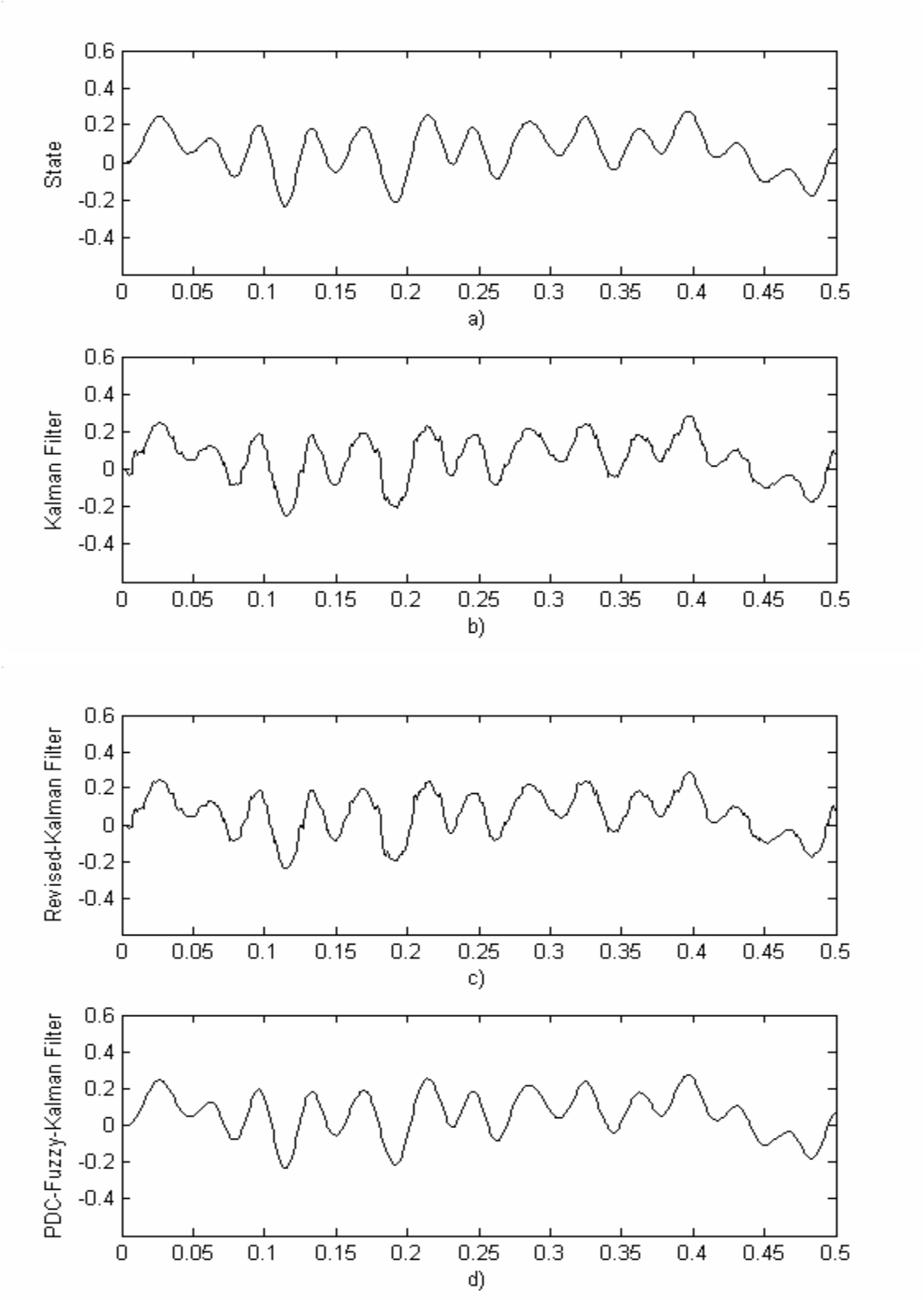


Figure 6.30 Position Estimate beyond the Assumed Noise Covariance (cm)

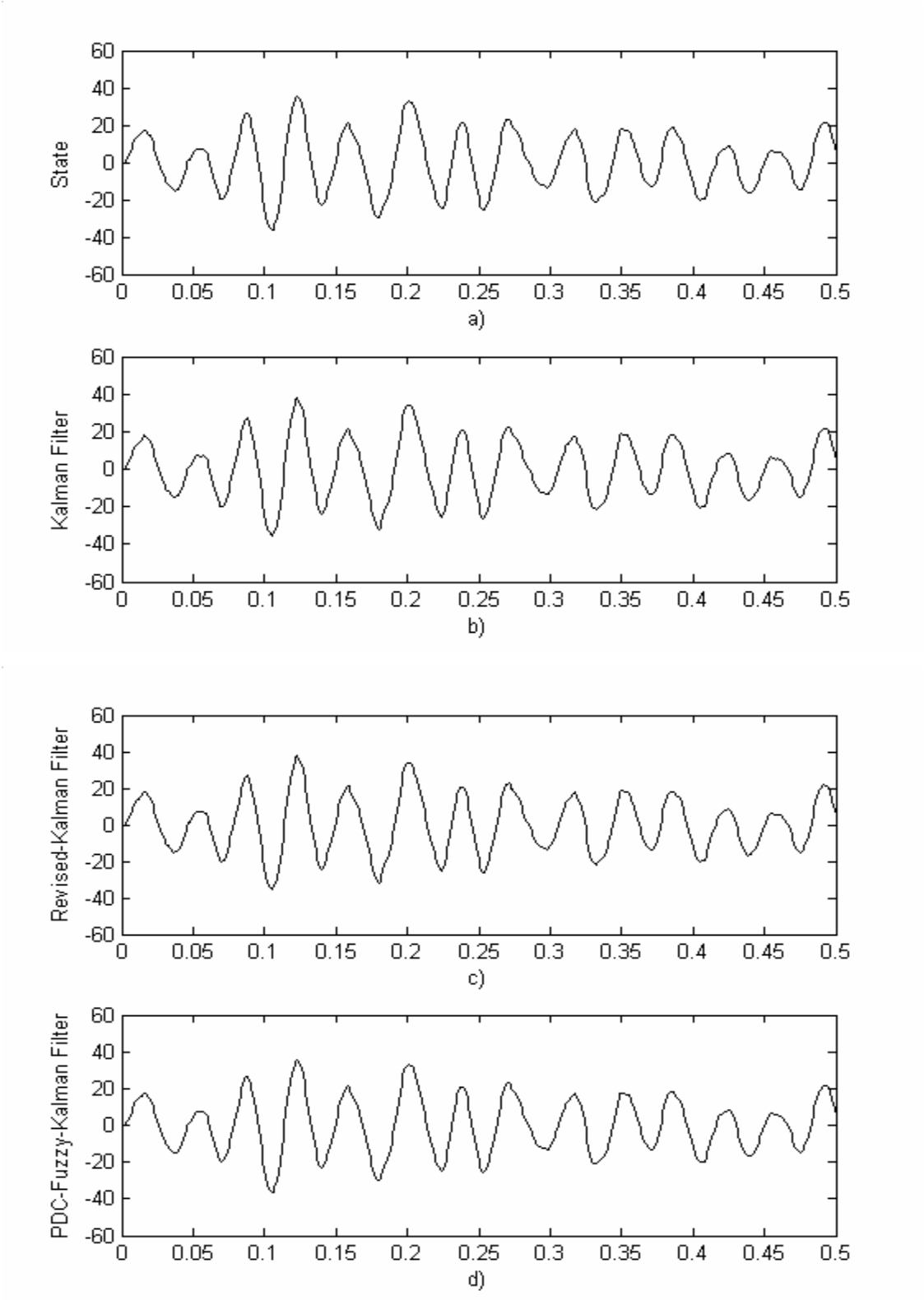


Figure 6.31 Velocity Estimate beyond the Assumed Noise Covariance (cm/s)

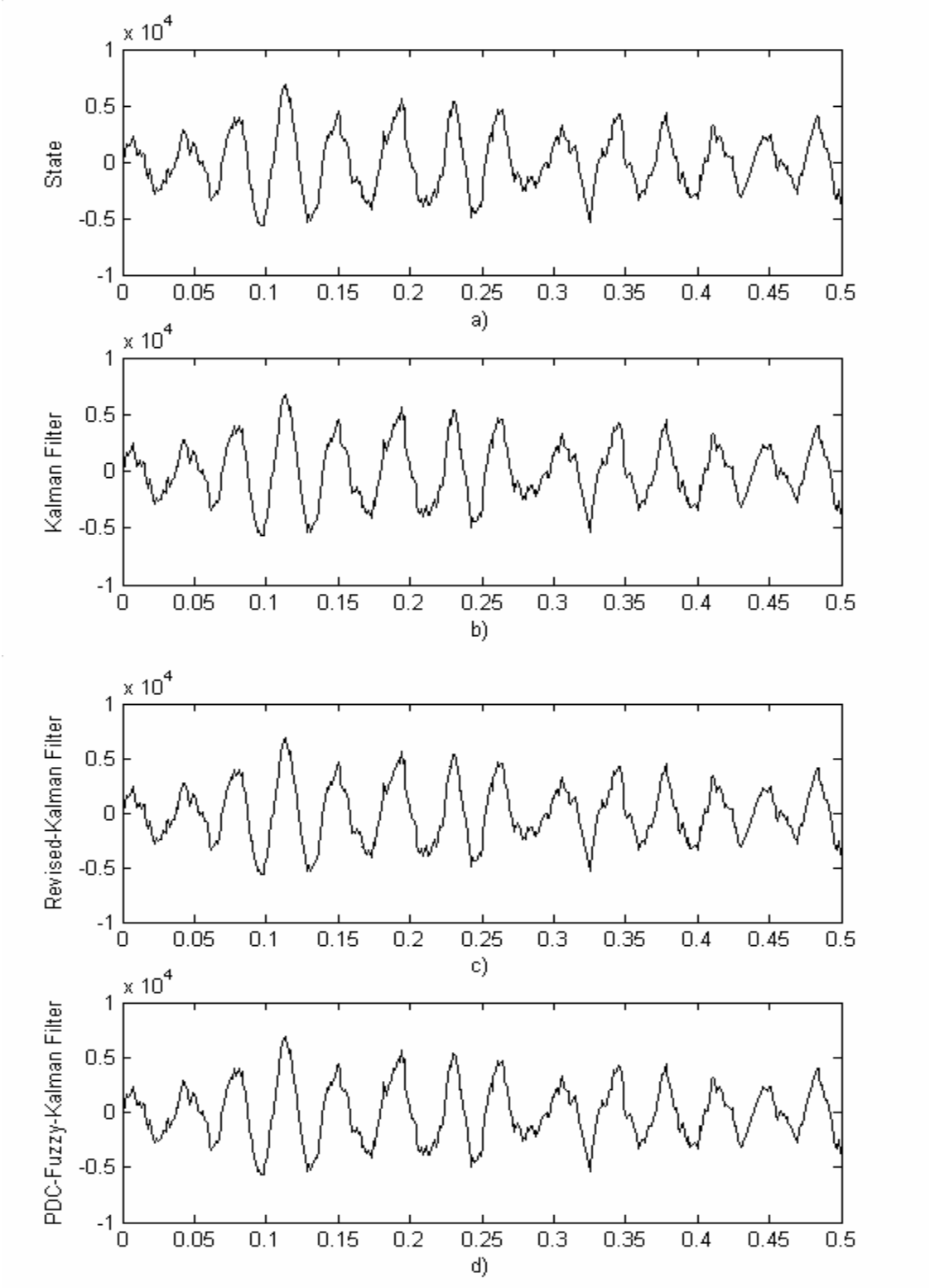


Figure 6.32 Acceleration Estimate beyond the Assumed Noise Covariance (cm/s^2)

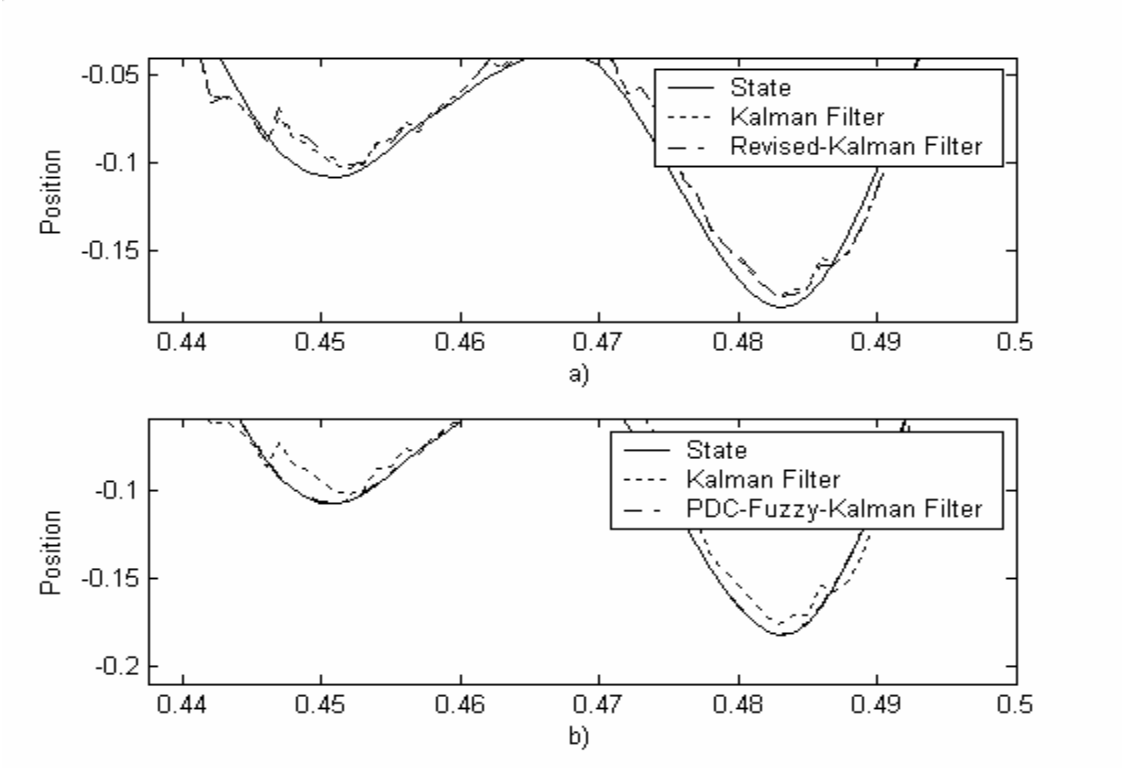


Figure 6.33 Expanded View for Comparing Position Estimations (cm)

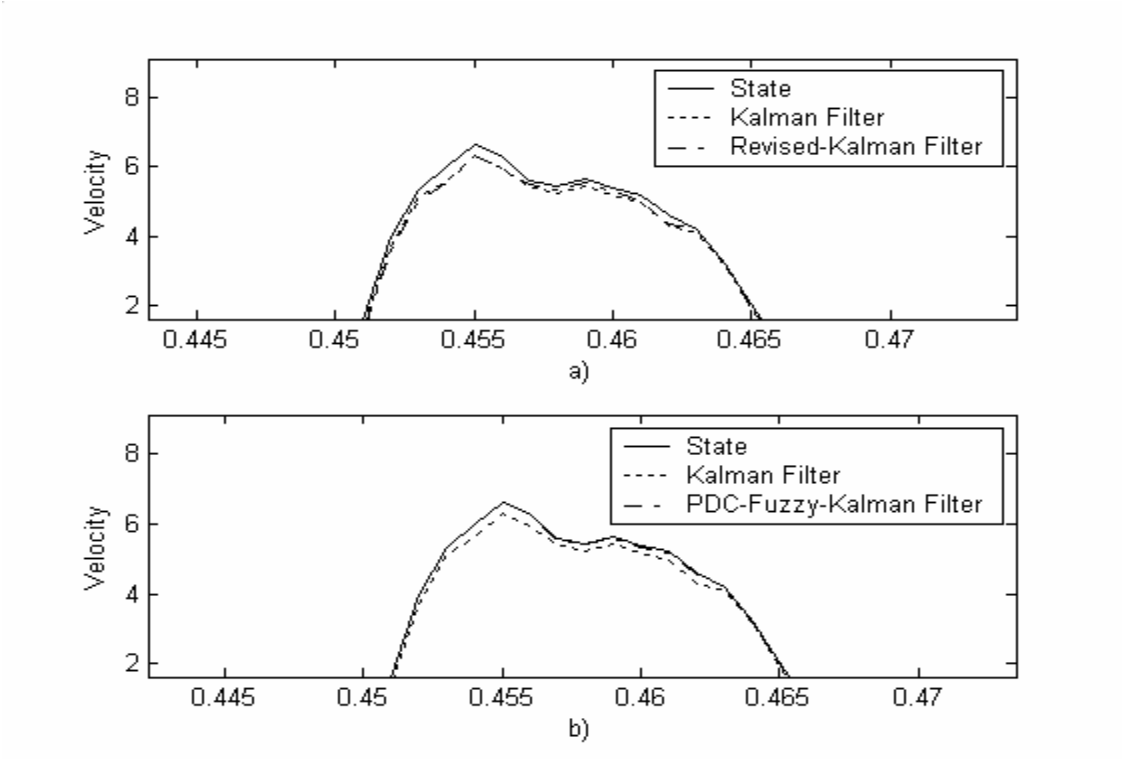


Figure 6.34 Expanded View for Comparing Velocity Estimations (cm/s)

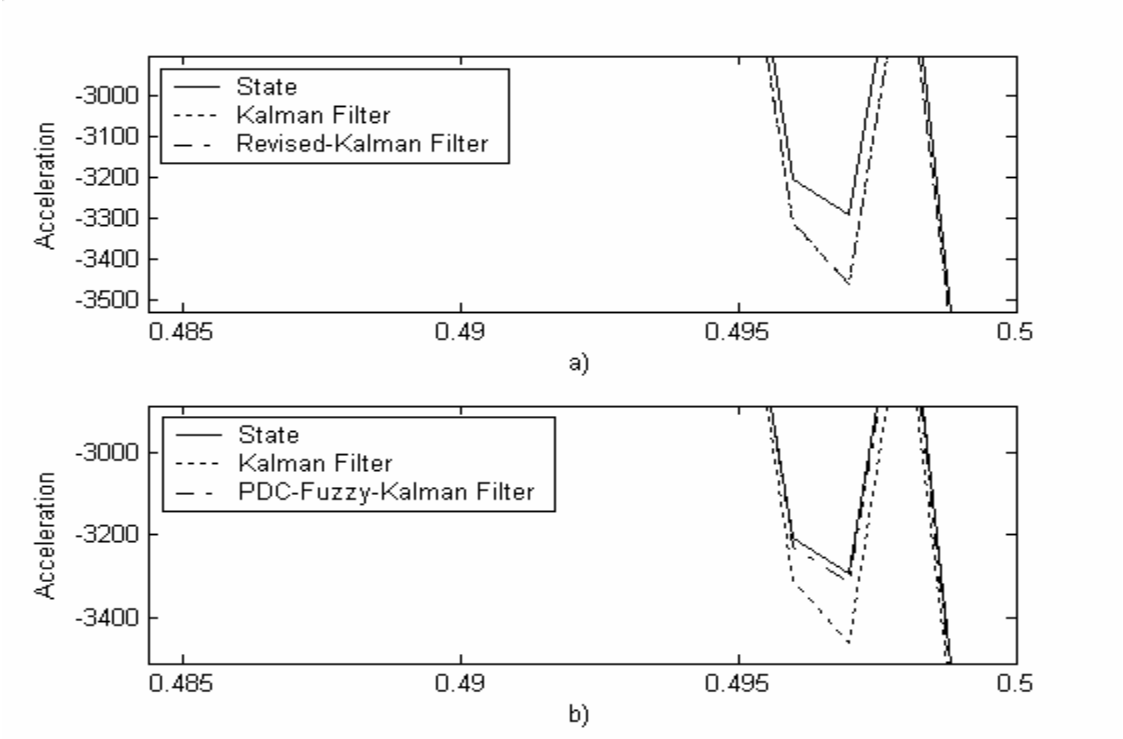


Figure 6.35 Expanded View for Comparing Acceleration Estimations (cm/s^2)

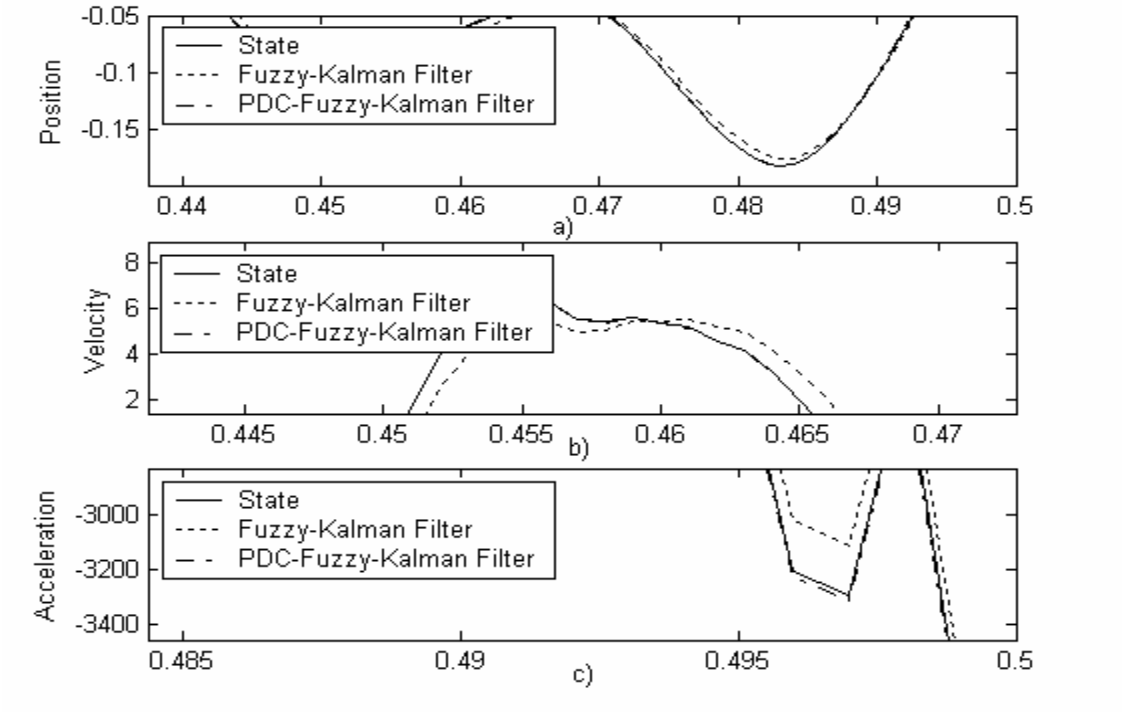


Figure 6.36 Expanded View for Comparing Two Fuzzy Kalman Filters

Chapter 7

Conclusions

The Kalman filter needs to be tuned according to the choice of an assumed system matrix, noise and the level of uncertainties. If conditions change to invalidate the tuning of the filter such as an increase in the amplitude of the measurement noise, then the amplitude of the estimates degrade. In this thesis the Fuzzy-Kalman filter is proposed to allow for the adaptation of the estimation process to changing system characteristics and noise. This filter constitutes a fusion of fuzzy logic and the Kalman Filter. For this fusion, the Kalman Filter concept has been revised into a form referred to as the Revised-Kalman Filter. The fundamental conceptual difference between the Kalman filter and its revised form is in the calculation of the corrective term. In the former the corrective term is a function of a gain and the a priori estimation error between the current computation step. Whereas, in the latter revised formulation, the a posteriori estimation error from the previous iteration step is used. In both cases, the associated gain multiplying the error term is an optimal derivation in the least square sense. In a simulation example, albeit limited, it is shown that the Revised-Kalman filter is as effective as the Kalman filter, but lends itself better to the application of fuzzy logic through the Sugeno-Takagi and parallel distributed compensator concepts. Further to a simulated example, it is found that the estimate from a tuned Kalman filter can be more accurate than the estimate from the Fuzzy-Kalman. However once outside the tuning envelope, the results of simulation show that the Fuzzy-Kalman outperforms the Kalman filter. This is because the Kalman

filter assumes that the system noise and the measurement noise are white and as such uncorrelated to the states and inputs. This assumption simplifies the derivation of the Kalman filter. In reality however, the noise may not be white and may not be of zero mean, [20], thus invalidating the optimality of the Kalman Filter. Fuzzy logic can be used for compensating for departures from ideal conditions. In the Fuzzy-Kalman formulation presented in this thesis, adjustments to the estimates are made through the adaptation of the system and measurement noise covariance matrices. Measurement error covariance matrix that is used as a membership function is used in this adaptation. Furthermore, provisions are made for the adaptation of the system matrix involved in the initial prediction of the states.

In conclusion, the methodology presented in this thesis successfully combines the Kalman filter concept and fuzzy logic. The preliminary simulation study considered here indicates that this combination is better at handling uncertainties.

Future Research:

This thesis has not provided a comprehensive adaptation of the Fuzzy-Kalman with multiple memberships. It is recommended as a continuation of this research, that the Fuzzy-Kalman Filter is applied to a real system thus forcing a realistic context for consideration of membership functions.

References

1. Kim, H.S., “Investment Game over the Stock Market” (in Korean), Kugil Stock and Economy Research Center, Kugil Media Publishers Co., Korea, 2002.

2. Kim, J.H., “A Future Trader with A Sense without University Degree” (in Korean), The Korea Daily, January 28, 1999.

3. Mukaidono, M., “Fuzzy Logic for Beginners” (in Korean), The Electronic Times Publishers Co., Ltd., Korea, 1995.

4. Koivo, H., “Lecture Note for Soft Computing”, Helsinki University of Technology, 2003.

5. Kosko, B., “Fuzzy Thinking: The New Science of Fuzzy Logic”, Hyperion, New York, 1993.

6. Passino, K.M. and Yurkovich, S., “Fuzzy Control”, Addison-Wesley, 1998.

7. Tsoukalas, L.H. and Uhrig, R.E., “Fuzzy and Neural Approaches in Engineering”, Wiley-Interscience, 1997.

8. Lin, C.T. and Lee, C.S., “Neural Fuzzy Systems, A Neuro-Fuzzy Synergism to Intelligent Systems”, Prentice Hall, 1995.

9. Dixit, A., “John Nash - Founder of Modern Game Theory”, Princeton University, May, 2004 - Internet website address: http://www.princeton.edu/~dixitak/home/nashenco.pdf

10. Kwon, O.H. and Yoon, T.H., “Game Theory and Strategy” (in Korean), Bumhan Book Corporation, Korea, 2000.

11. Park, C.B., and Prouty, R.W., “Helicopter Aerodynamics” (in Korean), Kyungmoon Publishers, Korea, 1992.

12. Soh, K.H., “Symbolic Logic” (in Korean), Kyungmoon Publishers, Korea, 1985.

13. Institute for Gifted Students, “Logic and Mathematical Induction (in Korean), Lecture Notes in Mathematics”, Korea Advanced Institute of Science and Technology, 1999.

14. Suber, P., “Recursive Function Theory”, Philosophy Department, Earlham College, May, 2004. - Internet website address: http://www.earlham.edu/~peters/courses/logsys/recursiv.htm

15. Suber, P., “Mathematical Induction”, Philosophy Department, Earlham College, May, 2004. - Internet website address: http://www.earlham.edu/~peters/courses/logsys/math-ind.htm

16. Aaby, A., “Theories and Systems: Mathematics, Science, Religion and Engineering”, Department of Computer Science, Walla Walla College, May, 2004. - Internet website address: http://cs.wwc.edu/~aabyan/Philosophy/Epistemology/theories.html

17. Aaby, A., “Correspondence Theory”, Department of Computer Science, Walla Walla College, May, 2004. - Internet website address: http://cs.wwc.edu/~aabyan/Philosophy/Epistemology/Correspondence.html

18. Kwon, S.H., “Fuzzy Sets, Measures and Integrals” (in Korean), Vol. 1, No. 3, ICASE, Korea, 1995, pp. 39 - 51.

19. Welch, G. and Bishop, G., “An Introduction to the Kalman Filter”, UNC-Chapel Hill, June, 2000 - Internet website address: http://www.cs.unc.edu/~welch/kalman/kalmanIntro.html

20. Brown, R.G. and Hwang, P.Y.C., “Introduction to Random Signals and Applied Kalman Filtering: with MATLAB exercises and solutions”, Wiley, 1997.

21. Layne, J.R. and Passino, K.M., “A Fuzzy Dynamic Model Based State Estimator”, Fuzzy Sets and Systems, Vol. 122, 2001, pp. 45 – 72.

22. Sasiadek, J. Z. and Wang, Q., “Fuzzy Adaptive Kalman Filtering for INS/GPS Data Fusion”, AIAA Guidance, Navigation, and Control Conference and Exhibit, August 1999.

23. Sasiadek, J. Z. and Wang, Q., “Sensor Fusion Based on Fuzzy Kalman Filtering for Autonomous Robot Vehicle”, Proceedings - IEEE International Conference on Robotics and Automation, v 4, 1999, p 2970-2975.

24. Habibi, S.R. and Singh, G., “Derivation of Design Requirements for Optimization of a High Performance Hydrostatic Actuation Systems”, International Journal of Fluid Power, Vol. 1, No. 2, October 2000.

25. Pytelkova, R. and Husek, P., “Synthesis of Takagi-Sugeno Fuzzy Systems, Advances in Fuzzy Systems and Evolutionary Computation”, World Scientific and Engineering Society Press, 2001.

26. Park, S.R., “ The Essence of Econometrics” (in Korean), Center for Corporate Studies, Korea Economic Research Institute, April, 2004 - Internet website address: http://www.parksr.net/

27. Watanabe, M., “Causal Inference with Non-experimental Data from National Institute on Drug Abuse Center” (in Japanese), Department of Economics, Toyo University, Japan, April, 2004. - Internet website address: http://stat.eco.toyo.ac.jp/~michiko/hakone/index.htm

28. Bai, D.S., “Inference, Lecture Note for Engineering Statistics 2” (in Korean), Korea Advanced Institute of Science and Technology, 2003.

29. Kim, S.J., “Basic Probability Theory, Lecture Note of Reliability Engineering” (in Korean), Pukyong National University, Korea, 2003.

30. Kobayashi, T., “Stationary Process, Lecture Note for Signal Processing” (in Japanese), Waseda University, Japan, 2003.

31. Park, M.Y. , Choi, H.S. and Sugeno, M., “Fuzzy Control System” (in Korean), Daeyoung Publishers Inc., Korea, 1990.

32. Ogata, K., “Modern Control Engineering”, 3rd Edition, Prentice Hall, 1997

33. Kim, D.H., “The Limits of Artificial Intelligence and Computer” (in Korean), Sungshin Women’s University, Korea, 2004

34. Kim, C.J., “Markov-Switching Models with Endogenous Explanatory Variables”, April, 2004. - Internet website address: http://www.econ.washington.edu/user/startz/Working_Papers/endogenousSwitching.pdf