A MARKOV CHAIN APPROACH TO ELECTROCARDIOGRAM MODELING AND ANALYSIS

by

Peter Charles Doerschuk

S.B., Massachusetts Institute of Technology
(1977)

S.M., E.E., Massachusetts Institute of Technology
(1979)

SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE
DEGREE OF

DOCTOR OF PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 1985

Signature of Author_____
Department of Electrical Engineering and Computer Science
February 8, 1985

Certified by_____
Alan S. Willsky
Thesis Supervisor

Certified by_____
Robert R. Tenney
Thesis Supervisor

Accepted by_____
Arthur C. Smith
Chairman, Departmental Committee on Graduate Students

A MARKOV CHAIN APPROACH TO ELECTROCARDIOGRAM MODELING AND ANALYSIS

by

Peter Charles Doerschuk

Submitted to the Department of Electrical Engineering and Computer Science on February 8, 1985 in partial fulfillment of the requirements for the Degree of Doctor of Philosophy

Abstract

A novel class of models of the electrocardiogram using interacting Markov chains is developed and used as a basis for signal processing. The modeling methodology emphasizes a balance between the inclusion of physiological detail and practicality for signal processing. In order for signal-processing algorithms based on the model to achieve accurate, detailed classification of the electrocardiogram, it is necessary to include physiological detail in the model. On the other hand, in order to make the signal processing practical, the models are restricted by imposing spatial, temporal, and hierarchical decompositions.

A signal-processing algorithm for a wave tracking problem relevant to rhythm classification is proposed. The algorithm is decomposed to mirror the spatial decomposition of the model. Limited simulations indicate that reasonable performance may be attainable.

Thesis Supervisor: Alan S. Willsky, Ph.D.
Title: Professor of Electrical Engineering

Thesis Supervisor: Robert R. Tenney, Ph.D.
Title: Lecturer in Electrical Engineering

Acknowledgements

I would like to thank my thesis supervisors, Professors Alan S. Willsky and Robert R. Tenney, for their unflagging interest over many years in this project. Professor Willsky first introduced me to the problem of electrocardiogram rhythm classification and encouraged me to pursue research in the area. Professor Tenney made crucial contributions to the modeling methodology (especially to the structure of the Markov chain) and to the signal processing techniques (especially the estimation goals, the subdivision of the estimators, and the modeling of communication). In addition, his help in defining the thesis scope and direction is greatly appreciated. The knowledgeable reader will recognize their mark on the thesis document itself; hopefully, their efforts will ease the reader's task.

I would also like to thank my thesis readers, Professors Sanjoy K. Mitter and Roger G. Mark. Professor Mitter's encouragement and interest in my education throughout my graduate school career and his faith that I could do something interesting have been invaluable to me. In addition, his support as Director of the Laboratory for Information and Decision Systems allowed me to finish this document in a timely manner. Professor Mark's generosity in providing the use of the Biomedical Engineering Center's DEC PDP 11/44 computer made it possible to pursue this thesis.

While the extensive software has been downplayed in the following chapters, several thousand hours of CPU time have been consumed in writing the code and generating the results. At many times during the course of the computer work I have benefited greatly from the advice and friendship of David A. Israel, Paul Albrecht, and George B. Moody. That the B.M.E.C. computer is a pleasant, productive computation facility is largely to their credit.

I would like to express my sincere appreciation for Arthur J. Giordani's care and patience with my myriad illustrations. His effort has much improved the presentation of this work.

I am sincerely grateful for the generous support, first of the Fannie and John Hertz Foundation and then of the M.D. - Ph.D. Program at Harvard University (funded in part by Public Health Service, National Research Service Award 2T 32 GM07753-06 from the National Institute of General Medical Science). I thank them for their continued confidence in me. In addition, Professor Willsky's time and a portion of the drafting expenses were defrayed by Air Force Office of Scientific Research Grant AFOSR-82-0258 while Professor Tenney's time was supported by the Department of Electrical Engineering and Computer Science at M.I.T.

Contents

1. Introduction

Many investigators have studied automatic rhythm interpretation of electrocardiograms (Thomas, 1979; Feldman, 1977; Oliver, 1977; LeBlanc, 1973; Cox, 1972; Proceedings of the IEEE Computers in Cardiology Conference, 1974-1984+).[1] However, at present there is no completely successful system capable of giving detailed rhythm classifications. We attempt to improve on previous work by applying statistical signal-processing ideas to models of the observed electrocardiogram signal where the models represent the rhythms in detail.

This thesis presents a model of the electrocardiogram (ECG) that is oriented toward detailed rhythm classification. Based on this model, a subproblem within the automatic rhythm-classification problem is formulated mathematically. Signal-processing algorithms solving this problem are developed. In simulations, these algorithms provide satisfactory performance. Evaluation on real ECGs has not been performed.

The ECG modeling methodology emphasizes the role of timing and causality since these characterize a rhythm. Furthermore, the methodology emphasizes that complex rhythms can be built up out of interactions between a few simple building blocks. By changing the building blocks, some of these modeling ideas may also be useful in other applications.

---

[1] In this thesis, references are cited by first author and date.

The emphasis on the problem rather than the possible methods of solution required the development of new signal-processing techniques. These techniques emphasize the structure in the model. As with the modeling methodology, these ideas may also be useful in other applications.

Signal processing based on a mathematical model of the observed signal, as outlined above, was called model-based signal processing in Willsky (1982). This approach differs somewhat from the more usual approach to ECG processing. More specifically a common element in many investigations is an attempt to mimic, at least in part, the human expert's approach to the problem. Due to this qualitative difference in approach, we do not review the bulk of the ECG signal-processing literature. Rather, the relevant literature is that concerning the modeling of the ECG, especially models oriented toward signal processing. This literature is reviewed in Section 3.1.

1.1. Description of the Problem

The ECG, an example of which is shown in Figure 1.1, is the surface recording of electrical-potential fluctuations due to the currents that flow in the heart muscle as it contracts. The term rhythm refers to the sequence and relative timing of the contraction of different parts of the heart. A disturbance of the normal rhythm is called an arrhythmia.

In a normally functioning heart, the ECG is made up of a sequence of discrete waveforms called P, Q, R, S, and T waves as illustrated in Figure 1.2. The waveforms differ greatly in signal to noise ratio (SNR) and, in more complex ECG's from arrhythmic hearts, multiple waveforms

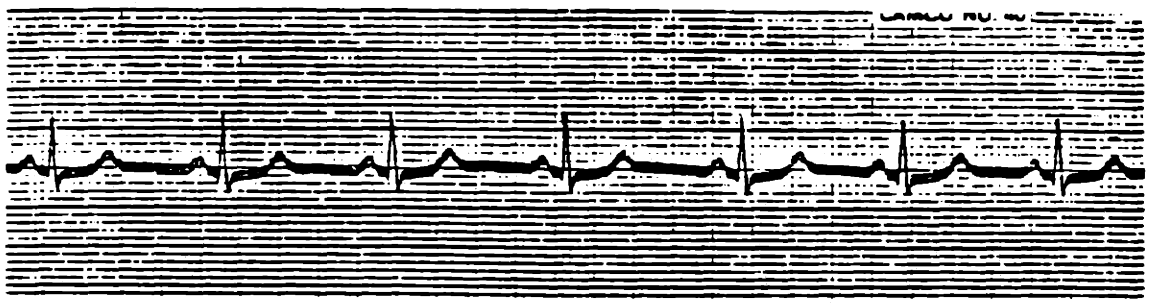


Figure 1.1 The Normal Electrocardiogram.
(From Schluter (1981), Figure 3a, p. 31).

can be superimposed. In each subject, different occurrences of the same wave are relatively similar though somewhat different from those of another subject. Even in an arrhythmic heart, the waveforms can often be classified as P, Q, R, S, and T though often several subtypes of a waveform must be introduced.

In a qualitative sense, the goal of rhythm interpretation is to monitor the behavior of the electrical conduction system of the heart. This conduction system controls the contraction of the heart muscle. Defects in the conduction system are associated with changes in the pattern in time of the observed waveforms and with the shape of each wave. Therefore the goal of rhythm-interpretation signal processing is to determine the correct pattern of timing and shape--the pattern that is actually occurring in the data--and to estimate several continuous parameters which characterize the pattern. An example of a continuous parameter would be heart rate, which primarily affects the pattern through a time normalization.

Arrhythmias are often divided into two categories--<u>persistent</u> and <u>transient</u> (Gustafson, 1978a, 1978b). A persistent rhythm is a rhythm where the pattern possesses some regulatity or repetitive property which

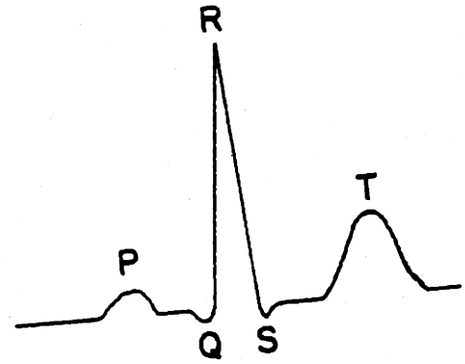


Figure 1.2 Waveform Definitions for One Beat of an Idealized, Normal ECG.

lasts for at least 6-10 heart beats. A transient rhythm is an abrupt change or irregularity lasting less than about 6 heart beats. Common occurrences in ECGs are ongoing persistent rhythms interrupted by transient events. An example of such a case is shown in Figure 1.3.

The shape of an ECG waveform is called its morphology. As indicated previously, morphology analysis is of use in rhythm interpretation. For example, morphology is important in distinguishing PVC beats[2] from normal beats. However, the subject of morphology analysis also addresses a large number of questions that are not relevant to rhythm interpretation. For example, the location of tissue damaged in a heart attack, the enlargement of one or more of the chambers of the heart, and diseases of the heart muscle itself are all reflected by changes in morphology and therefore are among the topics studied in morphology analysis, though their connection with rhythm

[2] See Chapter 2 for an introduction to cardiac behavior and the definition of terms such as this.

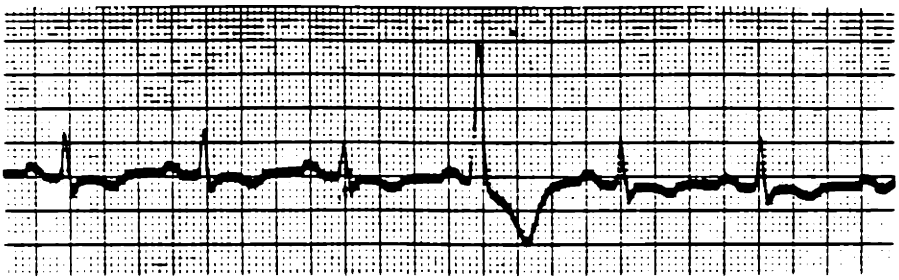


Figure 1.3 An Isolated PVC Beat Interrupting Normal Rhythm. (From Chou (1979), Figure 18-3B, p. 416).

analysis is rather distant. In our work on rhythm analysis, we address only those aspects of morphology analysis that are directly relevant to rhythm analysis.

1.2. Approach

Our approach to modeling emphasizes a tradeoff of physiological detail and practicality for signal processing. We have emphasized the event structure of the ECG, as these events completely characterize cardiac rhythms, and thus their characterization is the goal of signal processing. Furthermore, we have chosen to use models that match the spatial structure of the heart and that allow us to focus on two issues, namely timing of events in different parts of the heart and control, which corresponds to the electrical interactions of different portions of the heart (e.g. contraction of one part of the heart can cause a different part of the heart to contract).

Another reason for using models which mirror the spatial structure of the heart is our desire to obtain computationally feasible solutions to estimation problems. Our approach to signal processing is to mirror this spatial decomposition in the architecture of the estimator. More

specifically, the overall estimator is broken into subestimators corresponding to each of the submodels. Each subestimator is primarily concerned with the submodel to which it corresponds. An important issue is how to represent, for a given subestimator, the remainder of the model. The processing within each subestimator consists of a fixed sequence of passes through the data. A second important issue is the communication within subestimators and among the several subestimators between successive passes. If we used a single estimator based on the complete model (which is the natural approach if the model lacks structure) and considered models of the complexity described in this thesis, it would be more difficult to achieve computationally feasible solutions to estimation problems.

1.3. Thesis Outline

In the following chapter we give an overview of cardiac anatomy and physiology. This information provides the physical basis for our models. Then, in Chapter 3, we review and critique other approaches to ECG modeling and present our approach which we feel represents a contribution to the modeling of ECGs for signal-processing purposes. Moreover, the approach to modeling and the mathematical structure used may be applicable to other event-oriented systems, but that possibility is not explicitly explored.

Chapters 2 and 3 are optional for readers primarily interested in the signal processing since the mathematical formalism of the model is reviewed in the first section of Chapter 4. Succeeding sections of Chapter 4 describe our estimation goals and performance measures.

Chapter 5 presents our design methodology. Chapter 6 describes a series of case studies which are an attempt to justify our approach. Finally, Chapter 7 presents our conclusions and several directions for further research.

2. Cardiac Anatomy and Physiology

In this chapter, we present an outline of cardiac anatomy and physiology, concentrating on those aspects that are important for understanding cardiac rhythms. This information forms the physical foundation for our mathematical models of the electrocardiogram. As references, we have relied on Chou (1979), Katz (1977), and Marriott (1977).

2.1. Cardiac Anatomy and Mechanical Function

The heart is a four-chambered organ (right and left atria and right and left ventricles) which is divided into two pumping units. The right atrium and right ventricle form a unit which receives deoxygenated blood at low pressure (5 mmHg) from the systemic venous system and pumps it into the moderate-pressure (cyclicly 10 to 25 mmHg) pulmonary artery. From the pulmonary artery, the blood flows to the lungs and is oxygenated. The left atrium and left ventricle form a unit which receives oxygenated blood at low pressure (5 mmHg) from the pulmonary venous system and pumps it into the high-pressure (cyclicly 80 to 120 mmHg) aorta. From the aorta, the blood flows peripherally to the entire body. Because the left ventricle pumps into a higher pressure system than the right ventricle, the left ventricle has thicker, more muscular walls.

Both the right and left pumping units operate in the same cycle. In a normal heart, the cycle is:

(1) The atria are passively filled by the returning venous blood which, after the ventricle relaxes from the previous cycle's contraction, flows on through the atria, the tricuspid valve (right pumping unit) or the mitral valve (left pumping unit) and into the relaxed ventricles, partially filling the ventricles.

(2) The atria contract, emptying their contents into the ventricles through the tricuspid valve or the mitral valve. This completes the filling of the ventricles.

(3) The ventricles contract. Increased ventricular pressure closes the valves that allow communication with the atria (tricuspid and mitral valves) and opens the valves leading to the outflow arteries (pulmonic valve leading to the pulmonary artery in the right pumping unit, aortic valve leading to the aorta in the left pumping unit).

(4) While (3) is occurring, the atria relax.

(5) The ventricles relax. Decreased ventricular pressure allows the valves leading to the outflow arteries to close, thereby preventing backflow from the high-pressure outflow arteries into the low-pressure ventricles, and also allows the valves closing the connection with the atria to open, thereby starting ventricular filling.

In addition to following the same cycle, the right and left pumping units operate in synchrony.

2.2. Cardiac Control

The initiation and coordination of the cardiac function described above is performed by a complex electrical conduction system embedded in the muscular and structural elements of the heart. First we give a very phenomenological overview of the fundamental electrical process by which cardiac behavior is coordinated, that is, fluctuation of the electrical potential across cell membranes. Then we give an outline of the structure and function of the conduction system in a normal heart. In the following section, we look in more detail at the elements of the conduction system. Many of the molecular mechanisms of these processes are at least partially known but are not relevant to modeling for signal-processing purposes, and discussion of these is omitted.

2.2.1. The Fundamental Process

The elementary process by which the heart is coordinated is the fluctuation of the electrical potential across the cell membranes. These voltage fluctuations occur across the membranes of both the modified muscle cells of the conduction system and the muscle cells that perform the actual work of contraction. In the following description, the transmembrane potential is measured as interior potential minus exterior potential, which is the standard physiological convention.

The basic order of events for some patch of cell membrane is described in the following paragraphs. In its resting state, the cell membrane potential is approximately -90 mV. If an external depolarization wave arrives at this patch of cell membrane and the inward depolarizing current is sufficient to reduce the membrane potential to a

threshold voltage that is approximately -75 mV., then a series of events called an action potential ensues. A diagram of an action potential is shown in Figure 2.1.

At a given patch of cell membrane, the first event in a cardiac action potential is a rapid depolarization of this patch of cell membrane to approximately +20 mV. (region (a) in Figure 2.1). Thus the depolarization wave propagates. Furthermore, propagation is regenerative--irrespective of the magnitude of the initial depolarization current, so long as it was enough to move the transmembrane potential over the threshold, the depolarization achieves approximately +20 mV. Thus the propagating wave does not lose amplitude.

A complication of the regenerative propagation mechanism is significant to ECG rhythm interpretation. If the initial transmembrane voltage is less polarized, for example only -80 mV., then, if a suprathreshold depolarization wave arrives, the resulting depolarization of this patch of membrane is slower. Furthermore, the terminal value of the depolarization is unchanged so that the magnitude of the voltage change is reduced. This behavior has the effect of decreasing propagation velocity or even terminating propagation altogether. One mechanism which produces this partial depolarization is a bombardment of the patch of membrane with subthreshold stimuli.

The second event in a cardiac action potential is a brief, rapid repolarization to approximately 0 mV. (region (b) in Figure 2.1). The third is a long plateau at this membrane voltage (region (c) in Figure 2.1), and the fourth is a slower repolarization back to the resting voltage of approximately -90 mV. (region (d) in Figure 2.1). Thus the

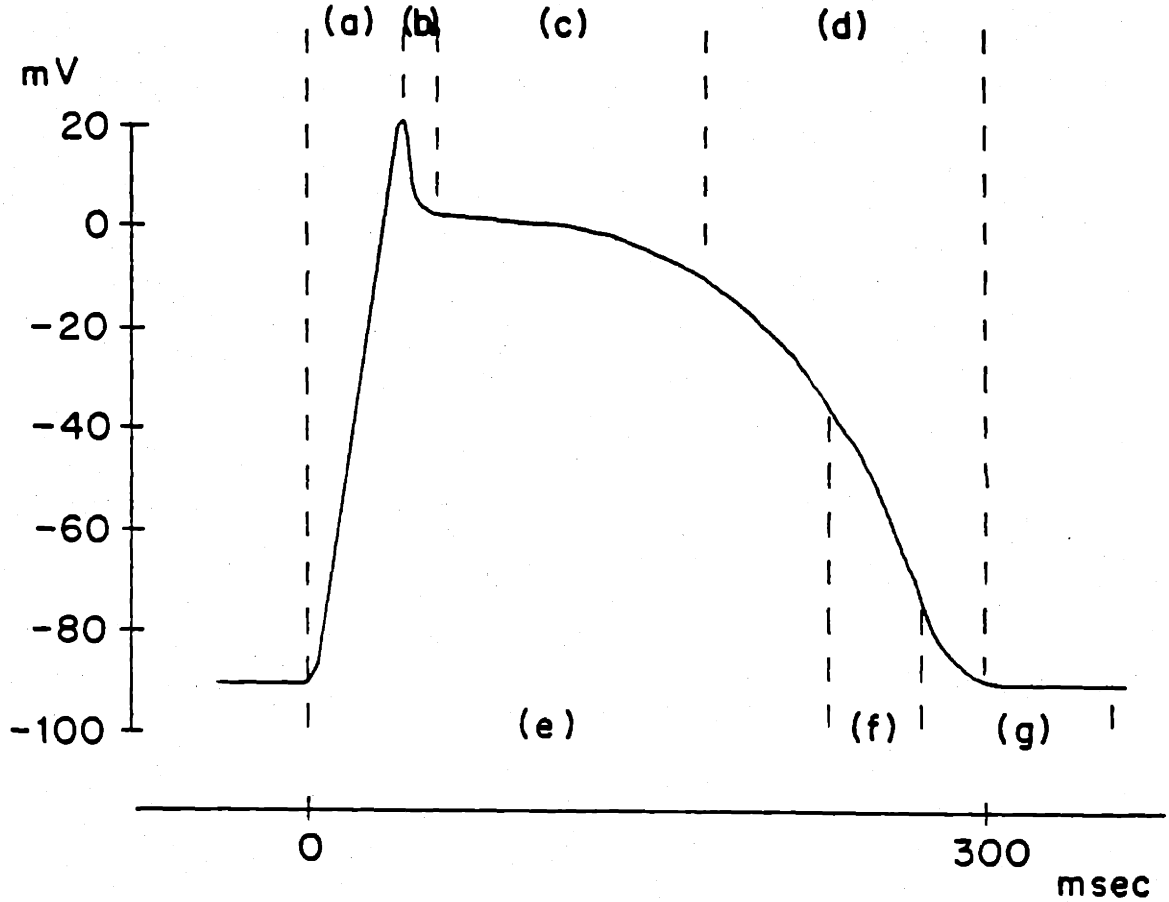


Figure 2.1 A Typical Cardiac Action Potential.

repolarization is automatic and does not require a second external trigger. The entire cardiac action potential takes about 300 msec.

A second, modified type of membrane also exists which, even when unexcited by external depolarization waves, does not maintain a -90 mV. membrane potential. Instead it slowly depolarizes until it reaches threshold and thereby triggers its own action potential. This depolarization is due to a slow ion leakage through the membrane. A cell with this type of membrane is called autorhythmic. Such cells form the pacemakers of the heart.

During the action potential, the patch of cell membrane undergoing the action potential is first absolutely and then relatively insensitive

to additional stimuli. This is called the refractory state and occurs for both types of cell membrane which were discussed above. From the initial stimulus until midway through the final repolarization (about 250 msec.) is the effective refractory period (region (e) in Figure 2.1). During this period a stimulus will not elicit a propagating action potential. However, late in this period, a stimulus can evoke a local response--a transient depolarization that is not propagated. This local response is important because it can effect the membrane's response to later stimuli by altering membrane properties temporarily. After the effective refractory period is the relative refractory period which lasts until nearly the end of the action potential (region (f) in Figure 2.1). During this period, only an abnormally strong stimulus will evoke a propagated response and the response will have both a slow and low-amplitude initial depolarization, and hence a low propagation velocity. The final period is the supranormal period which lasts a few tens of msec. past the end of the repolarization (region (g) in Figure 2.1). In this period, a subthreshold stimulus is able to elicit a propagated response which, however, has a low-amplitude initial depolarization. The net effect of these refractory periods is to disable, for typically 300 msec., the patch of membrane through which the action potential has propagated. This behavior is obviously important in understanding the cardiac control system.

2.2.2. Outline of the Conduction System

Now we outline the structure and function of the conduction system of the normal heart. The cells in the sinoatrial node (SA node), located in the right atria, are autorhythmic cells and undergo

spontaneous, cyclic action potentials which are propagated throughout the heart by the heart's conduction system and which control the muscular activity of the heart. On leaving the SA node, the propagating action potential causes the depolarization of the right and left atrial muscle (atrial myocardium). This causes the contraction of the atria (event (2) in Section 2.1). The propagation of this depolarization through the atrial muscle results in a potential fluctuation that can be observed in the ECG. In the ECG, this fluctuation is called the P wave (see Figure 1.2). After depolarizing the atria, the propagating action potential reaches the atrioventricular node (AV node) which, in a heart with normal anatomy, is the only electrical connection between the atria and the ventricles. In the AV node, the depolarization wave travels with a much reduced velocity, leading to a delay of 70-80 msec. in traversing the node. This delay allows the atria to complete the filling of the ventricles before the ventricles themselves begin to contract.

On leaving the AV node, the depolarization wave enters the specialized ventricular conduction system which will depolarize the right and left ventricular muscle (ventricular myocardium). The conduction pathway starts with the bundle of His which splits into the right and left bundle branches and then the left bundle branch splits into the anterior and posterior radiations. The bundle branches and radiations terminate in the many-branched Purkinje network which distributes the depolarization wave to the muscle cells of the inner side of the ventricular wall (endocardium). Finally, by muscle cell to muscle cell conduction, the remainder of the ventricular myocardium is depolarized. The depolarization of the ventricular myocardium causes the contraction of the

ventricles (event (3) of Section 2.1). Analogous to the atrial-myocardial depolarization generating the P wave of the ECG, the ventricular-myocardial depolarization generates the R wave (see Figure 1.2). Because the muscle mass of the ventricles is much greater than the muscle mass of the atria, the R-wave potential fluctuations have much greater amplitude than those of the P wave and hence the R wave has a much higher signal to noise ratio than the P wave. As indicated in Figure 1.2, the Q and S waves are respectively the initial and final fluctuations associated with the R wave and the entire waveform is referred to as the QRS complex.

Simultaneously with the ventricular contraction, the atrial myocardium repolarizes. This is event (4) of Section 2.1. While this repolarization also generates a surface potential fluctuation, it is smaller in amplitude than the P-wave fluctuation and is masked by the R-wave fluctuation generated by the ventricles. As discussed above, this repolarization is a local phenomenon, that is, it is not triggered by any global control mechanism but rather is the final event in the depolarization-repolarization cycle for any patch of cell membrane.

Finally, the ventricular myocardium repolarizes. This is event (5) of Section 2.1. This repolarization generates a surface potential fluctuation which, when measured in the ECG, is called a T wave (see Figure 1.2). This repolarization is also a local phenomenon, similar to the atrial repolarization.

2.3. Structure and Function of the Conduction System

With the outline of the previous section in mind, we now look in more detail at the conduction system.

An important concept is the gradient of autorhythmicity. Experimentally, it has been observed that many of the elements of the conduction system (in addition to the SA node) exhibit spontaneous, cyclic cell-membrane depolarizations that can then be propagated through the conduction system to depolarize the entire heart (i.e. they are autorhythmic). However, the rate of these depolarization cycles is different for different elements. Those elements that are more distal in the conduction system, that is, those elements that are later in the depolarization sequence outlined above,[1] have lower rates. This relationship between location and rate is the "gradient". For example, the SA-nodal cells cycle at 60-100 cycles per minute while Bundle of His cells cycle at 35-45 cycles per minute. Thus there exits a competition between the SA node and other, distal, autorhythmic centers for control of the heart.

The competition discussed in the prior paragraph is usually resolved in favor of the SA node by the following mechanism. When an autorhythmic cell is depolarized by the arrival of an external depolarization wave, the timer controlling when the next spontaneous depolarization will occur, which is physically a slow ion leakage across the cell

[1] The converse of distal is proximal. Thus a proximal structure is a structure that is depolarized early in the normal depolarization sequence and structure x is proximal to structure y if structure x is depolarized, in the normal depolarization sequence, before structure y.

membrane (see Subsection 2.2.1), is reset. Thus, for example, if a Purkinje cell is depolarizing at a rate of 20 cycles per minute and is 2.5 seconds into a cycle when an external wave of depolarization arrives, the next spontaneous depolarization of the cell will not occur in 0.5 seconds but rather in 3.0 seconds.

When a node is reset, a phenomenon called stunning can also occur where the time until the next spontaneous depolarization is increased over the nominal time. The stunning phenomenon can be seen in the ECG in some situations where the SA node itself is reset. In the normal heart, because of the reset phenomenon, the SA node is able to retain control of the heart in spite of the competition from the other autorhythmic centers by continually resetting them before they have an opportunity to spontaneously depolarize.

While we have described unidirectional propagation of the depolarization wave through the conduction system in what is called the antegrade direction, the system is also capable of conduction in the reverse direction, called the retrograde direction. This is very important in cases when the cell membranes of the conduction system are not in a refractory state and a depolarization wave is initiated in a distal structure.

We now consider the elements of the conduction system in more detail.

(1) Sinoatrial node

This node is located in the wall of the right atrium near the entrance of the superior vena cava (one of the main conduits for

systemic venous return). It is autorhythmic with a rate that is typically in the range of 60-100 cycles per minute. The actual rate is under the control of the sympathetic and parasympathetic nervous innervation of the node and the blood-borne hormone levels in the node. This node, as discussed above, is the primary pacemaker in a normal heart.

(2) Internodal tracts

These are functionally defined, in contrast to anatomically defined, preferred conduction pathways between the SA and AV nodes and between the two atria. There are three tracts: anterior, middle, and posterior. The middle and posterior tracts serve only as preferred conduction pathways between the SA and AV nodes while the anterior tract branches so that it can also served as a preferred electrical pathway between the two atria. This connection serves to synchronize the contraction of the two atria. If it is broken, the left atrial contraction lags the right atrial contraction.

(3) Atrial Myocardium

The atrial myocardium, in addition to its mechanical role, acts as a part of the conduction system. The depolarization wave is spread directly from the SA node and also from the internodal tracts by cell to cell conduction in the atrial myocardium with a velocity of 1.0 m/sec.

(4) Atrioventricular Node (AV node) and the Bundle of His

These structures are collectively called the AV junction. In the anatomically normal heart, the bundle of His is the only electrical path from the atria to the ventricles. Depolarization waves enter

the bundle of His via the AV node which is a small mass of specialized myocardium located above the coronary sinus on the posterior wall of the right atrium. There are three anatomically defined subdivisions of the AV node. From most proximal to most distal they are atrionodal, nodal, and nodal-His. Autorhythmic activity has been demonstrated in the atrionodal and nodal-His regions of the AV node (normal rate of 45-60 cycles per minute) and in the bundle of His (normal rate of 35-45 cycles per minute). It has not been demonstrated in the nodal region of the AV node (Chou, 1979, pp. 393, 397). As discussed above, the propagation velocity for a depolarization wave in the AV node is slow (0.05-0.1 m/sec) leading to a delay before the action potential reaches the bundle of His, where the propagation velocity is 1.0-2.0 m/sec, and, eventually, the ventricles.

(5) Anatomic Variants in Atrioventricular Conduction

There are at least three anatomic variants that affect the electrical connection between the atria and the ventricles.

(a) Bundle of Kent

The bundle of Kent is a direct connection between the atria and ventricles that completely bypasses the AV node and the bundle of His. Thus in a subject with this anomaly, both the normal AV-nodal delay is avoided and the various regions of the ventricular myocardium are excited in an abnormal sequence. Because of the abnormal excitation sequence, the QRS-complex morphology is abnormal.

(b) Bypass Fibers of James

The bypass fibers of James are a direct connection between the atria and the distal third of the AV node or the bundle of His. Thus in a subject with this anomaly, the normal AV-nodal delay is decreased or absent. However, the ventricular myocardium is activated in the normal sequence via the bundle of His. Therefore, the QRS-complex morphology is normal.

(c) Bypass Fibers of Mahaim

The bypass fibers of Mahaim are a direct connection between the lower AV node or the bundle of His and the ventricular septum (the muscular wall dividing the left and right ventricles). Thus, in a subject with this anomaly, the normal AV-nodal delay is present but the ventricular myocardium is excited in an abnormal sequence. Therefore, the QRS-complex morphology is normal.

Because these three anomalies have a common embryologic basis, mixtures are common. The frequency of this class of anomalies has not been accurately estimated anatomically. However, the incidence of the Wolff-Parkinson-White syndrome, which is explained in terms of these anomalies, is estimated at 0.15 to 0.2 percent of the general population (Chou, 1979, p. 489). In order to illustrate the generality of our modeling approach, we propose a model for this syndrome in Subsection 3.7.4.

(6) Left and Right Bundle Branches, Anterior and Posterior Radiations

The Bundle of His divides into the left and right bundle branches, and then the left bundle branch divides into the anterior and posterior radiations. The left bundle branch and its radiations activate the Purkinje network of the left ventricle while the right bundle branch activates the Purkinje network of the right ventricle. The depolarization-wave propagation velocity in these structures is 1.0-2.0 m/sec. They contain autorhythmic cells whose normal rate is similar to the normal rate of cells in the Bundle of His.

(7) Purkinje Network

The Purkinje network branches from the left and right bundle branches and the anterior and posterior radiations. It activates the interior surface of the ventricular myocardium (endocardium). The depolarization-wave propagation velocity in the Purkinje network is 4.0 m/sec. It contains autorhythmic cells whose normal rate is the slowest such rate of any class of cells in the heart.

(8) Ventricular Myocardium

The ventricular myocardium, in addition to its mechanical role, acts as a part of the conduction system. The depolarization wave started by the Purkinje network in the endocardium propagates from cell to cell in the ventricular myocardium with a velocity of 0.5 m/sec until all of the ventricular myocardium has been depolarized.

2.4. Origin of the ECG

The flow of charged ions across the cell membrane, which is the physical mechanism for all depolarizations, is a time-varying impressed current density in a conducting medium. In the quasi-static limit this current density gives rise to a (time-varying) potential field. Potential differences between fixed, chest-wall sampling locations then give the ECG. The electromagnetic origin of the ECG is discussed quantitatively and in some detail in Plonsey (1969). The cell-membrane area of the myocardium is much greater than the cell-membrane area of the remainder of the conduction system. Furthermore, the electrodes in the standard ECG recording arrangement are relatively distant from the entire heart. Myocardial depolarization thus dominates changes in the distribution of the impressed current density and hence in the voltage fluctuations seen in the standard ECG recording arrangement.

Clearly, changes in the time course of the impressed current density will result in changes in the ECG fluctuations and such changes can occur because the depolarization propagates through the myocardium in unusual directions. As discussed in Section 1.1, the analysis of the shape of the ECG voltage fluctuations is called morphology analysis. While morphology and rhythm analysis are interrelated, in this work we are separating them to the greatest extent possible. (We only use morphology information when it is important for rhythm diagnosis. For example, the correct classification of a slightly premature R wave as belonging to a normally-conducted beat versus a premature ventricular contraction requires morphology information).

2.5. Mechanisms for Abnormal Cardiac Control

The evolution of arrhythmias is strongly influenced by three general mechanisms: retrograde conduction, reset/stunning, and the gradient of autorhythmicity, all of which are physiologically normal and have been discussed above. These properties interact with two broad categories of physiologically abnormal properties--decreased conduction and increased or decreased autorhythmicity--to generate arrhythmias.

Decreased conduction can occur in several different forms which can also be mixed. The following forms are all invoked to explain various arrhythmias:

(1) Total block of all depolarization waves.

(2) Unidirectional block of all depolarization waves coming from a particular direction. This is explained in terms of an asymmetric severity of tissue damage (see Katz, 1977, pp. 310-314).

(3) Decreased propagation velocity.

(4) Increased refractory time.

Increased (decreased) autorhythmicity refers to an increased (decreased) rate of autorhythmic depolarizations.

Typical causes for these abnormalities are drugs and decreased blood flow to the affected portion of the cardiac tissue. The abnormalities can be transient, lasting from minutes to days, or permanent. See the references for details.

Explanations of arrhythmias in terms of the previously discussed causes (e.g. the increased refractory time of a certain volume of myocardium) are at a low level of abstraction. That is, the discussion focused on specific cellular causes of these phenomena. At a higher level of abstraction (i.e. mechanisms that could have several different cellular origins), a common explanation for many cardiac arrhythmias is the idea of reentrant depolarization waves. There are many detailed, cellular mechanisms that have been proposed. Three typical mechanisms are the following.

(1) Inhomogeneity of Repolarization

An area of myocardium with delayed repolarization adjoins an area with normal repolarization. An initial wave of depolarization propagates through both areas. Because of the asymmetry in repolarization times, the normal myocardium recovers its excitability before the delayed-repolarization myocardium has repolarized. This allows the delayed-repolarization myocardium to reexcite the normal myocardium. Thus the depolarization wave reenters the normal myocardium and eventually this local depolarization can evolve into a second general depolarization.

(2) Unidirectional Block with Conduction Delay

An area of unidirectional block with conduction delay can lead to a reentrant depolarization wave if the initial wave is incident in the blocked direction. In Figure 2.2, the depolarization wave is incident at p. It is blocked in branch 1 by the unidirectional block between A and B. However, it is free to propagate through branch 2 and then around to reach branch 1 from the opposite

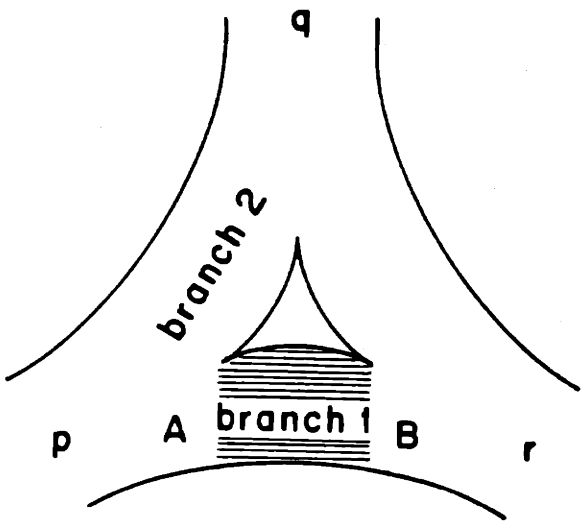


Figure 2.2 A Mechanism for Reentrant Depolarization.

direction. (Along the way, the depolarization wave excites the q and r branches). In this direction, the depolarization wave is not blocked between A and B. However, its arrival at A has been delayed by the circuitous path taken. Thus when it reaches A, the normal myocardium on which it is now incident is repolarized. Hence the depolarization wave is able to reenter this myocardium and eventually reexcite at least the p branch. Depending on the exact timing, this process can continue for multiple loops, or it may terminate. While this geometric pattern of muscle fibers seems complex, it is a reasonable idealization for some patterns actually seen in the myocardium.

(3) Slow Repolarization

In the same Figure 2.2 used in part (2), replace the unidirectional-block area with an area of slow repolarization. Now, if this area has not had time to repolarize after the prior depolarization wave when the current wave arrives, the same sequence of events described in part (2) can occur. Now however,

one needs to assume that the slow repolarization region has regained excitability by the time that the detoured depolarization wave reaches it for the second time (i.e. after the wave has passed through branch 2).

Note that a microscopic reentrant system has many of the properties of an autorhythmic cell. Thus there is an overlap between the classes of events that can be explained in terms of microscopic reentry and in terms of increased autorhythmicity. In explanations of some other events, the reentrant system is macroscopic. An example of a macroscopic reentrant pattern would be one including both the atria and the ventricles through the bundle of Kent.

2.6. Summary of Rhythm Causation and Implications for Modeling

Based on a study of the references cited previously, we have concluded that cardiac rhythms can usually be explained in terms of just a few mechanisms. These mechanisms are:

(1) directional conduction,

(2) conduction delay,

(3) reset/stunning,

(4) refractory period,

(5) autorhythmicity, and

(6) unusual anatomic structure.

Therefore, these are the properties that our mathematical formalism must be able to capture.

Furthermore, we have observed the importance of interactions between the abnormal and normal parts of the heart. That is, many of the changes in an arrhythmic ECG are due to how the abnormal substructure affects the normal substructures rather than to a direct change in the ECG caused by the depolarization of the abnormal substructure. It is possible to pursue an approach in which one proposes completely new, ad hoc models for arrhythmias. However, we choose to model arrhythmias by introducing anomalies into submodels in a manner that exactly parallels the cause of the arrhythmia in the actual heart. Therefore, our model must capture interactions so that we may develop models for arrhythmias by making changes in the model of normal rhythm, where the changes should be conceptually limited to those parts of the model corresponding to the abnormal substructures.

2.7. Lewis or Ladder Diagrams

A Lewis diagram, also called a ladder diagram, is a simple way, in common clinical use, to diagram the activities of the cardiac conduction system. We will have occasions to use these diagrams in subsequent chapters. The vertical axis is the location of the depolarization wavefront, and the horizontal axis is time. Four lines indicating three subdivisions are used to indicate location on the vertical axis. The uppermost line represents the SA node, and the three spaces represent, from top to bottom, the atria, the AV junction, and the ventricles.

Thus a ladder diagram for several beats conducted in the normal manner appears as in Figure 2.3: the depolarization front originates in the SA node, rapidly traverses the atria (nearly vertical line) slowly traverses the AV junction (angled line) and finally, rapidly traverses the ventricles (nearly vertical line).

2.8. Terminology

The following definitions are common terminology used in describing ECGs.

(1) Coupling Interval: The length of time between a particular depolarization of some structure and the most recent previous depolarization.

(2) x-y Interval: the x-y interval is the length of time between the x event or wave and the subsequent y event or wave.

(3) Escape Beat or Rhythm: A beat or rhythm initiated by a distal pacemaker because either all proximal pacemakers failed to depolarize or because the depolarization was not transmitted to the distal structure.

(4) Ectopic Pacemaker: Any pacemaker, except the SA node, which is successful in initiating the depolarization of some structure.

(5) Fusion Beat: A depolarization of a specified structure in which the depolarization originates from more than one source.

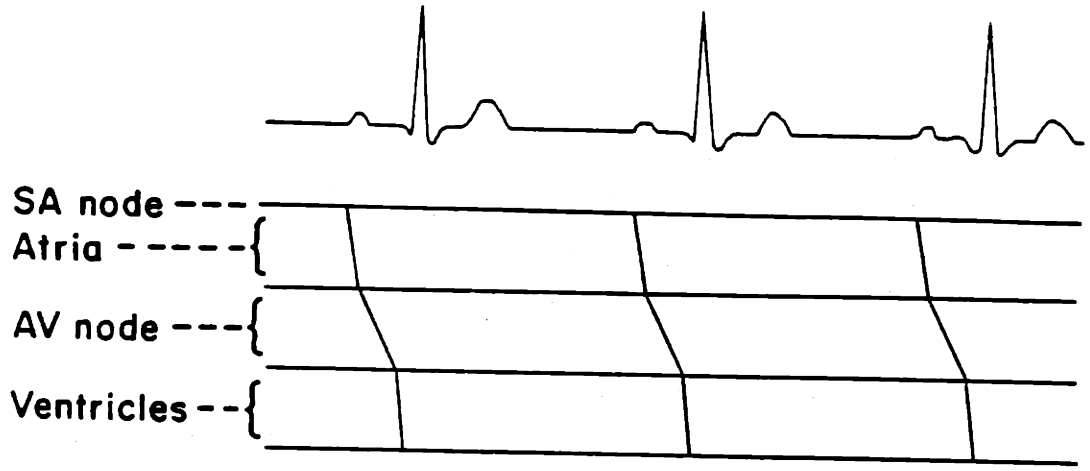


Figure 2.3 A Ladder Diagram for a Sequence of Normal Beats with the Corresponding Idealized ECG.

(6) Aberrant Conduction: An aberrantly-conducted depolarization is a depolarization that does not propagate through the standard conduction pathways of some structure.

(7) Supra-ventricular: An adjective referring to all structures proximal to the ventricles.

(8) PVC: An abbreviation for Premature Ventricular Contraction, which is a contraction of the ventricles for which the source of the depolarization is within the ventricles. Since it precedes the next expected (i.e. due to the underlying persistent rhythm) ventricular contraction, it is called premature.

3. Electrocardiogram Models

In this chapter we describe a new modeling methodology for ECGs. After a review of past literature on ECG modeling, we describe our approach to modeling and give five examples. Each example considers a different rhythm. The fifth example, which is considered in significantly more detail than the prior four, includes simulations to demonstrate the descriptive power of our models. The ideas of Chapter 2, especially in the summary in Section 2.6, are basic to this chapter.

3.1. Literature Review

This review of the ECG modeling literature has two purposes.

(1) We attempt to motivate the structure of our model by describing physiologically oriented models. Simultaneously we try to point out why these other models are not suitable for signal processing.

(2) We illustrate what we believe are limitations of the signal-processing-oriented models described in the past. We believe that some of these shortcomings are rectified in our models.

An underlying theme throughout this review is that models created for one specific purpose are often not appropriate for some other purpose.

3.1.1. Physiological Models

Electromagnetic models of the heart were originally developed in order to understand what the ECG and later the magnetocardiogram (MCG) were actually measuring. Text and review article references include (Geselowitz, 1979), (McFee, 1972), (Plonsey, 1966, 1969, 1971, 1979), (Tripp, 1979), and (Wikswo, 1979). As such, these models are oriented toward deriving spatial transfer functions from quasi-static electromagnetic sources (impressed currents across cell membranes) to measured variables (surface potentials for ECG and surface magnetic fields for MCG). Extensively studied questions include source models (Geselowitz, 1967; Plonsey, 1969, 1974; Spach, 1979), the effects of chest geometry and inhomogeneities (Barr, 1977; Cuffin, 1977; Hosaka, 1976; Rudy, 1979; Rush, 1971), measurement number and location (Barr, 1971; Lux, 1978), and inverse methods (i.e. calculating impressed sources from observations instead of visa versa) (Baker, 1974; Baldwin, 1979; Barr, 1969, 1970; Brody, 1972; Cuffin, 1978; Martin, 1972, 1975; Miller, 1974; Schloss, 1971).

Excepting some inverse calculations, these models do not prescribe how the source evolves over time--any impressed current source distribution can be used. Thus, while they are important for relating the depolarization of some volume of tissue to its ECG manifestations, they themselves do not provide rhythm models.

The remaining models described in this review are models for ECG behavior over a period of time. That is, in terms of the prior paragraph, the evolution of the source over time is prescribed for at least the duration of a single beat.

Many investigators have described finite-element models for different aspects of cardiac behavior. Miller and Geselowitz (Miller, 1978a, 1978b) described a single-beat model for body-surface potential maps. It combines a realistic finite-element model of the electromagnetic properties of the torso with experimentally determined depolarization-time and action-potential-duration data.

Vinke and his collaborators (Vinke, 1977) described a deterministic finite-element model of ventricular timing for a single normal beat. Each element has three states--resting, active, and refractory. The different depolarization-wave velocities are modeled in moderate detail.

Moe and Mendez (Moe, 1966) developed a stochastic two-dimensional model of atrial depolarization in atrial fibrillation. The form of the model is similar to that of Vinke (Vinke, 1977) except that the refractory state, which is no longer a trapping state, is modeled in much greater detail and there is no conduction system. The stochastic aspect is modeled by a parameter in the refractory period (for each element) which is constant throughout the simulation.

Cohn and his associates (Cohn, 1982) developed a model combining a finite-element timing model (similar to Vinke, 1977) and a finite-element electromagnetic model (similar to Miller, 1978a, 1978b). This is a deterministic model of a single beat and was used to study T waves.

Smith (Smith, 1982) developed a model combining finite-element timing and finite-element electromagnetic models. Unlike Cohn (Cohn, 1982), this is a model for a complete rhythm. Smith's interest is in the mechanisms of arrhythmias. Again, each element can be in one of

three states, and is modeled using a random, constant parameter for the refractory period. The electromagnetic model contains a very simple torso model since Smith was concerned with rhythm and not morphology.

The previously described models have hundreds to thousands of elements. Models have also been proposed with only tens of elements, where each element models a specific substructure in the cardiac anatomy. Rosenberg and his associates (Rosenberg, 1972) proposed a model of normal rhythm constructed from 13 dipoles. The dipoles are of fixed location and orientation but variable moment which, along with the inter-dipole timing, is fixed by a coupled collection of second-order relaxation oscillators. Each oscillator and its dipole model one anatomic substructure of the heart. The uncoupled oscillation frequencies of the relaxation oscillators follow the gradient of autorhythmicity. Thus, when they are coupled, the SA-nodal oscillator entrains all of the oscillators modeling more distal structures. By adjusting the properties of the relaxation oscillators, Rosenberg is able to adjust the time course of the dipole moments such that they generate quite realistic ECG beats.

In Zloof (Zloof, 1973) this model was further developed. By introducing more complex relaxation oscillators for the AV junction, the overall model is able to simulate various types of first-, second-, and third-degree AV block. The type of AV block generated depends on the value of a continuous parameter.

The Rosenberg-Zloof model was further modified in Thiry (Thiry, 1974, 1975). In this work, the timing mechanism is separated from the mechanism determining the dipole magnitudes. The dipole magnitudes were

then determined based on idealized geometry and experimental data.

These models, which were developed for other purposes, have major drawbacks if one were to consider using them as the basis for signal processing. In order to place our approach to modeling in perspective, we now list several of the limitations (as far as signal processing is concerned) that are common to one or more of the models are described previously.

(1) Models which assume deterministic behavior preclude the consideration of ensembles of waveforms that are close in a probabilistic sense but differ in detail.

(2) The ECG morphology information is deeply embedded in the model and is tightly coupled to the timing information. While this coupling is more realistic than a hierarchical arrangement which decouples the timing from the ECG-morphology properties, it makes the signal processing more complex. Specifically, adapting the morphology generated by the model to the actual morphology observed in the subject is a very complex undertaking, which may not have a unique solution and which will also alter the timing. Furthermore, the absence of this hierarchical structure obviously precludes our taking advantage of it to devise efficient, hierarchical algorithms.

(3) The model is (probably) not completely observable from the surface-electrical-potential output. That is, knowing the output and the structure and parameters of the model is not sufficient information to uniquely reconstruct the value of all internal variables. Furthermore, even if the model were completely observable,

estimation based on the model would involve estimation of far too many quantities, most of which are of no direct importance for rhythm diagnosis.

Note that several features of these models are important for our purposes in and approach to modeling:

(1) The models capable of generating rhythms are often stochastic. The way in which randomness is introduced in the models described to this point is typically limited (e.g. by a stochastic initial choice of parameters; the vast number of interacting variables in these models then evolve in a deterministic but very chaotic appearing way). In our approach we introduce randomness in a more fundamental way.

(2) The models often have different levels of detail for timing effects than for electromagnetic effects. Furthermore, these two aspects are sometimes treated hierarchically: the timing model drives the electromagnetic model which generates waveforms through a mechanism completely independent of the timing model. Our models will have the same features.

(3) The models are divided into tens to thousands of interacting submodels. Our models will be divided, though into a far smaller number of <u>functional</u> submodels.

(4) Interactions between the submodels of (3) are infrequent but strong. Again, this will be a feature that is highlighted in our models.

3.1.2. Signal-Processing Models

We divide signal-processing models into two broad categories:

(1) models that specify the ECG on a digitizer sample by sample basis and

(2) models that specify the ECG on a wave-arrival event by event basis.

3.1.2.1. Digitizer-Sample Models

Many authors have used digitizer-sample models of individual ECG beats. For example, Marcus (Marcus, 1982) and Uijen and his associates (Uijen, 1979) used such models to evaluate the performance of QRS-complex detectors. Sornmo and his collaborators (Sornmo, 1981) used such a model to evaluate the performance of QRS-complex feature extractors for pattern recognition. Murthy and coworkers (Murthy, 1979) used such a model to develop cepstral-based features for discriminating different QRS-complex morphologies by pattern recognition. Because they do not model cardiac rhythms, we do not review these models.

Several authors have considered digitizer-sample models of complete rhythms (Borjesson, 1982; Haywood, 1970; Richardson, 1971). These models are oriented toward R wave detection. They make very limited rhythm assumptions and do not include other waves. The fundamental problem with the rhythm assumptions is the lack of a simple underlying physical model to generate the probability mass function (pmf) on the wave arrival times. While the actual physical mechanism may be describable in simple terms, it may generate a very complex R-wave arrival distribution, which one cannot practically assume as the first principle

since it would create an unmanageable number of free parameters. Therefore, when the pmf is simply assumed, one is driven to use a very simple pmf. For similar reasons, one is forced to consider only one type of wave arrival. Thus, while these models may be of value in improving R-wave detection and characterization performance (perhaps combined with methods for QRS-complex morphology analysis as mentioned above), the level of aggregation and simplification involved precludes their use in extracting more detailed information concerning cardiac rhythms.

3.1.2.2. Wave-Arrival Models

The remaining models to be described in this review specify the ECG on a wave-arrival event by wave-arrival event basis.

Gersch and his associates (Gersch, 1970, 1975) and separately Tsui and Wong (Tsui, 1975) have developed three-state Markov chain models of R wave arrival times in various arrhythmias. The states of the chains are the intervals between successive R waves, coarsely discretized into three ranges: S (short), R (regular), and L (long).

In Gersch's work (Gersch, 1970, 1975), the sample size is fixed and a maximum-likelihood estimate of the correct rhythm is made. In (Gersch, 1970), a formula for an information-theoretic distance between the various pairs of rhythm models is derived. In that paper, six rhythms are considered:

(1) atrial fibrillation,

(2) atrial premature contractions and ventricular premature contractions,

(3) bigeminy,

(4) sinus tachycardia with occasional bigeminy,

(5) sinus tachycardia, and

(6) ventricular tachycardia.

Transition probability matrices are estimated from training data in which QRS complexes were detected by a threshold on the first derivative of the voltage with respect to time. The resulting R-R intervals were then grouped into the three ranges, S, L, and R and transition rates were estimated. Application of the algorithm to the training data resulted in 100 percent correct performance.

In (Gersch, 1975) a bound, exponential in sample size, is developed on the probability of misclassification in a binary hypothesis test between two Markov chain models. In this paper, two rhythms are considered:

(1) Atrial fibrillation (AF).

(2) Atrial fibrillation with occasional PVC's (AFOCC). "Occasional" is defined as less than 1 PVC per 25 R-R intervals.

Transition probability matrices were estimated as in (Gersch, 1970).

Two experiments were performed and reported. In the first experiment, the data used to estimate the transition probability matrices was

classified by the algorithm. In summary, detection was correct in 30 out of 42 cases (71%) when the sample size was 100 R-R intervals and correct in 17 out of 21 cases (81%) when the sample size was 200 R-R intervals.

In the second experiment, the AF transition probability matrix was fixed and the AFOCC transition probability matrix was varied in order to consider how changes in the relative self-entropy of the AF and AFOCC hypotheses effected the probability of error. The actual transition probability matrices (tpms) used were not completely reported. Extensive simulations were performed which indicated that performance increased rather dramatically when the difference in the self-entropies increased.

Finally, in a third incompletely reported experiment, the sensitivity of the probability of error to small changes in the definition of the R, S, and L symbols in terms of the observed interval durations was evaluated. This was done using the real data. Gersch (Gersch, 1975) summarized the results by stating that "classification error performance was not very critically dependent upon the transformation parameters."

Tsui and Wong (Tsui, 1975) use the same type of three-state Markov chain utilized by Gersch (Gersch, 1970, 1975). However, instead of considering maximum likelihood estimation of the correct rhythm, as is done by Gersch (Gersch, 1970, 1975), Tsui considers only binary hypothesis testing problems and uses Wald's sequential probability ratio test (SPRT). He derives a formula for the expected number of R-R intervals observed before the SPRT terminates and uses this as a measure of the separability of the two hypotheses.

Using unreported methods, Tsui generates probability transition matrices for three rhythms:

(1) atrial fibrillation,

(2) normal sinus rhythm, and

(3) premature atrial and ventricular contractions in the presence of normal sinus rhythm.

Calculation of the bound for the expected number of observations until termination of the SPRT for each pairwise combination of these three rhythms indicated that the average number of observations to achieve classification would not exceed 20 observations when the probability of error was restricted to less than 0.1 percent. If the bound is tight, then this formulation of the problem leads to a signal-processing algorithm that is rather slow to make a decision from the application point of view. No experimental results on real data were reported.

For an application of Gersch's work (Gersch, 1970, 1975) to the detection of atrial fibrillation, see Shah et al. (Shaw, 1977). For an extension of the Gersch-Tsui-Wong model see White (White, 1976). For a deterministic finite-automata approach using the S, R, and L symbols as input see Hristov et al. (Hristov, 1971).

Note four limitations of the Gersch-Tsui-Wong model:

(1) It cannot correctly model all R-R interval patterns. For example, consider second-degree AV block Type II with 4:3 conduction. This is shown schematically in Figure 3.1. In terms of S, R, and L, the rhythm is:

... R R R L R R R L R R R L ...

In the Gersch-Tsui-Wong model, every occurrence of R (for example) is equivalent to every other since it corresponds to a unique state and the transition probabilities out of that state are constant. However, in the 4:3 rhythm exhibited above, each R is obviously not equivalent. This point is also sketched by Strand (Strand, 1973). Therefore consideration of more complex states is necessary.

(2) The transformation to S, R, and L obscures certain rhythms. For example, in an interpolated PVC, two S intervals would be observed. If more detail were retained, it would be observed that these two intervals add up to one of the usual R-R intervals. However, this detail is lost.

(3) The model has no physical basis. It is essentially an elegant, empirical way to account for certain statistics in the observed interval sequences but it is not a description of why these statistics occur.

(4) Several further limitations are described in a general critique of a number of methods at the close of this subsection. Briefly, the limitations are the inadequate observation models which do not allow for missed waves or extra detections, the failure to use rhythm information to assist in wave detection, and the lack of a rational way to add additional waves (e.g. P waves) to the model.

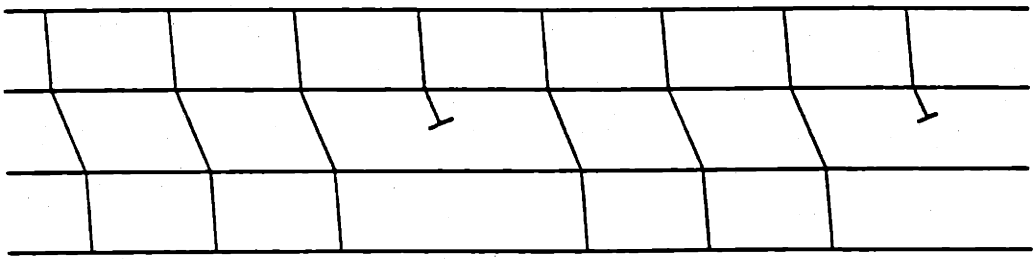


Figure 3.1 Ladder Diagram for Second-Degree AV Block with 4:3 Conduction.

In the previously-discussed ECG models, in which the ECG was modeled in terms of wave-arrival events, the inter-event times were coarsely discretized into three levels. Now we consider several models in which the inter-event times are not discretized. Because of the vast increase in the number of possible inter-event times in this approach, the Markov chain models of Gersch-Tsui-Wong are no longer practical since they require a state for each possible inter-event time.

Grove, Haywood, Richardson, Murthy, and Harvey (Grove, 1978; Haywood, 1977; Richardson, 1976) have considered autoregressive and one-lag nonlinear Markov prediction models for R-R intervals. The two models are:

(1) Autoregressive:

$$x(n)-\mu=\sum_{m=1}^{p} a_m(x(n-m)-\mu)+e(n) \quad .$$

e(.) a zero-mean white Gaussian sequence with variance σ^2 and x(.), e(.) $\in$ R

(2) Nonlinear One-Lag Markov:

$$x(n+1)=f(x(n))+e(n)$$

e(.) a zero-mean white Gaussian sequence with variance σ^2 and x(.), e(.) $\in$ R

In their work with the nonlinear one-lag Markov model, the steady-state probability density for the state of the process is identified with the experimental histogram of R-R intervals. This work was motivated by noting the bimodal character of some such R-R interval histograms. By fitting the model in this way, the investigators ignore the temporal patterns in the data. That is, they match the statistics on x(n), although the important statistics are the statistics on x(n) conditioned on x(n-1), x(n-2), ... Furthermore, the nonlinear mechanism has no physiological interpretation.

The work of Grove, Haywood, and Richardson with autoregressive models seems inconclusive. The model is fit by least squares techniques. The results are complicated by the question of nonstationary statistics. Furthermore, the empirical time-series-analysis approach does not use the available a priori information concerning the dynamics of ventricular rhythms, such as in ventricular bigeminy. Correspondingly, the parameters of the autoregressive model have no physiological significance. In none of these papers are any applications made of the models.

Most rhythm analysis systems have some model of the R-R interval duration in order to predict whether the just observed interval is or is

not "normal". A standard approach is to use the average of the interval durations of a certain number of "normal" preceding beats (Feldman, 1971; Schluter, 1981; Shah, 1977). This obviously only makes sense for a limited class of rhythms. Other authors have models to address very specific subquestions within a larger rhythm analysis algorithm. The work of Shah (Shah, 1977) on atrial fibrillation was noted previously. To detect atrial fibrillation, Schluter (Schluter, 1981) uses the prediction errors for a set of estimators that estimate the current R-R interval by computing averages at different lags of prior intervals. Large errors for all estimators imply that the rhythm is atrial fibrillation. In none of the investigations mentioned above is the model in any sense a complete description of all that is known a priori about a cardiac rhythm

Gustafson and his collaborators (Gustafson, 1977, 1978a, 1978b, 1978c, 1979, 1981) have developed an extensive set of models and signal-processing algorithms for ECG interpretation. This thesis is motivated in large part by the successes and limitations of their work. Therefore, it is described in more detail than the other investigations mentioned in this review. However, even the following description includes only a portion of the rhythm classification part of their work.

Gustafson's models are linear vector Markov models. The components of the vector state are different interval durations (e.g. R-R, P-R). Detection of the ECG wave-arrivals necessary to compute the intervals is described in (Gustafson, 1975) and (Wang, 1976) for R waves and (Gustafson, 1979, 1980) for P waves. In the early work (Gustafson, 1977, 1978a, 1978b) only R wave events are considered while in later work

(Gustafson, 1978c, 1979, 1981) P wave events are also included.

In the work based on R-R intervals only, four persistent rhythm classes are defined. Each class corresponds to several separate cardiac arrhythmias whose subclassification usually depends on more information than is available from R-R intervals alone. The exception is the differentiation of sinus bradycardia, normal sinus rhythm, and sinus tachycardia which are distinguished by heart rate within a single model. The four models are Small Variation, Large Variation, Period-Two Oscillator, and Period-Three Oscillator. We describe only the Small Variation model (which is required for the discussion of the transient rhythm models) and the Period-Two Oscillator model. Conceptually, the other two are similar.

In rhythms modeled by the Small Variation model, the R-R intervals exhibit small random deviations around their mean value, which is constant. Among the clinical syndromes included in this class are normal sinus rhythm, sinus tachycardia, and sinus bradycardia. The model is:

$$x(k)=x(k-1)$$

$$y(k)=x(k)+v(k)$$

$$\mathrm{Cov}\{v(k)\}=R_S$$

where the observation $y(k)$ is the observed value of the k^{th} R-R interval and the observation noise sequence $v(.)$ is a scalar zero-mean white Gaussian random sequence.

In the Period-Two Oscillator, the R-R intervals are alternately long and short. Among the clinical syndromes included in this class are ventricular bigeminy and second degree AV block with 2:1 conduction. The model is:

$$x(k)=\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}x(k-1)$$

$$y(k)=[1\ 0]x(k)+v(k)$$

$$\mathrm{Cov}\{v(k)\}=R_2$$

where the observation y(k) is the observed value of the k^{th} R-R interval and the observation noise sequence v(.) is a scalar zero-mean white Gaussian random sequence.

In the work of Gustafson et al., four transient rhythm models are defined as well. These transient events are modeled relative to the underlying normal rhythm pattern. Each occurrence of a transient rhythm is a single isolated change that can extend over one to three beats. The change affects the underlying persistent rhythm, it is not simply a superimposed event. The four transient rhythms are Rhythm Jump, Non-Compensatory Beat, Compensatory Beat, and Double Non-Compensatory Beat. We describe only the Non-Compensatory Beat model. Conceptually, the other three are similar.

Let Θ be the unknown time of occurrence of the transient, γ is the unknown amplitude of the transient, and $\delta(i,j)$ is the Kronecker delta function. The observation y(k) is the observed value of the k^{th} R-R

interval and is given by the observation equation of the Small Variation model above. In events modeled by the Non-Compensatory Beat model, there is an isolated short or long R-R interval. The mean R-R interval, excluding this interval, remains constant. Clinical events included in this class are an atrial premature contraction which resets the SA node ($\gamma<0$), an isolated failure of AV conduction ($\gamma>0$), or sinus arrest ($\gamma>>0$). The model is:

$$x(k)=x(k-1)+\gamma(\delta(\theta,k)-\delta(\theta,k-1)).$$

All of the persistent and transient rhythm models, including those not described above, share several desirable features.

(1) None of these models actually define the lengths of the intervals. Rather, they just state that a specific pattern exists. The optimal estimator estimates the actual lengths for any particular ECG. Thus the lengths are not locked into the model which, for example, allows the models to avoid requiring parameters reflecting heart rate.

(2) The state of the models in some cases retains several of the most recent R-R intervals. This is necessary to model oscillatory phenomena, which cannot be done if the state is only the last R-R interval as in the methods described previously.

(3) All the parameters have physiological meaning; none are set by empirical curve-fitting techniques.

Using standard statistical signal-processing techniques (Kalman Filter, Anderson, 1979; Multiple Model Hypothesis Test, Lainiotis, 1973; and Generalized Likelihood Ratio Test, Willsky, 1976) and insight into the behavior of the different models and algorithms, Gustafson and his collaborators have developed rhythm-classification algorithms for these classifications that are quite successful.

In order to achieve a finer classification of arrhythmias, it is necessary to include some information concerning morphology or atrial activity. In (Gustafson, 1978c, 1979, 1981), the previous work is generalized to include P waves. Now there are two types of wave-arrival events--P waves and R waves. The modeling framework and signal-processing techniques are carried over intact from the prior work based on R-R intervals alone.

This work does not encompass all possible patterns of P waves and R waves. Rather, it is oriented toward rhythms for which the analysis of P-R intervals is appropriate. For instance, sinus arrhythmia is included but third-degree AV block is not. The focus on this class of rhythms allows Gustafson to continue using models in which the time index is a counter of R wave arrivals. This becomes clearer when the basic outline of the algorithm is presented:

(1) Find, using heuristic techniques described in (Gustafson, 1975) and (Wang, 1976), the next R wave.

(2) Using statistical techniques (Gustafson, 1979, 1980), subtract out the R wave and locate all P waves (before, in, and after the R wave) in the beat.

(3) Define the P-R interval for the beat to be the interval calculated from the P wave whose location is nearest the R-wave location. The P wave used in the P-R interval calculation may come from the set of P waves detected in (2) for one of the neighboring beats if, for example, no P wave is detected in the current beat. Finally, if required by the previously described rules, a single P-wave location can be used to calculate the P-R interval for more than one R wave.

(4) Use the R-wave-P-wave pair of (3) as input to the rhythm analysis.

The models for this work are much less convincing than the models based on only R wave arrivals for several reasons:

(1) Because each increment of the discrete time index k in the persistent rhythm model now generates a new P and a new R observation, the models are not able to properly model rhythms with unequal P and R rates. Thus, for example,

(a) There is no model for second-degree AV block with 2:1 block that includes the extra (unconducted) P waves.

(b) Similarly, for third-degree AV block, there is no model that includes the P waves.

(2) The transient rhythm models, being additive terms to the persistent rhythm model for normal rhythm as before, share the same difficulty of having to predict paired P and R waves. Thus, in the "Interpolated R" model, which models an interpolated PVC, the PVC is also forced to have a P wave associated with it. Furthermore, even

without proper modeling of phenomena like retrograde conduction, there are a large number (seven) of transient models and some of the subdivisions are not very meaningful physiologically.

(3) The persistent rhythm models do not have the proper causal relationships. That is, in the model, the R-R intervals are the basic cycle and the P wave locations are specified in terms of P-R intervals. Physiologically, this is backwards.

Finally, we have two observations that apply to both the R-R interval only and the R-R interval plus P wave models:

(1) Transient rhythms are only considered relative to the normal-rhythm persistent rhythm model. Clinically, this is the most important case. However, it is also necessary, and very difficult, to detect certain transient rhythms relative to non-normal-rhythm persistent rhythms. For example, the detection of PVC's relative to a persistent rhythm of atrial fibrillation is important.

(2) Transient rhythms due to the same cardiac abnormality but whose (secondary) ECG manifestations differ must be modeled separately. For example, whether a premature beat resets the SA node or not separates into two different transient events.

In summary, the work of Gustafson and his collaborators on persistent and transient rhythm classes described by R wave arrivals is very elegant and successful. The extension to classes requiring P wave arrivals is limited in the rhythms it can model and its relationship to underlying cardiac phenomena is, in places, quite tenuous. Throughout, a striking aspect of Gustafson's work is the absence of free parameters

requiring empirical estimation.

We now discuss several features common to many of the models described above.

(1) All models described are wave-arrival event models. Therefore, from a signal-processing point of view, these are models of the output of a wave-detector preprocessor. When evaluating an ECG, cardiologists use their accumulated rhythm information to assist in wave detection, especially for the low amplitude P waves. Thus there is feedback of high-level rhythm information into the wave-detection preprocessor. Similar feedback occurs in optimal signal-processing schemes based on the true observable quantities--the ECG voltages. This feedback does not occur in any of the investigations reviewed here because a preprocessor-postprocessor structure is implied. (One exception is a simple ad hoc type of feedback, discussed previously, which is used by Gustafson (Gustafson, 1979) to aid in the detection of P waves). This is not surprising since all of the investigations described here emphasize R waves for which feedback is not very important. Our interest in more detailed rhythm classification implies a concern with P waves for which, as mentioned previously, such feedback is essential.

(2) Closely related to (1) is the issue of how to model the observation. Specifically, in reality, there are two types of errors the preprocessor can make: small timing errors and missed or extra arrivals. However, for the models described here, at most small timing errors are included. This is a natural simplification in the case of R wave arrivals but it is inappropriate for P wave

arrivals.

(3) All of the models described here use R wave arrivals as the "time" index. To model general arrhythmias in detail, it is necessary to model a variable number of P waves for each R wave. In all of the models described here, such a change would require the addition of detail in the state space. However, none of the approaches gives any guide to how the enlargement should be performed, and in any case, one would expect such extensions to have an even more tenuous connection with cardiac behavior.

3.1.3. Summary of the Literature Review

From the perspective of this thesis, none of the existing models with a signal-processing orientation models the ECG in sufficient detail. Furthermore, it is not clear how to extend them. We believe that the difficulty in extending the existing models is the inability to rationally increase the level of internal detail. For example, the Gersch-Tsui-Wong approach does not provide any guidance concerning what new states should be added to the Markov chain if P waves are to be included in the model. On the other hand, the models with a physiological orientation, while they contain the detail, are unmanageably complex for signal processing. The goal of our modeling is to achieve a balance between physiological accuracy and level of detail appropriate for signal processing. In fact our modeling methodology, which is described next, allows great flexibility in trading off model detail versus complexity--from very complex models which are closely related to cardiac behavior to more aggregate models that focus on the major events to

be detected and tracked.

3.2. Introduction to the ECG Model

We have tried to incorporate basic cardiac anatomy and physiology into our models of the ECG. This approach is motivated by three observations.

(1) Estimator intelligibility: If an estimator is based on such a model, it is generally straightforward to explain the behavior of the estimator in terms of cardiac structure.

(2) Low dimensional parameterization: It is generally easier to achieve a parsimonious parameterization of a model if the parameterization is done in terms of physically meaningful variables rather than some ad hoc curve-fitting parameterization. Also, it is generally good engineering practice to develop designs with a minimum of adjustable parameters where, furthermore, each parameter can be directly related to process behavior.

(3) Incorporation of a priori information: The anatomy and physiology of the heart constrain the signals that it can produce. The incorporation of these constraints into a model is of obvious value in achieving high levels of estimator performance.

However, unlike a physiologist, we are concerned with models for signal processing. Therefore, we are concerned with computational feasibility and we desire that the model include only enough detail to allow successful signal processing. In order to achieve computational feasibility, we emphasize decompositions. More specifically, we

emphasize spatial, time-scale, and hierarchical decompositions. In order to control the level of detail, we have developed a class of models suitable for a wide range of detail. The appropriate level of detail can then be chosen based on signal-processing goals.

The spatial and time-scale decompositions have a physiological basis. The basis of the spatial decomposition is the anatomic division of the heart into subunits which have relatively few interactions. The basis of the time-scale decomposition is the different rates at which events within an anatomic subunit and interactions between anatomic subunits occur.

The hierarchical decomposition we have developed is not entirely accurate physiologically but is flexible enough to allow us to mimic cardiac behavior quite accurately. Its two levels are based on a separation of discrete events from the generation of ECG waveforms. The discrete events are the interactions between anatomic subunits of the heart and the initiation of waves in the ECG. This hierarchy separates the high-level events we wish to detect and track from the actual observed voltages, and this separation is useful for signal-processing purposes.

A three-submodel example of a spatial and hierarchical decompositions is shown in Figure 3.2. The square boxes at the upper level of the hierarchy comprise the discrete-state physiological model, which captures the sequential evolution of high level events in the heart. Each submodel represents a functional anatomic structure (e.g. the atria, the ventricles, etc.). The directed solid lines indicate the initiator and recipient of control inputs, which we call interactions.

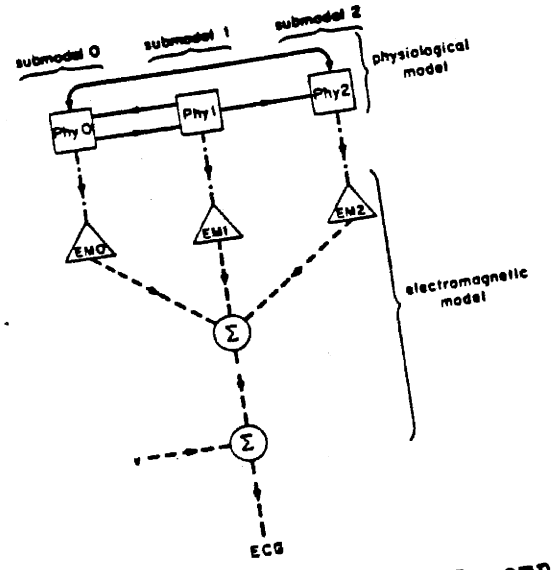


Figure 3.2 Spatial and Hierarchical Decompositions.

For example, the transmission of a depolarization wave from atria to AV node might be modeled via an interaction in which the present state of a submodel representing the atria causes a transition in the AV-nodal submodel, representing the initiation of the AV-nodal depolarization.

The triangular objects in Figure 3.2 comprise a part of the <u>electromagnetic model</u> which models the actual observed waveforms. Each submodel corresponds to the generation of the ECG contribution from a particular anatomic structure of the heart (e.g. P waves from the atria). (The remainder of the electromagnetic model describes how these contributions are combined into the ECG and how the ECG is observed). The dashed lines indicate the control of the electromagnetic level by the physiological level of a single submodel. These inputs are used to initiate the generation of waveforms in the observed ECG. For example, the occurrence of a particular transition in the physiological portion of an atrial submodel might initiate the generation of a P wave in the corresponding electromagnetic submodel. Note that the electromagnetic level does not affect the physiological level and that no

electromagnetic submodel affects any other electromagnetic submodel. Note also that the unidirectional interactions between levels occur only between the two levels of a single submodel (i.e. events in a submodel corresponding to the atria cannot initiate an R wave).

Finally, recall from Section 2.6 that the abnormal aspects of an arrhythmic ECG occur because there is some abnormal anatomic substructure in the heart which

(1) makes a direct abnormal contribution to the ECG and/or

(2) interacts with other normal parts of the heart causing them to make an abnormal contribution to the ECG.

In our models of arrhythmias, we take the same approach. That is, we begin with a normal rhythm model which is transformed into an arrhythmia model by altering the appropriate submodel. The contribution of the altered submodel and its interactions with the unaltered submodels create the arrhythmic ECG. In order for the interactions to occur, we often must generalize the normal submodels. The alterations are to include properties which were left out initially because, in the normal rhythm, they were unobservable or did not occur.

3.3. The Upper Hierarchical Level

In this section we discuss the upper hierarchical level, which we have called the physiological model. Because this level is concerned with discrete events, we have chosen a Markov chain model. Based on the spatial decomposition ideas, we have chosen a highly structured class of chains described in the following.

The state space of the Markov chain is the cross product of a set of spaces corresponding to the "states" of subprocesses which comprise the overall chain. Each subprocess corresponds to one of the anatomic subunits of the heart. Furthermore, there is a direct correspondence between each state of a subprocess and a physical state of the corresponding anatomic subunit. We call each subprocess a submodel. We often refer to an element in the subprocess "state" space as a state.

In this chapter we present an informal discussion of the mathematical formalism associated with our modeling approach. We made this choice so that we can provide a clear picture of the important features of our models without distracting mathematical detail. In Section 4.1 we give a detailed presentation as a prelude to the discussion of estimation algorithms. However, no where do we give a rigorous presentation.

Let x_n be the state of the overall Markov chain which consists of a set of N subprocess denoted x_n^i, i=0,...,N-1 and let p(n) be the pmf on x_n. Since x_n is a Markov chain, p(n+1)=Ap(n) where A is a stochastic matrix. That is, the elements of A are the values of $p(x_{n+1}|x_n)$ as x_n and x_{n+1} range over all allowed values.

A key feature of our models is that $p(x_{n+1}|x_n)$ has a great deal of structure. Specifically:

(1) Given x_n, the transitions of each of the component subprocesses are independent. That is,

$$p(x_{n+1}|x_n)=p(x^i_{n+1},\ i=0,\ldots,N-1|x_n)$$

$$=\prod_{i=0}^{N-1} p(x^i_{n+1}|x_n).$$

(2) For each subprocess there are far fewer values of $p(x^i_{n+1}|x_n) = p(x^i_{n+1}|x^j_n,\ j=0,\ldots,N-1)$ than there are values of x^j_n, $j\neq i$. That is, we assume that $p(x^i_{n+1}|x_n) = p(x^i_{n+1}|x^i_n,h^i_n)$ where $h^i_n = h^i(x^j_n,\ j\neq i)$ denotes the net interaction of all other subprocesses with the i^{th} subprocess. Here h^i_n may take on values 0, 1, 2, ... but the assumption is that the number of possible values of h^i_n is far less than the number of possible values of $\{x^j_n,\ j\neq i\}$.

(In an abuse of terminology, we often refer to the individual subprocess as Markov chains).

Before discussing the connection between the two levels of our hierarchy or the mathematical structure of the lower hierarchical level, we now consider a very much simplified model for normal rhythm in order to fix these ideas about interacting chains. This model has two submodels, corresponding to a division of the heart into two anatomic substructures. The first anatomic substructure, which we call the SA-atrial substructure, is composed of the SA node and atria. The second, the AV-ventricular substructure, is composed of the AV junction and ventricles. As in the normal heart, the SA-atrial submodel originates interactions with the AV-ventricular submodel. This interaction corresponds to a supra-ventricular depolarization propagating through

the AV junction in an antegrade direction. For simplicity we assume that the AV-ventricular submodel never initiates an interaction with the SA-atrial submodel. This means that no retrograde conduction through the AV junction is modeled. The two submodel chains are shown in Figure 3.3. Because we discuss this example further after presenting the connection between the physiological and electromagnetic models, these connections are included in Figure 3.3. The interpretation is as follows.

In the SA-atrial submodel (Figure 3.3(b)), the state transition from 0 to 1 represents the firing of the SA node and the atrial depolarization. State 1 represents the SA node immediately after repolarization. State 0 represents the SA node when its membrane potential is just below the firing threshold. The time required for the state to travel from state 1 to state 0 models the random time required for the slow ion leakage to decrease the membrane polarization from fully repolarized to threshold. Finally, by assuming that the atrial conduction velocity is infinite (an idealization), state 1 also represents the excitation of the AV node by the atrial depolarization.

Now consider the AV-ventricular submodel (Figure 3.3(c)). That state 1 (in the SA-atrial submodel) represents the excitation of the AV node is reflected in the differing probabilities assigned in the AV-ventricular submodel depending on whether the SA-atrial-submodel state is or is not in state 1. AV-ventricular-submodel state 0 represents the fully repolarized state. If the AV-ventricular submodel is in that state and the SA-atrial submodel moves into state 1, then the AV-ventricular-submodel state transitions into state 1. This transition

models the excitation of the AV node by the atrial depolarization. If the AV-ventricular substructure is not receptive to a depolarization (i.e. is refractory), then the submodel state will not be in state 0 and the change in the probabilities due to the SA-atrial-submodel state occupying state 1 will have no effect on the evolution of the AV-ventricular subprocess. Then, in the AV-ventricular submodel, the time required for the state to travel through states 1 and 2 represents the (deterministic) AV-junctional delay time. The transition from state 2 to state 3 represents the initiation of ventricular depolarization. Finally, the time required for the state to travel through states 3, 4, and 5 represents the (random) AV-junctional and ventricular repolarization time. After the repolarization is completed, the state traps in state 0 awaiting another excitation from the SA-atrial submodel.

3.4. The Lower Hierarchical Level

We now discuss the lower level in our hierarchy, which we call the electromagnetic model. The spatial decomposition that was imposed on the upper hierarchical level is also imposed on the lower hierarchical level. As in the upper hierarchical level, we use the term submodel for the decomposed elements. The reason that the hierarchical decomposition carries through is that the individual waveforms in the ECG that are modeled by the electromagnetic level are each due to a single anatomic subunit.

As mentioned previously, the strict hierarchical nature of the physiological/electromagnetic model decomposition is reflected in the facts that:

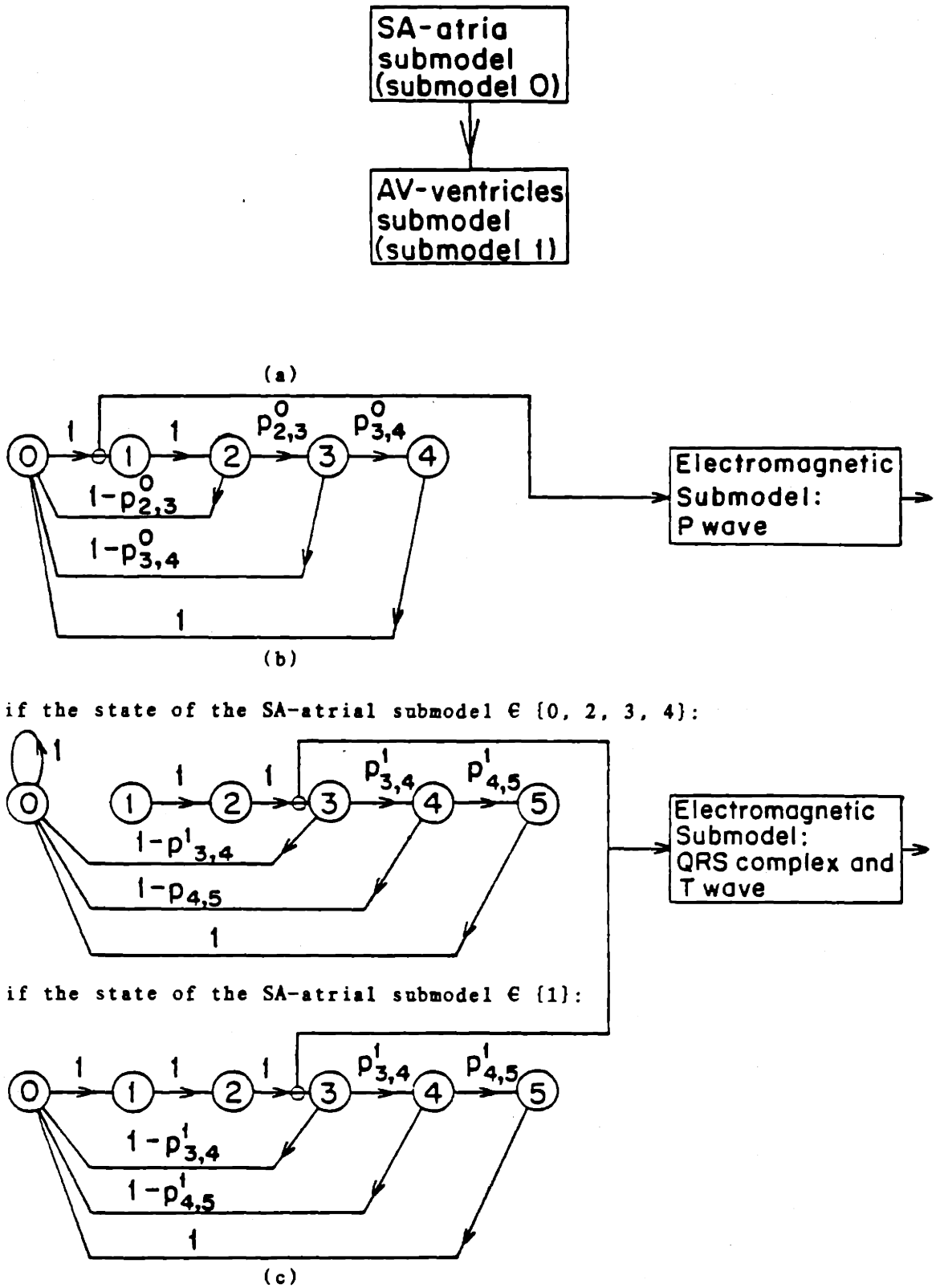


Figure 3.3 A Simple Model for Normal Rhythm.
(a) Submodel Block Diagram.
(b) SA-atrial Submodel.
(c) AV-ventricular Submodel.

(1) submodels within the electromagnetic model do not interact,

(2) each electromagnetic submodel is driven only by the corresponding physiological submodel, and

(3) no electromagnetic submodel affects any physiological submodel.

These relationships are illustrated diagrammatically in Figure 3.2.

Each state of each physiological submodel has a physical interpretation. Certain transitions between states correspond to the initiation of waves, so these transitions are used to drive the corresponding electromagnetic submodel. The output of each of the electromagnetic submodels is a linear superposition of signals with shifted origins. The unshifted signals are called signatures. The origin is the time at which the initiating transition in the corresponding physiological submodel occurs. Each signature is a shift-invariant finite-durational deterministic function with additive white zero-mean Gaussian noise (signature noise) where the additive noise is independent from one occurrence of the signature to the next.[1] Finally, the outputs of the individual electromagnetic submodels are linearly superposed and the result is observed in additional, exogenous, white Gaussian noise (observation noise).

---

[1] The statement that the signature is a deterministic signal observed in additive white zero-mean Gaussian noise is exactly equivalent to the statement that the signature is a white Gaussian signal. That is, the signature could equivalently be described as a shift-invariant finite-durational white Gaussian signal. The reason for using the prior, unwieldy description is that it emphasizes the nonzero mean.

To be somewhat more precise, let $S^i_{j,k}$ be the signature from the i^{th} electromagnetic submodel when the i^{th} physiological submodel makes a transition from j to k. Let v be the white Gaussian observation noise. The observation y is then

$$y(t) = \sum_i \sum_n S^i_{x_{n-1},x_n}(t-n) + v(t).$$

To illustrate our ideas concerning the interaction of the physiological model with the electromagnetic model, let us continue with the discussion of the simple model of normal rhythm that is illustrated in Figure 3.3. In the SA-atrial submodel (Figure 3.3(b)), as discussed previously, the state transition from 0 to 1 represents the firing of the SA node and the atrial depolarization. Thus, as indicated in the diagram, the electromagnetic-model response to this transition is the P wave of the ECG. (We have chosen not to model the low-amplitude atrial repolarization wave since it is obscured by the much larger voltages of the succeeding QRS complex. However, if we wished to include it, we could, for example, model the repolarization wave as deterministicly coupled to the P wave. The electromagnetic-model response to this transition would then be the entire depolarization-repolarization sequence). The form used in Figure 3.3(b) will be our standard way of indicating what transition has a particular effect on the electromagnetic model. The electromagnetic-model response to the other transitions, for example to the state transitions from 2 to 3 or from 2 to 0, is identically zero and hence not indicated.

In the AV-ventricular submodel (Figure 3.3(c)), as discussed

previously, the transition from state 2 to state 3 represents the initiation of the ventricular depolarization. Hence the electromagnetic-model response to the corresponding transition is the QRS complex and the T wave. Here, we are modeling the QRS complex and T wave as deterministicly coupled waveforms--the ST interval duration is not random. Note that a more complex model of the same type could allow a random coupling. The electromagnetic-model response to the other transitions generated by the AV-ventricular submodel is identically zero and hence is not indicated.

In the following paragraphs, we make several comments on the electromagnetic model.

(1) Note that some anatomic subunits do not cause waves in the ECG (e.g. the AV node). For such subunits, the corresponding electromagnetic submodel does not exist. Similarly, in the remaining physiological submodels, most transitions do not correspond to the initiation of a new wave in the ECG. Rather, they model the timing between wave and interaction initiations. Therefore, most transitions of a physiological submodel have no effect on the corresponding electromagnetic submodel.

(2) The randomness in the signatures models the beat-to-beat variation in the morphology of the ECG waveforms for a single subject. The white-Gaussian assumption clearly limits the type of randomness that can be modeled by this mechanism. However, a second mechanism is also available. Consider a generalized electromagnetic submodel in which the deterministic portion of the signature is replaced by a random function. The random function is constructed by choosing

one of a finite set of fixed deterministic functions. The pmf for the choice among the elements in this set is taken to be time invariant and the choice is remade each time the signature occurs. This situation is shown in Figure 3.4(a). By augmenting the subprocess state space and making appropriate use of the choice pmf, we can lift the process of choosing from the electromagnetic model into the corresponding physiological submodel, as shown in Figure 3.4(b). Using identical methods, one can also include variation in the statistical properties of the additive noise component of the signature.

(3) The additive white Gaussian noise v that is added to the summed output of the electromagnetic submodels models all noncardiac contributions to the observed signal. Such noise sources include:

(a) electromyogram signals,

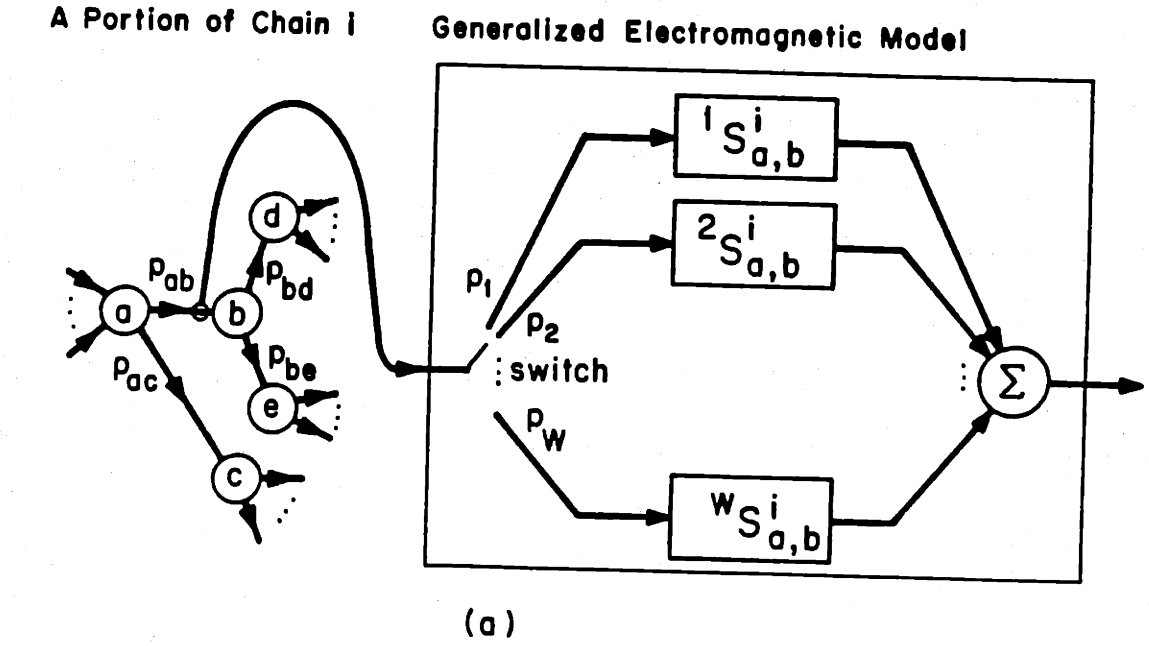
(b) electrode artifacts,

(c) 60 Hz. lead pickup,

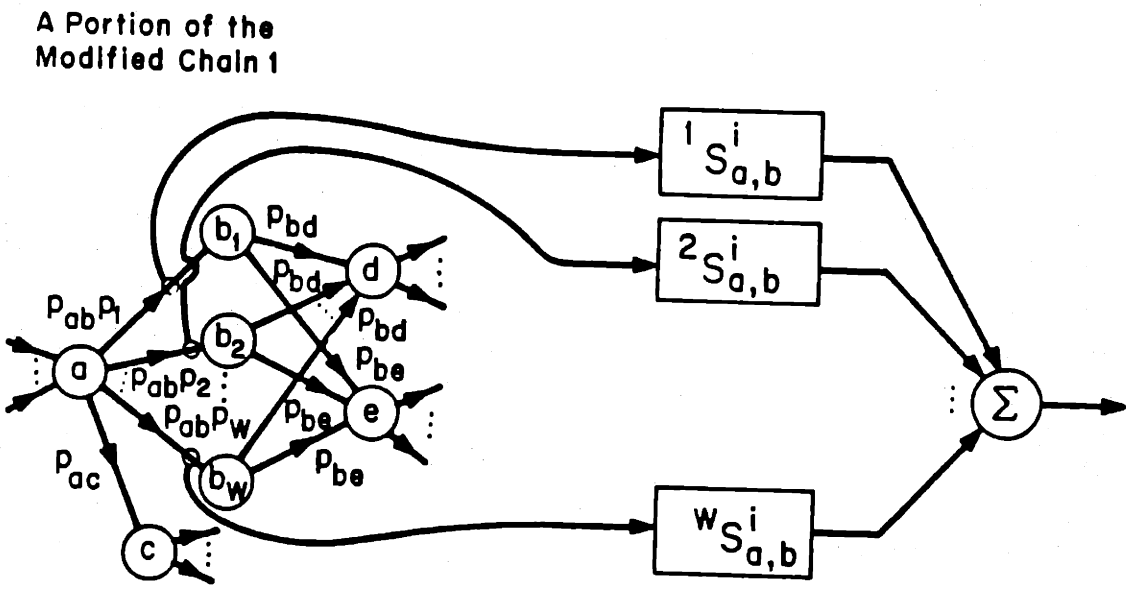
(d) digitization quantization error, and

(e) electronics noise.

The white-Gaussian assumption is clearly a great idealization. Given the linearity postulated throughout the electromagnetic model, this is a natural assumption and we conjecture that it may be sufficient for signal processing. However, the primary reason we have chosen to use this assumption is to avoid obscuring the main focus of our investigation--demonstrating the utility of our



(a)



(b)

Figure 3.4 Lifting the Generalized ELectromagnetic Model Into the Physiological Model.
(a) Physiological Model with Generalized Electromagnetic Model.
(b) Generalized Electromagnetic Model Lifted Into the Physiological Model.

approach to modeling and estimation. If a more complex noise model is warranted, it certainly can be included.

(4) The linearity assumption within each electromagnetic submodel has a physical basis. Specifically, the physical processes are low-field-strength electromagnetic processes so that Maxwell's equations with linear constitutive relations apply. These equations are linear in the charge and current sources. Furthermore, we can choose states and their physiological interpretations in each submodel of the physiological model such that each transition which has a non-zero effect on the electromagnetic model marks the depolarization of a different volume of heart tissue. Therefore, the response should also be linear in the transitions. The linearity assumption for the combination of the outputs of the electromagnetic submodels has the same justification.

(5) The dimension of the observation is not specified (i.e. we could consider one or several ECG leads at one time), and correlations between the components (i.e. different leads) are modeled by the off-diagonal elements of the covariance matrices that determine the Gaussian noises.

(6) The Markov chain cycle interval need not equal the signature sampling interval. Typically, the Markov chain cycle interval can be taken to be substantially larger than the signature sampling interval. This reflects the difference in time scale between interaction events (which set the Markov chain cycle interval) and events internal to the anatomic submodels (which set the signature sampling interval). It is for this reason that the variable t in the

previous equation defining y is not described in detail (i.e. it could be continuous or discrete and if discrete, perhaps with a sampling interval not equal to the Markov chain cycle interval). For an example, see the Wenckebach model of Section 3.8.

(7) Signatures can only be initiated at Markov chain cycles. However, by a suitable artifice (replicating signatures and inserting one or more leading zeros into the signature), signatures can be made to start at arbitrary signature samples. This is done in the Wenckebach model of Section 3.8.

(8) An individual signature may last several Markov chain cycles.

(9) A single state transition in the overall model may initiate as many signatures as there are submodels.

3.5. Level of Modeling Detail

The level of detail in our model is controlled in two places.

(1) The number of submodels can be varied. If there are hundreds to thousands of such submodels, then our approach is essentially a finite-element model of the heart. Such models were described in Section 3.1, and thus we can in fact view these models as (extreme) special cases of our models. Note that in such cases, the electromagnetic model is essentially trivial, as the physiological model contains all of the information, and individual state transitions generate small increments of signature waveforms.

(2) The number of submodel transitions which initiate non-zero waves in the ECG can be varied. That is, each wave can be broken up into subwaves which are each initiated by a different transition.

Not only can the level of detail be controlled, but we believe that a clinically appropriate level of detail is achieved at a practical level of complexity. Here we use the ladder diagram (see Section 2.7) as a definition of clinically appropriate detail. In order to substantiate this claim, we will describe several models in detail. In Section 3.7 we present four models which meet or exceed ladder diagram levels of detail and which are still reasonably simple, though physiological accuracy and not simplicity was the primary goal. In Section 3.8 we present a fifth model which also meets or exceeds ladder diagram levels of detail and which is quite simple. In this example, significant attention was paid to achieving as simple a model as possible which could produce realistic tracings of the Wenckebach rhythm.

As a final point, note that our model always specifies the probabilistic properties of the observed signals completely. That is, from our model it is theoretically possible to calculate the implied probability distribution on any set of events defined within the model. In practice, due to unavoidable inaccuracies in the model, it is undesirable to allow estimators to depend on to much of the detail implied by this specification. For example, if the estimator made critical use of the joint pmf for the k^{th} transition in submodel 1 and the $k+94^{th}$ transition in submodel 2, then the estimator would never be successful when processing real data. Therefore, the study of <u>robustness</u> (see Section 4.3.6) is an important issue in estimator performance.

3.6. The Microscopic Model--Structural Elements

The chains of our physiological model are typically constructed from structural elements of a very few types. These structural elements and their interconnections are immediately implementable as parts of chains. There are two fundamental structural elements, which are essentially elapsed time clocks, out of which three other structural elements are constructed.

The first structural element is the delay line (DL). When the state of a chain enters the first state of such an element, denoted i, it undergoes a random time delay and then arrives at the final state denoted o. The delay is called the transit time. The pmf on the transit time is specified and unaffected by events in the other chains.

(To describe complex models, some notation is helpful. We use superscripts to indicate submodels and subscripts to indicate DLs within a submodel. In block diagrams we use the symbol shown in Figure 3.5 for a DL, where i and o have the meanings defined previously).

Given a desired transit-time pmf, a basic issue is how to implement a DL that realizes this pmf. There are two major approaches, distinguished by the topology of the interconnections in the DL. That is, they are distinguished by the class of transitions which can have positive probability. In the tapped delay-line topology, no state can be occupied more than once in a single transit of the DL. Therefore, the set of transitions with positive probability can not contain any loops. In the feedback topology, there must be at least one state which can be occupied at least twice during a single transit of the DL. Therefore,

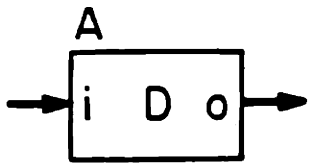


Figure 3.5 Delay Line Symbol.
i = initial state and o = final state.

there must be at least one loop. An example of each topology is shown in Figure 3.6.

Each topology has advantages and disadvantages. The feedback topology cannot represent transit-time pmfs with finite support and it is relatively difficult to choose the transition probabilities. The tapped delay-line topology can represent any transit-time pmf for bounded transit times and the calculation of the appropriate transition probabilities is straightforward. However, the number of states required may be large. Since no transit-time pmf can be exactly represented by both topologies, the choice of topology is intertwined with the choice of desired transit-time pmf. (However, the choice between the two topologies also has signal-processing implications. These implications, which led us to use the tapped delay-line topology exclusively, are discussed in Chapter 6).

The second structural element is the resettable delay line (RDL). We often use the term delay line as a generic name for both DLs and RDLs. The differences between the RDL and the DL are that there are two different mechanisms for the state to exit an RDL and an RDL has transition probabilities that are controlled, in a very simple and specific

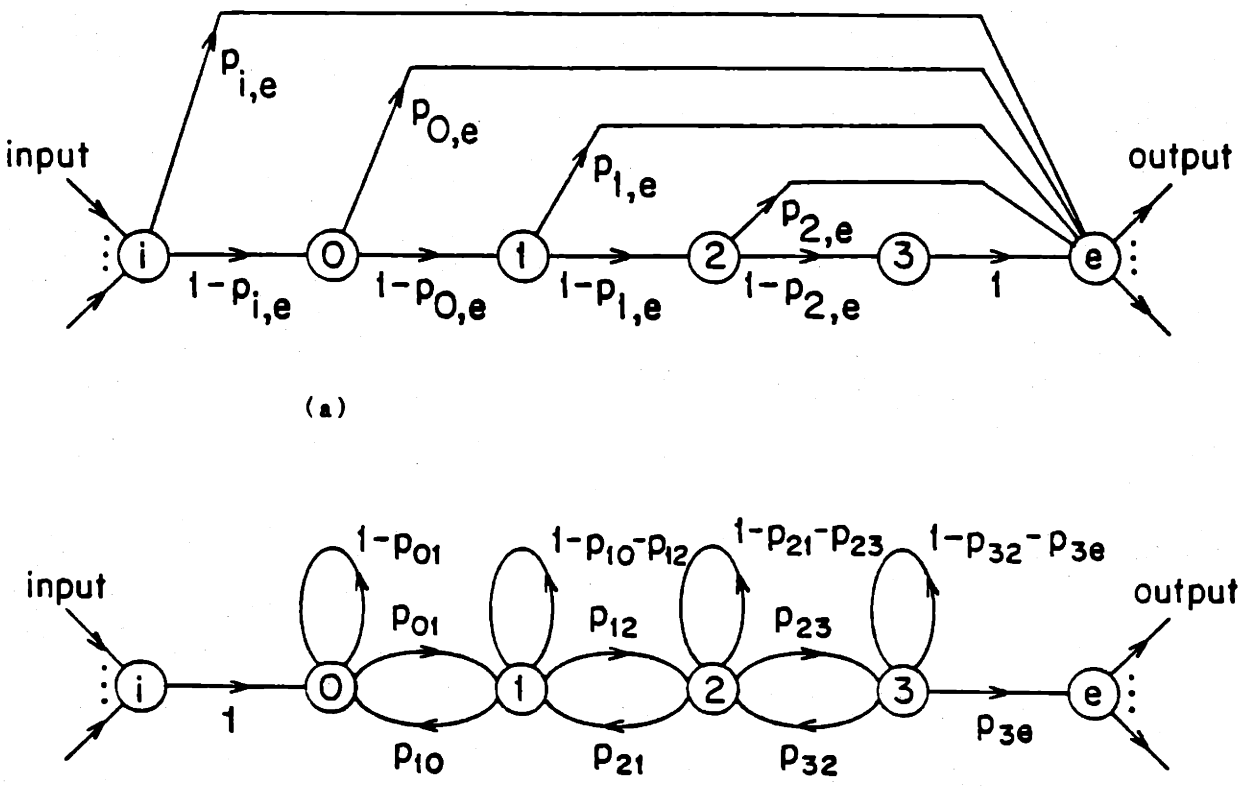


Figure 3.6 Examples of Feedback and Tapped Delay-Line Topologies.
(a) Feedback Topology.
(b) Tapped Delay-Line Topology.

way, by interactions initiated by another subprocess in the overall Markov chain. Specifically, the possible interactions impinging on a chain containing an RDL are divided into two classes denoted normal and abnormal. Within each class the transition probabilities in the RDL are constant. When the interaction is in the normal class, the RDL behaves as a DL, possibly exiting through a final state denoted state o. However, when the interaction is in the abnormal class, a second set of transition probabilities is used for the next transition. The second set of transition probabilities forces the state to leave the RDL and enter a state, external to the RDL, called the reset state.

As in the DL we denote the initial state as state i. As described

previously, the final state (assuming the impinging interactions remained in the normal class) is state o. We define a control input $\underline{c}$ which takes the value R (for "reset") if the current impinging interaction is in the abnormal class and $\bar{R}$ (for "not reset") if the current impinging interaction is in the normal class. In order to describe models containing RDL(s) we use the same notation introduced for describing DLs without distinguishing whether a particular delay-line is a DL or a RDL. In block diagrams we use the symbol shown in Figure 3.7 for a RDL, where "i", "o", and "c" are as defined previously and "r" is attached to the reset state, which is external to the RDL. The implementation issues for the RDL are the same as for the DL.

An RDL can be used, for example, to model the time-delay behavior of the AV node of a model where retrograde AV conduction is allowed. Such a model needs two RDLs, one for the antegrade conduction direction and one for the retrograde conduction direction. The reason for using two RDLs is to allow for the possibility of colliding depolarization waves, one from the atria and one from the ventricles. When such a pair of waves collide, they extinguish each other since both lack polarized tissue into which they can continue to propagate. Therefore, though the AV node is completely depolarized, neither depolarization wave reaches the other end of the AV node where it could excite the succeeding anatomic substructure. Such multisource depolarizations are referred to as <u>fusion</u> depolarizations. In the model, a fusion depolarization corresponds to a situation where the state is in either the antegrade or retrograde delay line and a wave enters from the opposite direction. In this case, the state is shunted to a common reset state that is the start of a third delay line modeling the refractory delay. Thus neither

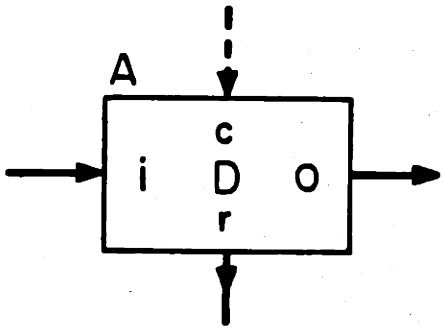


Figure 3.7 Resettable Delay Line Symbol.
i = input, o = normal output, r = reset output, and c = control input.

depolarization wave arrives at the other end of the delay line. With the level of physiological detail modeled here, this shunting occurs immediately after the initiation of the depolarization wave from the opposite direction--the time required for the two waves to meet is not included in the model.

The third structural element is the autorhythmic element. This element is constructed from DL(s) and/or RDL(s). The basic idea is to attach the input and output of a DL together, as in Figure 3.8. If a DL is used, then this specifies the entire chain. If a RDL is used, then it is also necessary to specify the identity of the reset state (see the following examples). The choice of DL versus RDL depends on what physiological process is being modeled.

The SA node can be modeled using such an element. If the SA node is to be modeled as non-resettable, then we base the model on a DL, as in Figure 3.9(a). The initial state (i) of the delay line corresponds to the polarized membrane potential state. The polarization decays toward the firing threshold due to ion leakage. The time required to decay to threshold is modeled by the DL transit time. Thus the DL transit-time pmf is the SA-SA interval pmf. After this delay, the

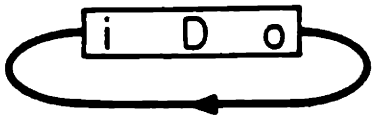
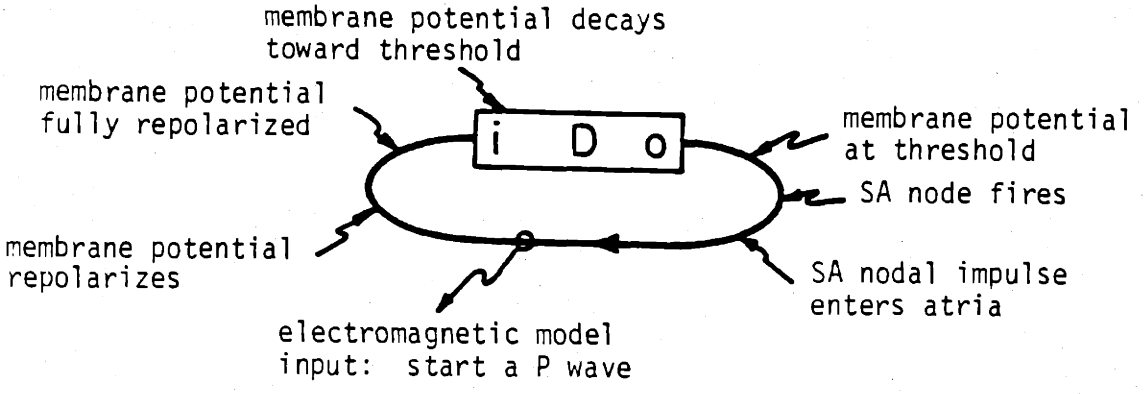


Figure 3.8 Basic Autorhythmic Element.

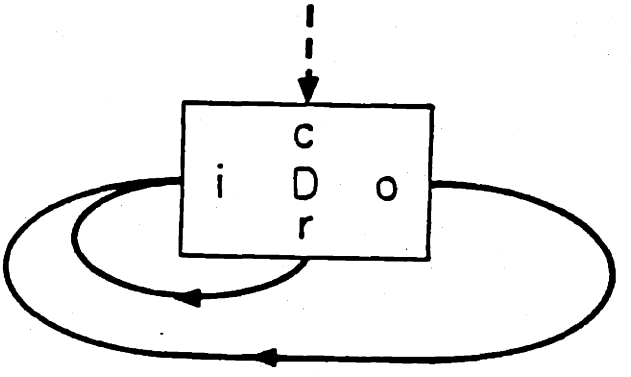
SA node fires, thereby starting a depolarization wave in the atria. This is modeled by the final state (o) of the DL. Occupancy of this state, through the standard interaction mechanism, alters the transition probabilities in the atrial submodel in order to model the initiation of a depolarization wave in the atrial submodel. Finally, the SA node repolarizes. This is modeled by the transition from the final state (o) to the initial state (i) of the DL.

In the case where the SA node is to be modeled as resettable but not stunnable, we use an RDL and make the reset state of the RDL the same as the initial state (i) of the RDL, as in Figure 3.9(b). Thus, when reset, the next SA impulse will occur a time t after the reset where t is a random variable chosen from the SA-SA interval pmf.

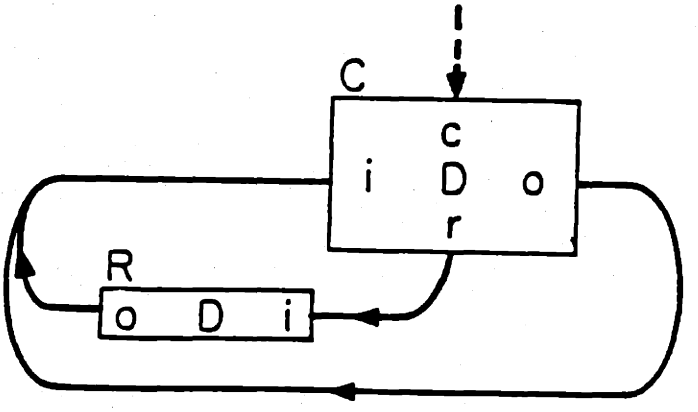
In the case where the SA node is to be modeled as resettable and stunnable, we also use a RDL (labeled "C") but now make the reset state of the RDL the same as the initial state of a second DL (labeled "S") whose final state is connected to the initial state of the first RDL, as shown in Figure 3.9(c). Thus, when reset, the next SA impulse will occur a time t after the reset where t is the sum of two random variables, one from the usual SA-SA interval pmf (delay line C) and one from the stun time pmf (delay line S). Note how all three of these autorhythmic submodels are constructed purely from DL and RDL structural



(a)



(b)



(c)

Figure 3.9 SA-Nodal Models.
(a) Non-Resettable SA-Nodal Model.
The comments indicate the correspondence between the physical processes and the model.
(b) Resettable But Not Stunnable SA-Nodal Model.
(c) Resettable And Stunnable SA-Nodal Model.

elements. This illustrates a basic property of our modeling methodology: we build models for complex phenomena by using interconnections of a small set of elementary models.

The fourth structural element is the passive transmission line (PTL). This element is constructed from a DL or a RDL. We describe the DL case in Figure 3.10. A PTL is a connection of a single state, called the resting state, to the initial state of a DL. The only allowed transition out of the resting state is into the DL. The transition probability for this transition, denoted p, depends on the value of the current interaction impinging on the submodel containing the PTL. The possible values of the impinging interaction are partitioned into two disjoint sets called the autonomous and nonautonomous sets. When the current impinging interaction is in the autonomous set, p=0. That is, the resting state is a trapping state. In the other case (nonautonomous), p>0. In most instances, we take p=1 in the nonautonomous case. This corresponds to a deterministic coupling of events in the chains that initiate and receive the interactions. In other instances, we take p<1 thereby allowing some randomness in the coupling.

A PTL can be used, for example, as a part of an overall submodel for atrial activity. Specifically, consider the SA-nodal submodel (SAN) and the atrial submodel (A) within an overall model of the heart corresponding to normal rhythm (Figure 3.11). In this model the SAN initiates interactions with the A, which is a PTL. For most states in the SAN, the interaction is in the autonomous set. Therefore, p=0 and the resting state is a trapping state in the A. This models the unexcited atria: whatever was happening, for instance the passing of the

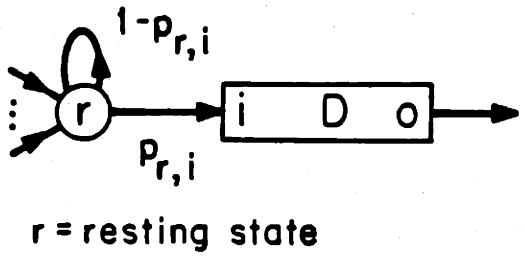
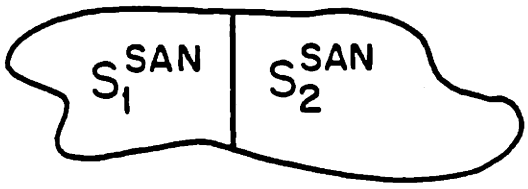


Figure 3.10 Passive Transmission Line.

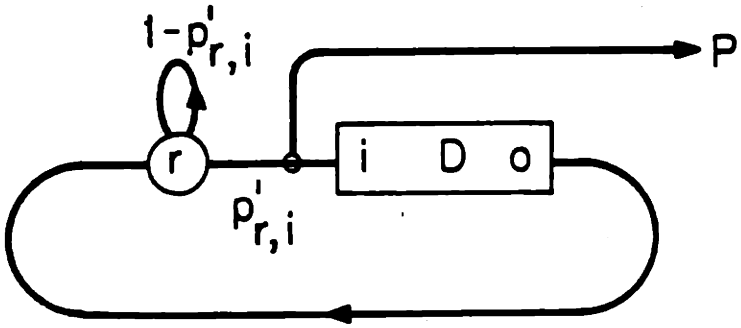
previous depolarization wave, eventually is completed, leaving the atria in the resting state. With the level of physiological detail that is modeled here, it would remain in the resting state forever unless the SAN initiates an interaction setting $p \neq 0$. (In actuality, even if the SAN does not change the interaction, i.e. in the absence of external excitations, the atria would be excited by an internal autorhythmic source, which might be modeled as a separate submodel if deemed important).

For the remaining states in the SAN, the interaction initiated with the A is in the nonautonomous set. Therefore, $p>0$ and, for simplicity in this discussion, we take $p=1$. An interaction in the nonautonomous class models the attempt of the SA node to excite the atria. Suppose that the atria are repolarized and receptive to a depolarization wave from the SA node. In the model this corresponds to the state of the A

SA-nodel submodel: x^0



Atrial submodel: x^1



$$p'_{r,i} = \begin{cases} 0 \text{ if } x_n^0 \in S_1^{SAN} \\ 1 \text{ otherwise} \end{cases}$$

Figure 3.11 A Portion of an Overall ECG Model for Normal Rhythm, Illustrating the Use of a PTL.

occupying the resting state. Therefore, p=1 forces a transition into the DL. This transition initiates a P wave from the electromagnetic model. On the other hand, suppose that the atria are not repolarized when the SA node attempts to excite the atria. In this case the excitation attempt fails. In the model, this corresponds to the state of the A not occupying the resting state. Therefore, p=1 has no effect.

In the previous example, p=1 when the interaction was in the nonautonomous set. A value of p less than one might be used in the situation where an arrhythmia is explained by postulating that a normal or an ectopic node is sometimes unable to excite a region even though, viewed as a whole, the region is not refractory. This behavior is due to inhomogeneities within the region and could be modeled by a finer discretization of the anatomic substructure. The use of p less than one is qualitatively an averaging of this finer detail to yield a coarser model which, however, has sufficient detail to allow us to achieve our signal-processing goals.

The fifth structural element is the bidirectional refractory transmission line (BDRTL), which is a complete submodel. As indicated in Figure 3.12, the BDRTL is constructed from three delay lines (labeled "A", "R", and "F") and one additional state labeled "r".

The state r corresponds to the repolarized resting state of the anatomic substructure. Two of the three delay lines (A and R) are RDLs. One RDL corresponds to antegrade conduction (A) while the other corresponds to retrograde conduction (R). In accordance with these correspondences, the BDRTL attempts to excite the submodel(s) corresponding to the adjacent distal (proximal) anatomic substructure(s) whenever the BDRTL state occupies state o_A (o_R), the final state of the antegrade (retrograde) conduction RDL. We refer to this as the antegrade (retrograde) output of the BDRTL. RDLs are used here in order to model the collision of two depolarization waves, one in the antegrade and one in the retrograde direction. The relationship between state r and the RDLs A and R is a simple generalization of the PTL structural

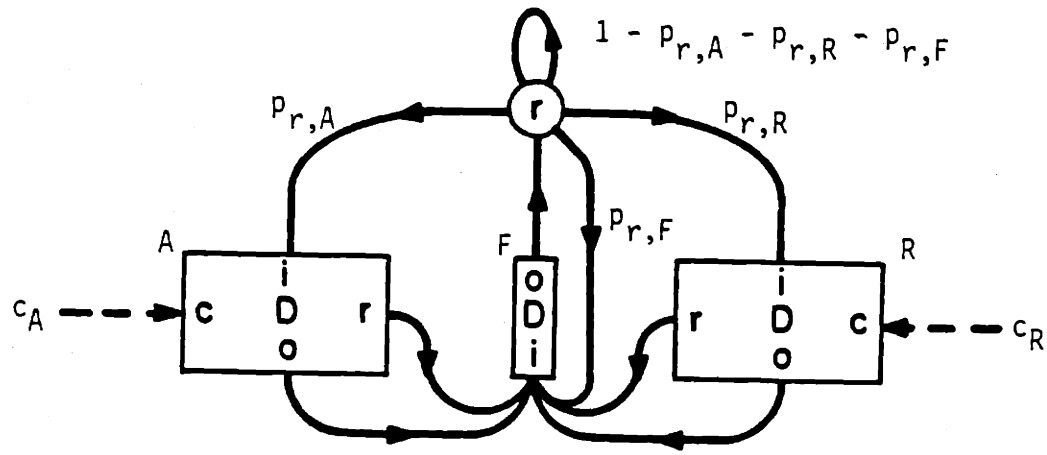


Figure 3.12 Bidirectional Refractory Transmission Line.

element. The third delay line (F), a DL, corresponds to the refractory period. A nonresettable delay line is used because, at the level of physiological detail that we are modeling, the duration of the refractory period is independent of all external events.

The state transition probabilities $p_{r,A}$, $p_{r,R}$, and $p_{r,F}$ and the RDL state transition probabilities (controlled exclusively through c_A and c_R) are the only probabilities that depend on the states of other submodels, that is, on the interactions impinging on the BDRTL. We now describe the effects of these impinging interactions. In the absence of external excitations, $p_{r,A} = p_{r,R} = p_{r,F} = 0$ and $c_A = c_R = \bar{R}$. If the BDRTL is excited from the antegrade direction but not simultaneously from the retrograde direction, then $p_{r,A} = 1$, $p_{r,R} = p_{r,F} = 0$, $c_A = \bar{R}$, and $c_R = R$. For the reverse case (i.e. excited from the retrograde direction but not simultaneously from the antegrade direction), the

values are $p_{r,R} = 1$, $p_{r,A} = p_{r,F} = 0$, $c_R = \bar{R}$, and $c_A = R$. Finally, if the BDRTL is simultaneously excited from both the antegrade and retrograde directions, then $p_{r,F} = 1$, $p_{r,A} = p_{r,R} = 0$, and $c_A = c_R = R$.

Depending on what anatomic substructure the BDRTL models, it may or may not contain transitions which generate a non-zero response in the electromagnetic model. If the BDRTL does contain such transitions, then there are three basic situations:

(1) Antegrade conduction without a reset (e.g. a normally conducted P wave from an atrial submodel).

(2) Retrograde conduction without a reset (e.g. a retrograde P wave from an atrial submodel).

(3) Reset antegrade or retrograde conduction which generates a fusion depolarization (e.g. a fusion P wave from an atrial submodel due to joint SA-nodal and retrograde-AV-junction depolarizations).

Though it is not the only possible choice, we have always used the transitions described in the following to initiate the various waveforms. The transitions r to i_A and r to i_R generate the nonreset antegrade-conduction and nonreset retrograde-conduction electromagnetic model responses respectively. In this case we are committed to generating the nonreset response before we know whether the conduction will actually be nonreset. Thus, if it is reset, we must have the transitions from the states internal to the RDL to the reset state (which is i_F) generate the desired fusion-depolarization response minus whatever part of the nonreset response that we wish to delete. This deletion can only be done exactly if two conditions hold:

(1) The RDL must have the tapped delay-line topology. Only in this case can the resetting transition represent the time interval since the delay line was entered and thus how far the resulting electromagnetic-model signature has progressed.

(2) The Gaussian signatures must have zero covariance and therefore actually be deterministic.

The fusion-depolarization response can be made to depend on the relative fractions of the substructure depolarized by each of the two depolarization waves. The relative fraction is reflected in the relative timing between the two excitations, which in turn is represented by the identity of the state occupied in the RDL when the RDL is reset.

The BDRTL models presented in the following section are less detailed models than those in the above discussion. Specifically, we have not separately included fusion depolarizations (from resetting the delay lines) in the electromagnetic model. Instead, we have taken the fusion-depolarization response to be identical with the response from whichever was the earlier of the two depolarization waves. With this simplification, the disadvantages of the approach discussed previously vanish, and we have used this approach in the models described in Sections 3.7 and 3.8.

In this section, we have discussed five structural elements, the building blocks of our submodels. In the following section (Section 3.7), we give four examples of detailed ECG models constructed from submodels which are in turn made up of these structural elements. Then, in Section 3.8, we present a fifth example which includes simulations.

The simulations indicate that our modeling approach can realistically capture complex cardiac behavior.

3.7. Examples of ECG Models

In this section we present models for four different conditions:

(1) Normal Rhythm,

(2) Normal Rhythm with Ectopic Focus PVCs,

(3) Ventricular Oscillation, and

(4) Wolff-Parkinson-White Syndrome.

The models described in this section are specified at the level of the structural elements described in Section 3.6. The details of the Markov chains and signature morphologies are not specified. Rather our objective here is to demonstrate the flexibility of our approach and its ability to include anatomic and physiological detail.

A simple, graphical, abstract notation for classes of models is helpful in describing both the models and (later) the estimators. This paragraph describes such a notation by example. Figure 3.13 describes a class of models, each of which has exactly four submodels. The boxes labeled C0, ..., C3 denote these submodels. The directed lines between boxes indicate the existence of an interaction in the indicated direction. Thus, for example, submodel C0 initiates an interaction with submodel C1. The number of values which the interaction can take on is not specified. The wavy lines terminating in S0, ..., S3 indicate that the submodel of the originating box contains one or more transitions

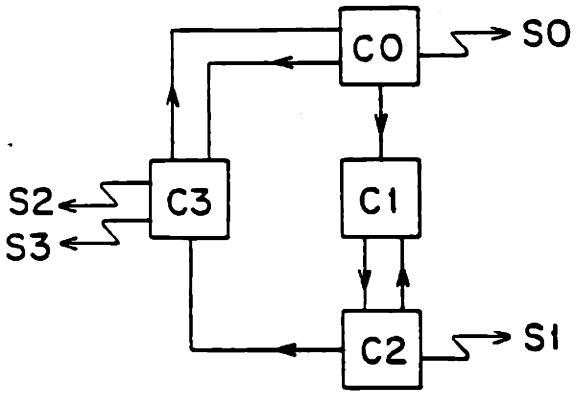


Figure 3.13 An Illustration of the Block Diagram Description of a Class of Models.

which initiate a signature whose name is the label at the end of the arrow.

3.7.1. Normal Rhythm

The clinical rhythm modeled here is a prototypical normally conducted rhythm. A block diagram of the model and a listing of the inter-submodel interactions is given in Figure 3.14.[2] As seen in Figure 3.14, we have divided the heart into four anatomic substructures--SA node, atria, AV junction, and ventricles--each of which is modeled by a separate submodel.

Qualitatively, the model behaves in the following manner. The SA-nodal submodel initiates a depolarization wave. This is the only way in which a depolarization can be initiated. The depolarization then propagates antegrade through the atrial submodel, producing the P wave; the

[2] In each subsection of Section 3.7, all figures appear at the end of the subsection.

AV-junctional submodel, which makes a zero contribution to the ECG; and finally the ventricular submodel, producing a QRS-T complex.

Because only antegrade conduction is included in the model, several aspects of the model should be noted. Specifically:

(1) submodel 0 is not resettable

(2) submodels 1 and 2, which would be BDRTLs if retrograde conduction were included, are instead simple arrangements of delay lines.

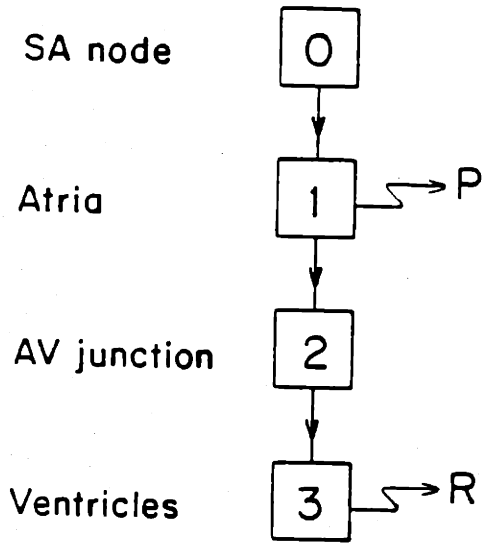
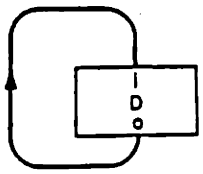
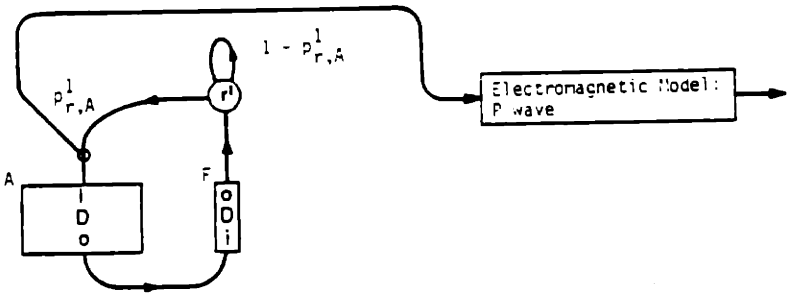


Figure 3.14 A Model For Normal Rhythm.
Part (a): Submodel Structure.

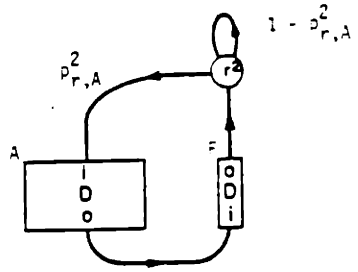
Submodel 0: SA-nodal Substructure



Submodel 1: Atrial Substructure



Submodel 2: AV-junctional Substructure



Submodel 3: Ventricular Substructure

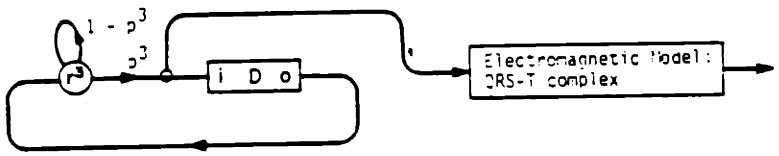


Figure 3.14 Continued.
Part (b): Block Diagram.

Submodel 0:

Submodel for the SA-nodal Substructure

This submodel is autonomous.

Submodel 1:

Submodel for the Atrial Substructure

$$p^1_{r,A} = \begin{cases} 1 \text{ if } x^0 \in \{o^0\} \\ 0 \text{ otherwise} \end{cases}$$

Submodel 2:

Submodel for the AV-junctional Substructure

$$p^2_{r,A} = \begin{cases} 1 \text{ if } x^1 \in \{o^1_A\} \\ 0 \text{ otherwise} \end{cases}$$

Submodel 3:

Submodel for the Ventricular Substructure

$$p^3 = \begin{cases} 1 \text{ if } x^2 \in \{o^2_A\} \\ 0 \text{ otherwise} \end{cases}$$

Figure 3.14 Continued.
Part (c): Intersubmodel Interactions.

3.7.2. Normal Rhythm with Ectopic Focus PVCs

The clinical rhythm modeled here is one in which the basic pattern is normal, but which contains superimposed occasional PVCs. The PVCs need not have constant coupling intervals to the prior QRS complex. Because of the non-constant coupling intervals, these PVCs are modeled as arising from an ectopic center rather than through a reentrant mechanism. In this model, all possible levels of retrograde conduction are modeled.

In order to develop a model for this arrhythmia, we have modified the normal-rhythm model in two ways. First, we have modified the part of the normal-rhythm model which corresponds to the part of the heart which exhibits the abnormal physiology. Therefore we replaced the ventricular submodel by a new ventricular submodel and an ectopic ventricular pacemaker submodel. Second, we have modified, as required, the remaining parts of the normal-rhythm model so that they can interact with the part modified in the first step. The primary purpose of these modifications is to allow retrograde conduction and resetting of the SA node.

A block diagram of the model and a listing of the inter-submodel interactions is given in Figure 3.15. As seen in Figure 3.15, we have divided the heart into five anatomic substructures--SA node, atria, AV junction, ventricles, and ectopic ventricular pacemaker--each of which is modeled by a separate submodel.

Qualitatively, the model's behavior is based on the competition for control of the heart between the two autorhythmic submodels--the SA-

nodal and the ectopic-ventricular-pacemaker submodels. Specifically, if the depolarization originating in the SA-nodal submodel propagates through the model and reaches the ventricular submodel, then a normal beat results and the ectopic-ventricular-pacemaker submodel is reset. On the other hand, if the depolarization originating in the ectopic-ventricular-pacemaker submodel propagates through the retrograde pathway of the ventricular submodel, then a PVC results. This depolarization may or may not be able to propagate all the way retrograde and reset the SA-nodal submodel, due to the possibility of collision with antegrade depolarization waves initiated by the SA-nodal submodel. Thus a very complex pattern is set up that depends on the transit-time pmfs of all the delay lines in the model.

The ventricular and ectopic-ventricular-pacemaker submodels require comment. First, note the symmetry between the SA-nodal and atrial pair of submodels and the ventricular and ectopic-ventricular-pacemaker pair of submodels. By reusing large portions of models in this manner we develop greater insight into the behavior of specific arrangements and thereby develop greater understanding of the functioning of signal-processing algorithms based on such models. Also, the development of a "library" of large fragments of models with well-understood behavior simplifies the development of models for new rhythms.

Second, note that a ventricular ectopic pacemaker can have a wide variety of locations in the ventricles. Its location will affect how it interacts with the ventricles and, through the ventricles, with the rest of the heart. That is, the location controls the depolarization-wave trajectory which in turn controls

(1) the morphology of the resulting PVC,

(2) the possibility of reentrant depolarization waves within the ventricles (a possibility not modeled here), and

(3) the timing of the interactions with the remainder (i.e. nonventricular part) of the heart.

In our models we can essentially control (1) and (3) independently: (1) through choosing the PVC signature in the electromagnetic model and (3) through choosing the structure and parameters of the physiological model. The dominant way in which the location of the focus affects the observable ECG is through the ECG morphology. Therefore, a model in which the timing is independent of, but the morphology is dependent on, the focus location may prove to be a sufficiently good approximation. Note that making the morphology dependent on focus location implies adapting the morphology model to individual subjects, which is something that a practical system must eventually deal with. Thus we have chosen the structure presented in Figure 3.15 because it is simple, plausible, has arisen previously, and, in our opinion, should prove to be adequate for our signal-processing purposes.

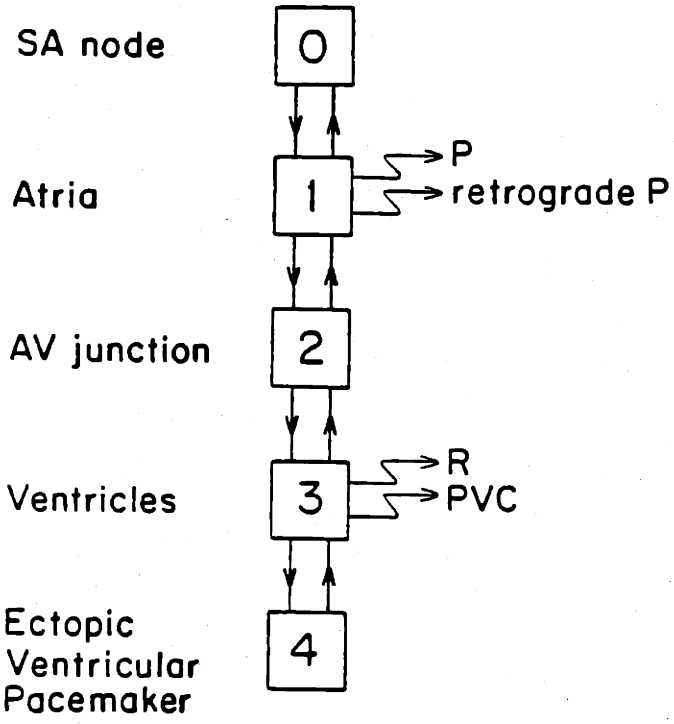
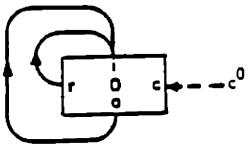
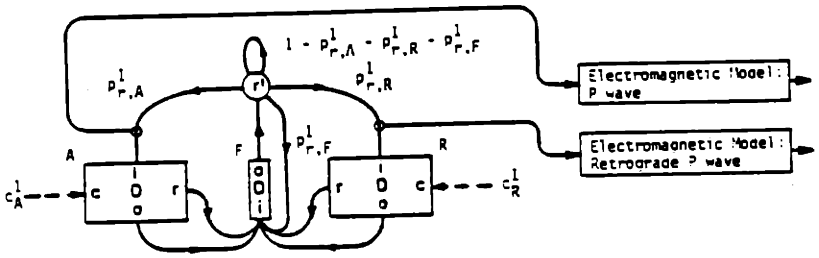


Figure 3.15 A Model For Normal Rhythm With PVCs Generated By An Ectopic Focus.
Part (a): Submodel Structure.

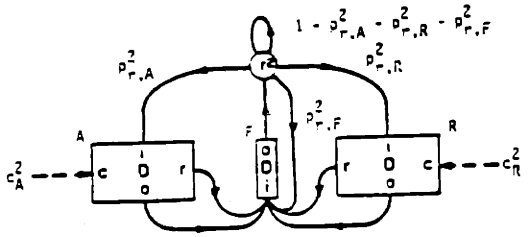
Submodel 0: SA-nodal Substructure



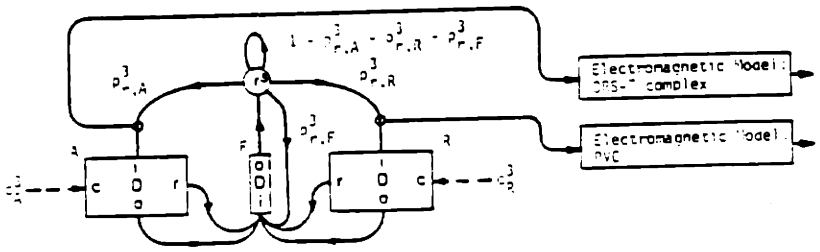
Submodel 1: Atrial Substructure



Submodel 2: AV-junctional Substructure



Submodel 3: Ventricular Substructure



Submodel 4: Ectopic-ventricular-pacemaker Substructure

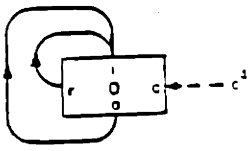


Figure 3.15 Continued.
Part (b): Block Diagram.

Submodel 0:

Submodel for the SA-nodal Substructure

$$c^0 = \begin{cases} \bar{R} \text{ if } x^1 \in \{o_R^1\} \\ R \text{ otherwise} \end{cases}$$

Submodel 1:

Submodel for the Atrial Substructure

$$p^1_{r,A} = \begin{cases} 1 \text{ if } x^0 \in \{o^0\} \text{ and } x^2 \in \{o_R^2\} \\ 0 \text{ otherwise} \end{cases}$$

$$p^1_{r,R} = \begin{cases} 1 \text{ if } x^0 \in \{o^0\} \text{ and } x^2 \in \{o_R^2\} \\ 0 \text{ otherwise} \end{cases}$$

$$p^1_{r,F} = \begin{cases} 1 \text{ if } x^0 \in \{o^0\} \text{ and } x^2 \in \{o_R^2\} \\ 0 \text{ otherwise} \end{cases}$$

$$c^1_A = \begin{cases} \bar{R} \text{ if } x^2 \in \{o_R^2\} \\ R \text{ otherwise} \end{cases}$$

Figure 3.15 Continued.
Part (c): Intersubmodel Interactions.

$$c_R^1 = \begin{cases} \bar{R} \text{ if } x^0 \notin \{o^0\} \\ R \text{ otherwise} \end{cases}$$

Submodel 2:

Submodel for the AV-junctional Substructure

$$p_{r,A}^2 = \begin{cases} 1 \text{ if } x^1 \in \{o_A^1\} \text{ and } x^3 \notin \{o_R^3\} \\ 0 \text{ otherwise} \end{cases}$$

$$p_{r,R}^2 = \begin{cases} 1 \text{ if } x^1 \notin \{o_A^1\} \text{ and } x^3 \in \{o_R^3\} \\ 0 \text{ otherwise} \end{cases}$$

$$p_{r,F}^2 = \begin{cases} 1 \text{ if } x^1 \in \{o_A^1\} \text{ and } x^3 \in \{o_R^3\} \\ 0 \text{ otherwise} \end{cases}$$

$$c_A^2 = \begin{cases} \bar{R} \text{ if } x^3 \notin \{o_R^3\} \\ R \text{ otherwise} \end{cases}$$

$$c_R^2 = \begin{cases} \bar{R} \text{ if } x^1 \notin \{o_A^1\} \\ R \text{ otherwise} \end{cases}$$

Figure 3.15 Continued.
Part (c) Continued.

Submodel 3:

Submodel for the Ventricular Substructure

$$p^3_{r,A} = \begin{cases} 1 \text{ if } x^2 \in \{o^2_A\} \text{ and } x^4 \in \{o^4\} \\ 0 \text{ otherwise} \end{cases}$$

$$p^3_{r,R} = \begin{cases} 1 \text{ if } x^2 \in \{o^2_A\} \text{ and } x^4 \in \{o^4\} \\ 0 \text{ otherwise} \end{cases}$$

$$p^3_{r,F} = \begin{cases} 1 \text{ if } x^2 \in \{o^2_A\} \text{ and } x^4 \in \{o^4\} \\ 0 \text{ otherwise} \end{cases}$$

$$c^3_A = \begin{cases} \bar{R} \text{ if } x^4 \in \{o^4\} \\ R \text{ otherwise} \end{cases}$$

$$c^3_R = \begin{cases} \bar{R} \text{ if } x^2 \in \{o^2_A\} \\ R \text{ otherwise} \end{cases}$$

Figure 3.15 Continued.
Part (c) Continued.

Submodel 4:

Submodel For The Ectopic-ventricular-pacemaker Substructure

$$c^4 = \begin{cases} \bar{R} \text{ if } x^3 \in \{o_A^3\} \\ R \text{ otherwise} \end{cases}$$

Figure 3.15 Continued.
Part (c) Continued.

3.7.3. Ventricular Oscillation

The clinical rhythm modeled here is a rhythm in which a normal beat is followed by a variable number of premature ventricular contractions. For any fixed set of parameters (i.e. delay-line transit-time pmfs and other probabilities), the number of PVCs which follow each normal beat will not vary significantly and the coupling interval between the normal beat and the first PVC and the coupling interval between the PVCs will both be approximately constant. Because of the constant coupling intervals, the premature beats are modeled through a ventricular reentrant-pathway mechanism. In this model, all possible levels of retrograde conduction are modeled.

In this rhythm, the abnormal physiology is the existence of a ventricular reentrant pathway. Therefore our modifications to the normal-rhythm model discussed previously are restricted to adding a reentrant ventricular pathway, implemented by a new ventricular submodel and the ventricular-reentrant-pathway submodel, and modifying the neighboring submodels so that they can interact appropriately.

A block diagram and a listing of the inter-submodel interactions is given in Figure 3.16. As seen in Figure 3.16, we have divided the heart into five anatomic substructures--SA node, atria, AV junction, ventricles, and reentrant ventricular pathway--each of which is modeled by a separate submodel. We have based the reentrant ventricular pathway on the unidirectional-conduction-block mechanism.

Qualitatively, the model operates in the following manner. The SA node initiates a depolarization wave which propagates antegrade through

the atria and the AV junction in the normal manner. On reaching the ventricular substructures, it is successful in depolarizing the ventricular substructure but is unsuccessful in depolarizing the reentrant-ventricular-pathway substructure due to the unidirectional conduction block. However, as the ventricular substructure depolarizes, eventually the reentrant-ventricular-pathway substructure is successfully excited in the retrograde direction. Conduction in the reentrant-ventricular-pathway substructure is very slow and while it occurs, the ventricular substructure repolarizes. Then the depolarization wave exiting the reentrant-ventricular-pathway substructure is able to reexcite the ventricular substructure in the retrograde direction. This produces a PVC and, depending on how far retrograde the depolarization can propagate, may produce other effects. Depending on the duration of the reentrant-ventricular-pathway-substructure refractory time, the reentrant-ventricular-pathway substructure can be reexcited by the retrograde ventricular depolarization leading to the possibility of multiple cycles.

As in the previous model (normal rhythm with ectopic focus PVCs), we have great latitude in choosing the submodel structures and the inter-submodel interactions required to create the reentrant pathway. We have chosen a simple but physiologically reasonable arrangement that can generate a large variety of ventricular oscillation rhythms.

The only submodel that requires explanation is the reentrant-ventricular-pathway submodel. State r^4 is the usual resting state modeling the polarized state awaiting an excitation. The excitation, which controls p^4, can come from either the retrograde output or the antegrade output of the ventricular submodel. This models the

assumption that the pathway has a unidirectional conduction block but, as a depolarization wave reaches all the ventricular tissue, will eventually be excited. Therefore, every time the ventricular submodel is depolarized, whether from the antegrade or from the retrograde direction, the ventricular submodel will try to excite the reentrant-ventricular-pathway submodel. The delay line labeled "C" models the conduction delay through the reentrant pathway. Thus, when the state occupies the final state of the delay line (state o_C^4), the reentrant-ventricular-pathway submodel attempts to reexcite the ventricular submodel. Finally, the delay line labeled "R" models the refractory-period duration.

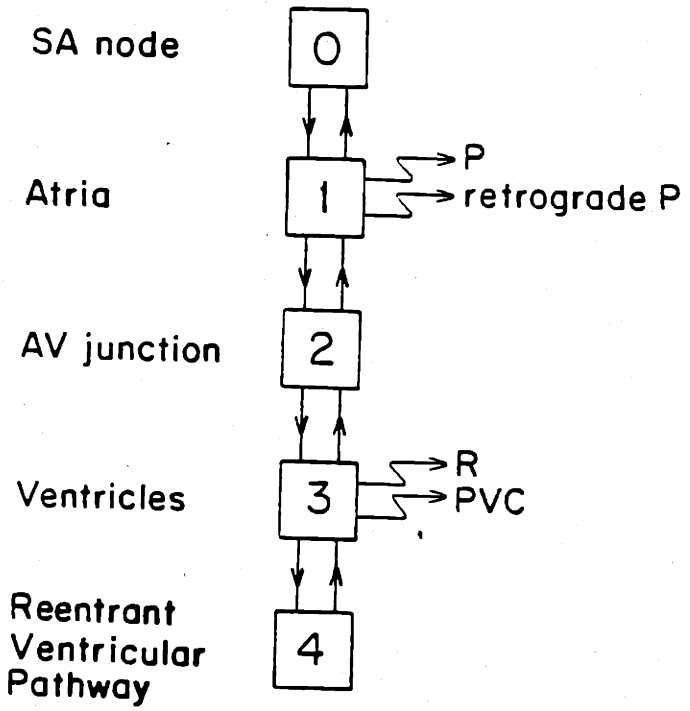
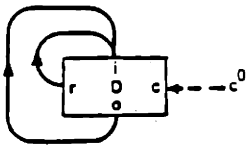
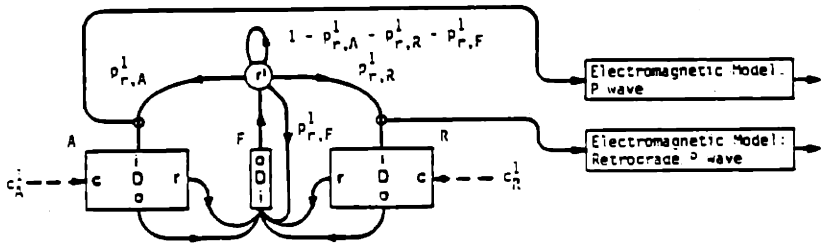


Figure 3.16 A Model For Ventricular Oscillation.
Part (a): Submodel Structure.

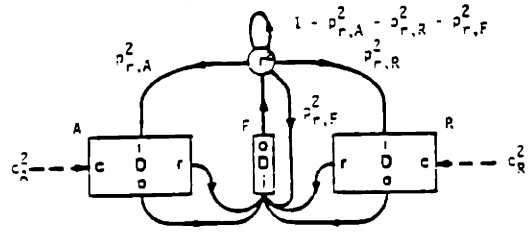
Submodel 0: SA-nodal Substructure



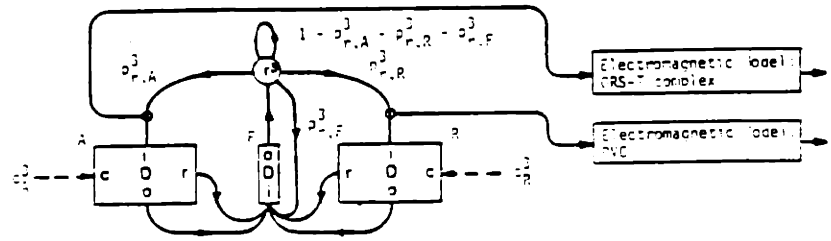
Submodel 1: Atrial Substructure



Submodel 2: AV-junctional Substructure



Submodel 3: Ventricular Substructure



Submodel 4: Reentrant-ventricular-pathway Substructure

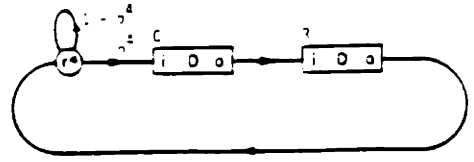


Figure 3.16 Continued.
Part (b): Block Diagram.

Submodel 0:

Submodel for the SA-nodal Substructure

$$c^0 = \begin{cases} \bar{R} \text{ if } x^1 \notin \{o_R^1\} \\ R \text{ otherwise} \end{cases}$$

Submodel 1:

Submodel for the Atrial Substructure

$$p^1_{r,A} = \begin{cases} 1 \text{ if } x^0 \in \{o^0\} \text{ and } x^2 \notin \{o_R^2\} \\ 0 \text{ otherwise} \end{cases}$$

$$p^1_{r,R} = \begin{cases} 1 \text{ if } x^0 \notin \{o^0\} \text{ and } x^2 \in \{o_R^2\} \\ 0 \text{ otherwise} \end{cases}$$

$$p^1_{r,F} = \begin{cases} 1 \text{ if } x^0 \in \{o^0\} \text{ and } x^2 \in \{o_R^2\} \\ 0 \text{ otherwise} \end{cases}$$

$$c^1_A = \begin{cases} \bar{R} \text{ if } x^2 \notin \{o_R^2\} \\ R \text{ otherwise} \end{cases}$$

Figure 3.16 Continued.
Part (c): Intersubmodel Interactions.

$$c_R^1 = \begin{cases} \bar{R} \text{ if } x^0 \in \{o^0\} \\ R \text{ otherwise} \end{cases}$$

Submodel 2:

Submodel for the AV-junctional Substructure

$$p_{r,A}^2 = \begin{cases} 1 \text{ if } x^1 \in \{o_A^1\} \text{ and } x^3 \in \{o_R^3\} \\ 0 \text{ otherwise} \end{cases}$$

$$p_{r,R}^2 = \begin{cases} 1 \text{ if } x^1 \in \{o_A^1\} \text{ and } x^3 \in \{o_R^3\} \\ 0 \text{ otherwise} \end{cases}$$

$$p_{r,F}^2 = \begin{cases} 1 \text{ if } x^1 \in \{o_A^1\} \text{ and } x^3 \in \{o_R^3\} \\ 0 \text{ otherwise} \end{cases}$$

$$c_A^2 = \begin{cases} \bar{R} \text{ if } x^3 \in \{o_R^3\} \\ R \text{ otherwise} \end{cases}$$

$$c_R^2 = \begin{cases} \bar{R} \text{ if } x^1 \in \{o_A^1\} \\ R \text{ otherwise} \end{cases}$$

Figure 3.16 Continued.
Part (c) Continued.

Submodel 3:

Submodel for the Ventricular Substructure

$$p^3_{r,A}=\begin{cases}1 \text{ if } x^2\in\{o^2_A\} \text{ and } x^4\in\{o^4_C\}\\ 0 \text{ otherwise}\end{cases}$$

$$p^3_{r,R}=\begin{cases}1 \text{ if } x^2\in\{o^2_A\} \text{ and } x^4\in\{o^4_C\}\\ 0 \text{ otherwise}\end{cases}$$

$$p^3_{r,F}=\begin{cases}1 \text{ if } x^2\in\{o^2_A\} \text{ and } x^4\in\{o^4_C\}\\ 0 \text{ otherwise}\end{cases}$$

$$c^3_A=\begin{cases}\bar{R} \text{ if } x^4\in\{o^4_C\}\\ R \text{ otherwise}\end{cases}$$

$$c^3_R=\begin{cases}\bar{R} \text{ if } x^2\in\{o^2_A\}\\ R \text{ otherwise}\end{cases}$$

Figure 3.16 Continued.
Part (c) Continued.

Submodel 4:

Submodel For The Ectopic-ventricular-pacemaker Substructure

$$p^4 = \begin{cases} 1 \text{ if } x^3 \in \{o_A^3, o_R^3\} \\ 0 \text{ otherwise} \end{cases}$$

Figure 3.16 Continued.
Part (c) Continued.

3.7.4. Wolff-Parkinson-White Syndrome

The model discussed in this subsection describes the Wolff-Parkinson-White Syndrome in terms of the ectopic atrial-ventricular conduction pathways which are the physiological cause of the syndrome. The primary modification that we have made to our normal-rhythm model in order to model this syndrome is the introduction of a submodel for the bundle of Kent. Furthermore, in order to easily model

(1) ventricles with two excitation sources (and which therefore have the possibility of undergoing a fusion depolarization) and

(2) a reentrant pathway made up of (primarily) the bundle of Kent and the AV junction,

we have subdivided the ventricular anatomic substructure into two anatomic substructures in the following manner.

(1) The portion of the ventricles responsible for the delta wave (the portion which lies at the distal terminus of the bundle of Kent) together with the portion of the ventricles lying between the distal termini of the bundle of Kent and the AV junction are lumped into one anatomic substructure called the delta substructure. (Note that the two volumes described above are not disjoint).

(2) The remainder of the ventricles excluded from (1) is a single anatomic substructure called the ventricular substructure.

Finally, we have incorporated retrograde conduction into the SA-nodal, atrial, and AV-junctional submodels.

A block diagram and a listing of the inter-submodel interactions are given in Figure 3.17. As seen in Figure 3.17, we have divided the heart into six anatomic substructures--SA node, atria, AV junction, bundle of Kent, delta, and ventricles--each of which is modeled by a separate submodel. Qualitatively, the model operates in the following manner. The SA node initiates a depolarization which propagates through the atria in the normal manner. It then enters both the AV junction and the bundle of Kent. However, because of the AV-nodal delay, transmission through the bundle of Kent is much faster than through the AV junction. Thus the bundle of Kent (rather than the AV junction) first excites the ventricles and this excitation occurs in the delta substructure. This generates the short P-R interval and the delta wave. Finally, due to the eventual arrival of the depolarization from the AV junction, the remainder of the ventricles (ventricular submodel) is depolarized through the normal pathway. Thus the entire ventricles (delta plus ventricular submodels) are excited from two sources (bundle of Kent and AV junction) to generate a fusion depolarization. However, the latter excitation (AV junction) becomes dominant because it feeds the preferred high-speed ventricular depolarization pathways (e.g. the Purkinje system). As seen in Figure 3.17, there is also the possibility of reentrant depolarizations, propagating in either direction, through a pathway made up from the bundle of Kent, delta, and AV-junctional submodels and also the possibility of bizarre QRS-T morphology due to depolarization of the ventricles (delta plus ventricular submodel) exclusively through the bundle of Kent. The actual occurrence of these later two phenomena depends on the details of the pmfs for the various delay-line transit times.

The delta and the ventricular submodels require further comment. The delta submodel is a BDRTL with an arbitrary choice for the antegrade/retrograde designations. We have assigned antegrade to conduction from the distal end of the bundle of Kent toward the distal end of the AV junction. From a design point of view, this substructure performs two tasks:

(1) It plays a major role in modeling the fusion depolarization of the entire ventricles (delta plus ventricular submodels).

(2) It closes the bundle-of-Kent-AV-junctional reentrant-pathway loop in both conduction directions.

The antegrade excitation source for the delta submodel is the antegrade output of the bundle-of-Kent submodel. This models the conduction pattern of a standard Wolff-Parkinson-White Syndrome beat and a link in the abnormal reentrant pathway in which conduction is antegrade in the bundle of Kent and retrograde in the AV junction. A successful excitation results in the transition r^4 to i_A^4. Therefore this transition generates the delta wave in the electromagnetic model. The depolarization described here eventually reaches all the ventricular tissue included in this submodel. At that point, the delta-submodel state occupies state o_A^4. Then the depolarization wave attempts to continue into the remainder of the ventricles and retrograde into the AV junction. If it excites the remainder of the ventricles (modeled by the ventricular submodel), a Q'RS-T with bizarre morphology will be formed due to the use of abnormal depolarization pathways. (Q'RS-T instead of QRS-T because this is the ECG fluctuation due to only a part of the ventricular myocardium--that part modeled by the ventricular submodel which does not

include the part modeled by the delta submodel). Therefore, when the delta-submodel state occupies state o_A^4, it attempts to excite, via the interaction mechanism, the retrograde path of the AV-junctional submodel and the bizarre Q'RS-T pathway (see the following) of the ventricular submodel.

The retrograde excitation source for the delta submodel is the antegrade output of the AV-junctional submodel. This models a link in the abnormal reentrant pathway in which conduction is antegrade in the AV junction and retrograde in the bundle of Kent. A successful excitation in the retrograde direction results in the transition r^4 to i_R^4. Therefore this transition generates the "retrograde delta wave" in the electromagnetic model. Note that the depolarization modeled here is a part of the normal AV-junction-initiated ventricular depolarization.

The present ventricular submodel is a slight generalization of the ventricular submodel of the normal-rhythm model. Instead of having only one transition which generates a non-zero response from the electromagnetic model, the present submodel has two. One transition generates the normal Q'RS-T morphology. The transition producing this wave (r^5 to n^5) is in what we call the <u>normal Q'RS-T pathway</u> (r^5 to n^5 to i^5 to o^5 to r^5). This pathway models excitation of the ventricles through the normal conduction pathway starting from the distal end of the AV junction. The second transition which generates a non-zero response from the electromagnetic model generates a bizarre Q'RS-T morphology. The transition producing this triplet (r^5 to b^5) lies in what we call the <u>bizarre Q'RS-T pathway</u> (r^5 to b^5 to i^5 to o^5 to r^5). This pathway models the excitation of the ventricles through an abnormal pathway starting at the

distal terminus of the bundle of Kent, traversing the tissue modeled by the delta submodel, and then continuing into the tissue modeled by the ventricular submodel. The normal Q'RS-T pathway is taken if the excitation comes from the AV-junctional submodel or simultaneously from both the AV-junctional and delta submodels. The bizarre Q'RS-T pathway is taken if the excitation comes exclusively from the delta submodel.

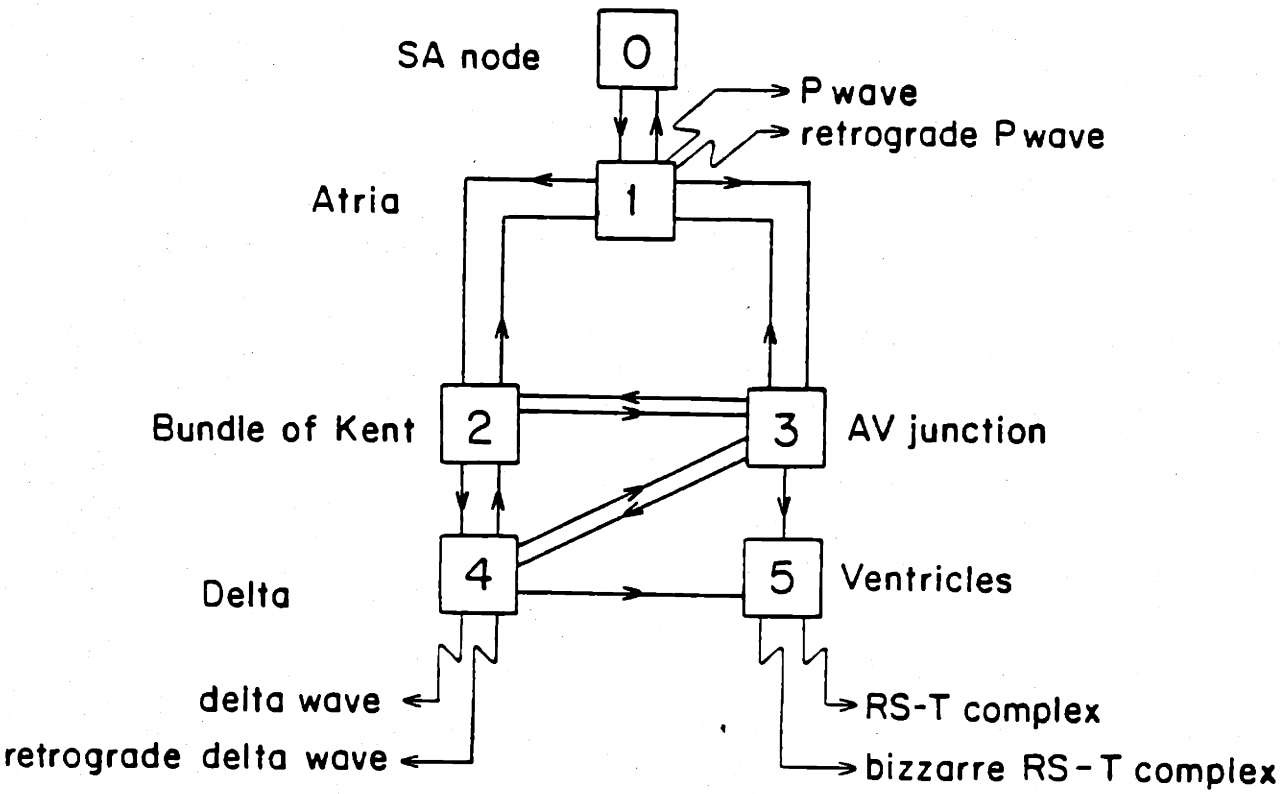
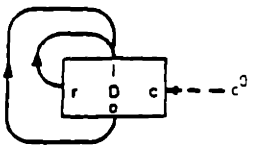
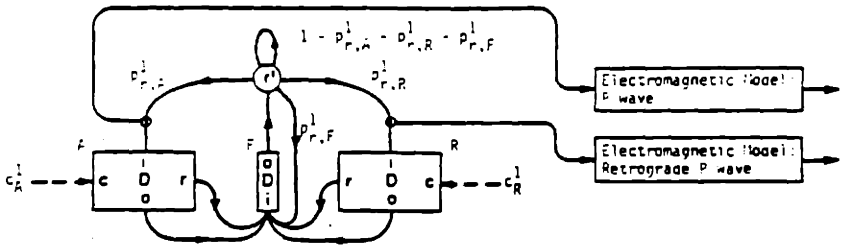


Figure 3.17 A Model For The Wolff-Parkinson-White Syndrome. Part (a): Submodel Structure.

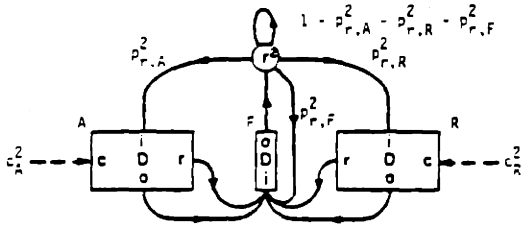
Submodel 0: SA-nodal Substructure



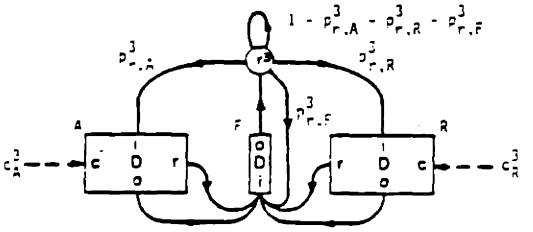
Submodel 1: Atrial Substructure



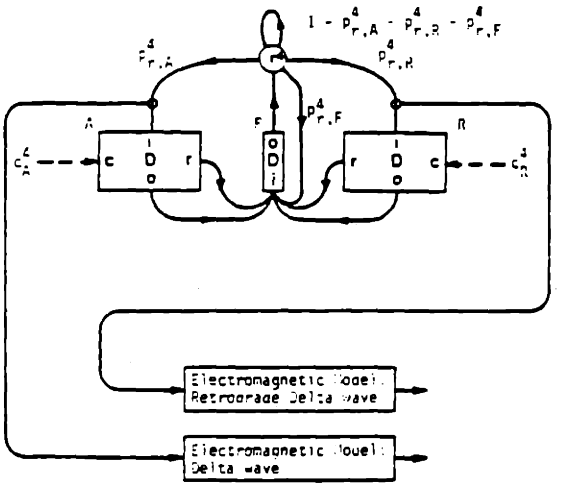
Submodel 2: Bundle-of-Kent Substructure



Submodel 3: AV-junctional Substructure



Submodel 4: Delta Substructure



Submodel 5: Ventricular Substructure

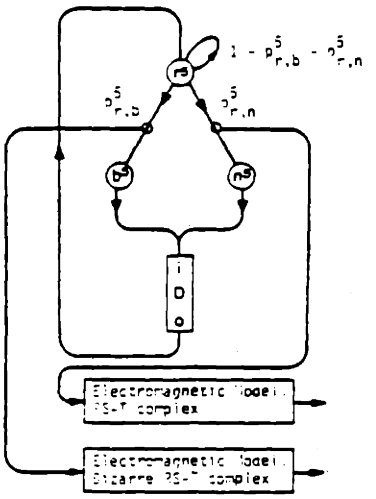


Figure 3.17 Continued.
Part (b): Block Diagram.

Submodel 0:

Submodel for the SA-nodal Substructure

$$c^0 = \begin{cases} \bar{R} \text{ if } x^1 \in \{o_R^1\} \\ R, \text{ otherwise} \end{cases}$$

Submodel 1:

Submodel for the Atrial Substructure

$$p^1_{r,A} = \begin{cases} 1 \text{ if } x^0 \in \{o^0\} \text{ and } (x^2 \in \{o_R^2\} \text{ and } x^3 \in \{o_R^3\}) \\ 0 \text{ otherwise} \end{cases}$$

$$p^1_{r,R} = \begin{cases} 1 \text{ if } x^0 \in \{o^0\} \text{ and } (x^2 \in \{o_R^2\} \text{ or } x^3 \in \{o_R^3\}) \\ 0 \text{ otherwise} \end{cases}$$

$$p^1_{r,F} = \begin{cases} 1 \text{ if } x^0 \in \{o^0\} \text{ and } (x^2 \in \{o_R^2\} \text{ or } x^3 \in \{o_R^3\}) \\ 0 \text{ otherwise} \end{cases}$$

$$c^1_A = \begin{cases} \bar{R} \text{ if } x^2 \in \{o_R^2\} \text{ and } x^3 \in \{o_R^3\} \\ R \text{ otherwise} \end{cases}$$

Figure 3.17 Continued.
Part (c): Intersubmodel Interactions.

$$c_R^1 = \begin{cases} \bar{R} \text{ if } x^0 \in \{o^0\} \\ R \text{ otherwise} \end{cases}$$

Submodel 2:

Submodel for the Bundle-of-Kent Substructure

$$p_{r,A}^2 = \begin{cases} 1 \text{ if } (x^1 \in \{o_A^1\} \text{ or } x^3 \in \{o_R^3\}) \text{ and } x^4 \in \{o_R^4\} \\ 0 \text{ otherwise} \end{cases}$$

$$p_{r,R}^2 = \begin{cases} 1 \text{ if } (x^1 \in \{o_A^1\} \text{ and } x^3 \in \{o_R^3\}) \text{ and } x^4 \in \{o_R^4\} \\ 0 \text{ otherwise} \end{cases}$$

$$p_{r,F}^2 = \begin{cases} 1 \text{ if } (x^1 \in \{o_A^1\} \text{ or } x^3 \in \{o_R^3\}) \text{ and } x^4 \in \{o_R^4\} \\ 0 \text{ otherwise} \end{cases}$$

$$c_A^2 = \begin{cases} \bar{R} \text{ if } x^4 \in \{o_R^4\} \\ R \text{ otherwise} \end{cases}$$

$$c_R^2 = \begin{cases} \bar{R} \text{ if } x^1 \in \{o_A^1\} \text{ and } x^3 \in \{o_R^3\} \\ R \text{ otherwise} \end{cases}$$

Figure 3.17 Continued.
Part (c) Continued.

Submodel 3:

Submodel for the AV-junctional Substructure

$$p^3_{r,A}=\begin{cases}1 \text{ if } (x^1\in\{o^1_A\} \text{ or } x^2\in\{o^2_R\}) \text{ and } x^4\in\{o^4_A\}\\ 0 \text{ otherwise}\end{cases}$$

$$p^3_{r,R}=\begin{cases}1 \text{ if } (x^1\in\{o^1_A\} \text{ and } x^2\in\{o^2_R\}) \text{ and } x^4\in\{o^4_A\}\\ 0 \text{ otherwise}\end{cases}$$

$$p^3_{r,F}=\begin{cases}1 \text{ if } (x^1\in\{o^1_A\} \text{ or } x^2\in\{o^2_R\}) \text{ and } x^4\in\{o^4_A\}\\ 0 \text{ otherwise}\end{cases}$$

$$c^3_A=\begin{cases}\bar{R} \text{ if } x^4\in\{o^4_A\}\\ R \text{ otherwise}\end{cases}$$

$$c^3_R=\begin{cases}\bar{R} \text{ if } x^1\in\{o^1_A\} \text{ and } x^2\in\{o^2_R\}\\ R \text{ otherwise}\end{cases}$$

Figure 3.17 Continued.
Part (c) Continued.

Submodel 4:

Submodel For The Delta Substructure

$$p^4_{r,A} = \begin{cases} 1 \text{ if } x^2 \in \{o^2_A\} \text{ and } x^3 \in \{o^3_A\} \\ 0 \text{ otherwise} \end{cases}$$

$$p^4_{r,R} = \begin{cases} 1 \text{ if } x^2 \in \{o^2_A\} \text{ and } x^3 \in \{o^3_A\} \\ 0 \text{ otherwise} \end{cases}$$

$$p^4_{r,F} = \begin{cases} 1 \text{ if } x^2 \in \{o^2_A\} \text{ and } x^3 \in \{o^3_A\} \\ 0 \text{ otherwise} \end{cases}$$

$$c^4_A = \begin{cases} \bar{R} \text{ if } x^3 \in \{o^3_A\} \\ R \text{ otherwise} \end{cases}$$

$$c^4_R = \begin{cases} \bar{R} \text{ if } x^2 \in \{o^2_A\} \\ R \text{ otherwise} \end{cases}$$

Figure 3.17 Continued.
Part (c) Continued.

Submodel 5:

Submodel for the Ventricular Substructure

$$p^5_{r,n} = \begin{cases} 1 \text{ if } x^3 \in \{o^3_A\} \\ 0 \text{ otherwise} \end{cases}$$

$$p^5_{r,b} = \begin{cases} 1 \text{ if } x^4 \in \{o^4_A\} \text{ and } x^3 \notin \{o^3_A\} \\ 0 \text{ otherwise} \end{cases}$$

Figure 3.17 Continued.
Part (c) Continued.

3.7.5. Summary of Section 3.7

In this section, we have presented four examples of our ECG models. However, in none of these models did we specify the delay-line transit-time pmfs of the physiological model or the morphology of the electromagnetic model. In the following section, we present one additional example, a model for the Wenckebach rhythm, in which these quantities are specified and for which simulations are included.

3.8. An Example ECG Model--Wenckebach

In this section we present a model for the Wenckebach arrhythmia. This arrhythmia is also referred to as Mobitz Type I. In the ECG, Wenckebach is characterized by a multibeat cycle in which the P waves are repeated at constant intervals but the P-R interval grows until, in the final beat of the cycle, the R wave is dropped. Then the cycle begins again with the P-R interval reset to its initial small value. The increase in the P-R interval from beat to beat is usually greatest at the beginning of the multibeat cycle. The multibeat cycle is typically three or four beats long.

Physiologically, the cause of Wenckebach is a defective AV node. Specifically, the AV node is such that it has a long relative refractory period. At the beginning of the multibeat Wenckebach cycle, the AV node is at rest. The first excitation occurs and is transmitted to the ventricles and the AV node enters its refractory period. Because the refractory period is prolonged, the second excitation from the atria reaches the AV node during its relative refractory period. The impulse is still able to excite the AV node and through it the ventricles since

the effective refractory period is past. However, the early excitation of the AV node has two effects: the following P-R interval and the following refractory period are both prolonged. Thus the third excitation occurs even earlier in the relative refractory period. This lengthening of the P-R interval and refractory period continues until finally a depolarization wave attempts to excite the AV node during its effective refractory period and is not conducted at all. This leads to the dropped R wave. Because the AV node is not excited during the dropped beat, the occurrence of the dropped beat gives the AV node time to complete its refractory period. Therefore, the P-R interval and refractory period for the succeeding depolarization of the AV node are reset to their initial (i.e. small) values.

As we have emphasized previously, our approach to ECG modeling is to use models of realistic physiological mechanisms to induce the arrhythmic behavior of the complete model. Thus, for Wenckebach, abnormalities are introduced into the submodel for the AV node. The overall model is made up of three submodels: SA-nodal/Atrial (SAN/A), AV-nodal (AV), and ventricular (V) submodels. Since retrograde conduction does not play a role in Wenckebach, it is not included in the model. The resulting submodels and their interactions are shown in Figure 3.18.[3]

A more detailed block diagram of the model, in terms of the structural elements of Section 3.6, is shown in Figure 3.19.

The SAN/A submodel is simply the autorhythmic structural element

[3] All figures and tables for Section 3.8 appear at the end of the section.

described earlier. The SAN/A submodel attempts to excite the AV submodel whenever the state of the SAN/A submodel occupies state 0. In this model we are in effect assuming that the transmission of the depolarization wave across the atria is instantaneous. Therefore, the initiation of the P wave, which occurs on transition out of state 0, and the attempt to excite the AV node, which occurs on occupancy of state 0, are essentially simultaneous. Physically, this amounts to lumping the time to depolarize the atria and the AV-junctional delay time together and calling the resulting time the AV-junctional delay time.

The AV submodel must exhibit the complex behavior described above. Specifically, early excitation (that is, excitation during the relative refractory period) must lead to a prolonged AV-junctional delay time followed by a prolonged refractory period. The correspondence between parts of the submodel and physiological events is as follows.

(1) AV-junctional delay time: a transit time through one of the four DLs labeled AV_1, ..., AV_4. The transit-time pmf for AV_1 is biased toward smaller values than is the pmf for AV_2. Likewise for the pair AV_2 and AV_3 and the pair AV_3 and AV_4. Thus, in the sense of typical transit times, $AV_1 < AV_2 < AV_3 < AV_4$.

(2) Attempt to excite the ventricles: occupancy of any of the states labeled 1, 2, 3, or 4.

(3) Effective refractory period: the total time spent in the DLs labeled AR_1, AR_2, AR_3, and AR_4. Note that in some situations, discussed in more detail in the following, this total time will not have contributions from <u>all</u> of these DLs.

(4) Relative refractory period: The total time spent in the RDLs labeled RR_1, RR_2, and RR_3. Note that the RDLs are reset if and only if the SAN/A submodel attempts to excite the AV submodel. Note also that each of the three RDLs has a different reset state. This is discussed in more detail in the sequel.

(5) Resting state: occupancy of state 0. Note that state 0 is a trapping state except when the SAN/A submodel attempts to excite the AV submodel. In that case the state, if it is occupying state 0, transitions into the delay line AV_1 with probability one.

We now describe a Wenckebach cycle. Initially the AV node is at rest: $x^1=0$. When the AV submodel is excited, the state transitions into the AV_1 DL. The transit time for this DL is the AV-junctional delay. A transit time from the AV_1 DL is biased toward shorter values than a transit time from any of the other AV_i DLs. Therefore, as desired, this is the shortest possible AV-junctional delay. After the AV-junctional delay, the AV submodel attempts to excite the V submodel: $x^1=1$. Then the AV submodel enters the effective refractory period. Note that the effective refractory period contains a transit-time contribution only from the AR_4 DL and therefore the effective refractory period is short. Following the effective refractory period is the relative refractory period, the total time spent in the RR_1, RR_2, and RR_3 DLs. If the next excitation of the AV submodel is sufficiently delayed, the AV submodel's state will pass through the three RDLs labeled RR_1, RR_2, and RR_3 and reenter the resting state (state 0). However, that is not what usually occurs. Rather, the refractory period duration is such that the next excitation of the AV submodel generally occurs during the

relative refractory period. More specifically, because this first beat of the cycle had a short AV-junctional delay (used delay line AV_1) and a short effective refractory period (avoided delay lines AR_1, AR_2, and AR_3), the next excitation of the AV submodel generally occurs while the AV submodel's state is in the final RDL of the relative refractory period, namely RR_3. Therefore, the excitation of the AV submodel forces the AV submodel's state to transition into the AV_2 DL, leading to a somewhat longer AV-junctional delay than in the first beat of the cycle and subsequently to a somewhat longer effective refractory period (AR_3 and AR_4 delay lines).

In this stage of the cycle the state is typically in the RR_2 DL when the following excitation occurs. Therefore, the state is reset into the AV_3 DL. This leads to a still longer AV-junctional delay (AV_3 delay line) followed by a longer effective refractory period (AR_2, AR_3, and AR_4 delay lines). The state is typically in the RR_1 RDL when the following excitation occurs. Therefore, the state is reset into the AV_4 DL producing a long AV-junctional delay followed by a long effective refractory period (AR_1, AR_2, AR_3, and AR_4 delay lines). In this part of the cycle the state is typically still in one of the effective-refractory-period DLs when the following excitation occurs. Therefore this excitation has no effect on the AV submodel. Rather, the state of the AV submodel continues through the effective-refractory-period DLs, the relative-refractory-period RDLs, and finally traps in the resting state (state 0). It remains in the resting state until the succeeding excitation occurs. Therefore, the excitation which should have started beat five does not get passed on to the ventricles. That is, beat five is dropped. Finally, because the state of the AV submodel is able to

reach the resting state (state 0), the Wenckebach cycle is restarted.

The V submodel is simply a PTL (state 0 and delay line RT) whose output is connected to a DL (R_V) whose output is in turn connected to the input of the PTL. The correspondence between parts of the submodel and physiological events is as follows:

(1) Ventricular depolarization corresponds to the transit time through the RT delay line.

(2) Ventricular refractory period corresponds to the time spent in the R_V delay line.

(3) The resting state corresponds to occupancy of state 0. Note that state 0, which is the resting state of the PTL, is a trapping state except when the AV submodel attempts to excite the V submodel. In that case the state, if it is occupying state 0, transitions into the RT delay line with probability one.

The operation of this submodel is obvious. The only point of note is that the two delay lines have transit-time pmfs such that the state of the V submodel has almost always returned to the resting state before the succeeding attempted excitation by the AV submodel occurs. Therefore, dropped beats are due to events in the AV submodel and not in the V submodel.

The actual Markov chains and signatures are shown in Figure 3.20. They were chosen based on an nominal heart rate of 60 beats per minute with a Markov chain cycle period of 1/25 second and a signature sampling period of 1/100 second. Note the multiple copies of the P wave

signature with one, two, or three leading zeros. These were introduced so that P waves could begin at any signature sample rather than at only every fourth signature sample (i.e. at a Markov chain transition). Similar remarks apply to the V and T waves in the V submodel.

Table 3.1 gives a summary of a few successive Wenckebach periods. Note the lengthening P-R intervals followed by a dropped beat. Note also that the model is not deterministic. For example, sometimes the Wenckebach cycle is four beats long and sometimes it is five.

Finally, Figure 3.21 gives the actual simulated ECG corresponding to the data in Table 3.1.

This completes the discussion of example ECG models based on the approach developed in this thesis.

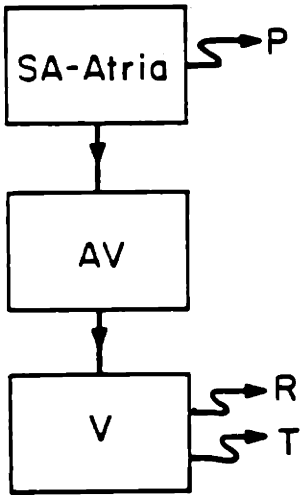
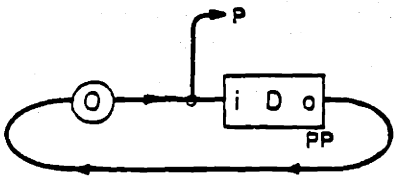
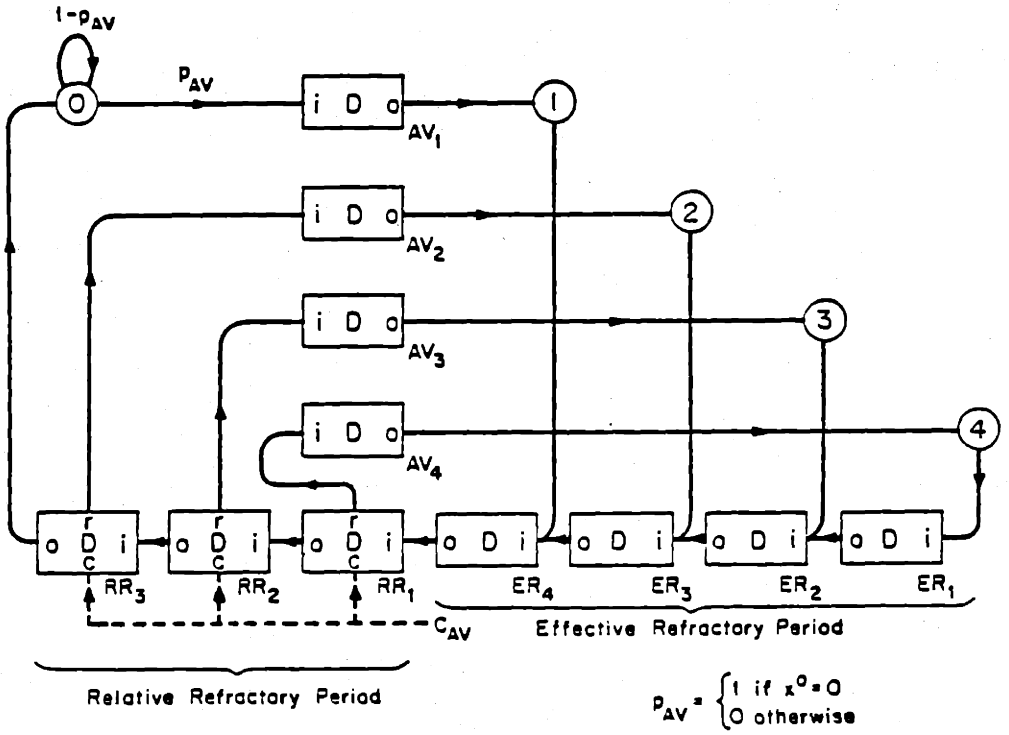


Figure 3.18 A Model For Wenckebach. Submodel Structure.

SA-Atria: x^0



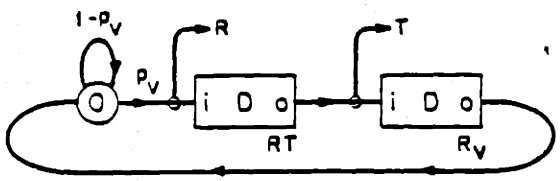
AV: x^1



$$p_{AV} = \begin{cases} 1 & \text{if } x^0 = 0 \\ 0 & \text{otherwise} \end{cases}$$

$$C_{AV} = \begin{cases} R & \text{if } x^0 = 0 \\ \bar{R} & \text{otherwise} \end{cases}$$

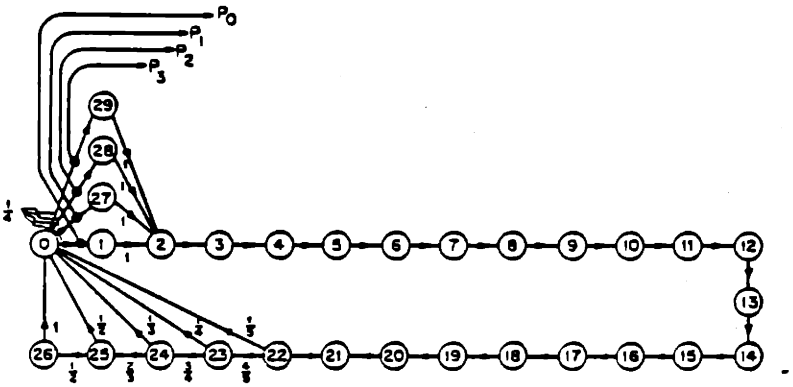
Ventricles: x^2



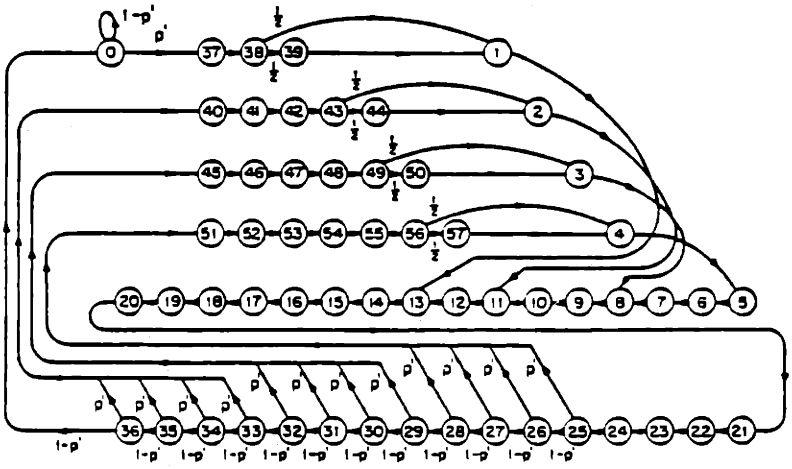
$$p_V = \begin{cases} 1 & \text{if } x^1 \in \{1,2,3,4\} \\ 0 & \text{otherwise} \end{cases}$$

Figure 3.19 A Model For Wenckebach. Block Diagram.

SA-Atrial: x^0

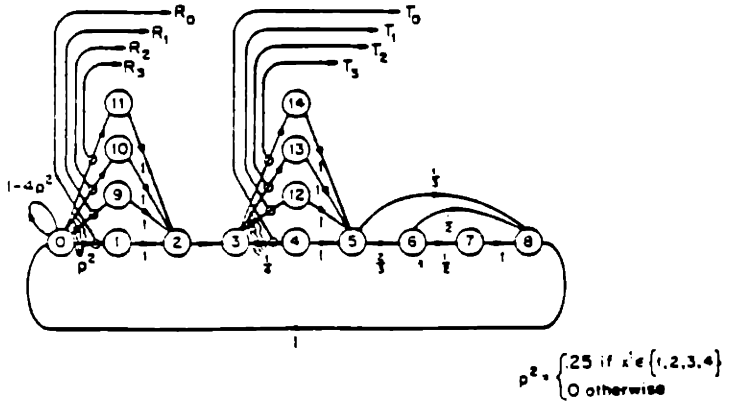


AV: x^1



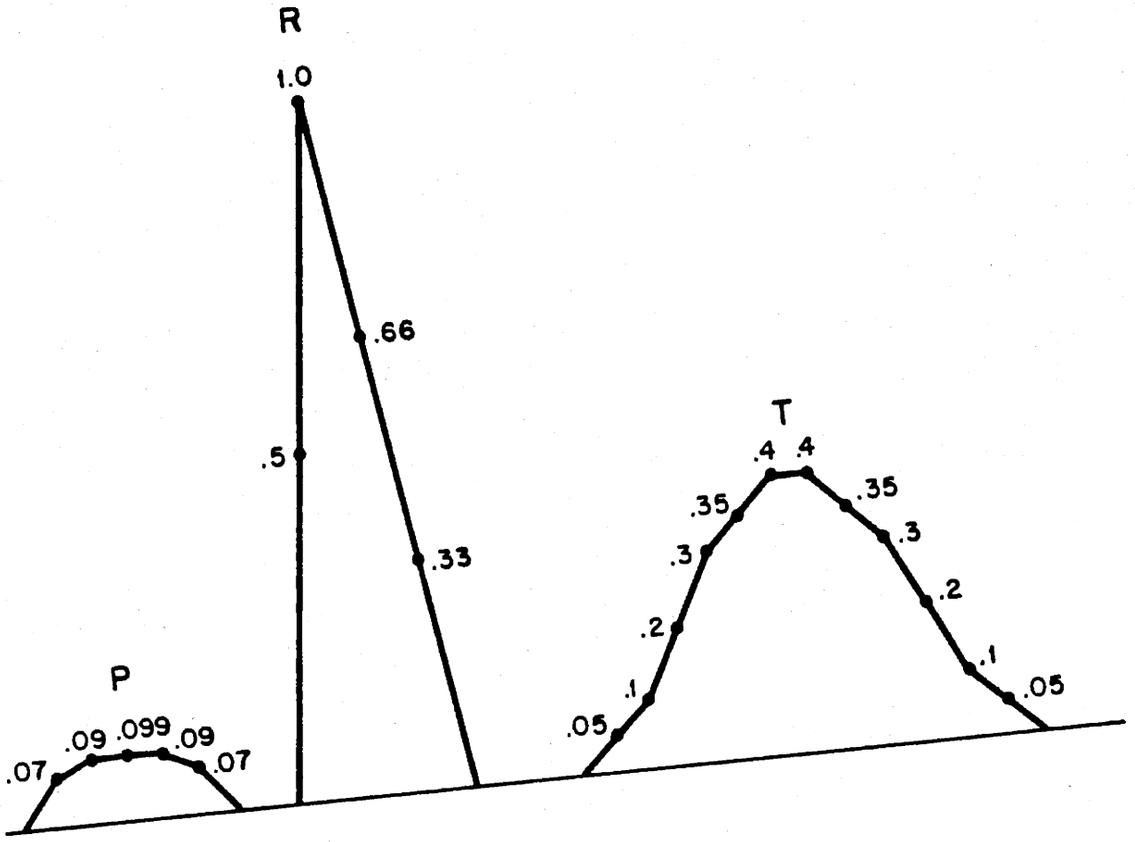
Ventricular: x^2

$p^1 \begin{cases} 1 \text{ if } x^0 \in \{0\} \\ 0 \text{ otherwise} \end{cases}$



$p^2 = \begin{cases} .25 \text{ if } x^1 \in \{1,2,3,4\} \\ 0 \text{ otherwise} \end{cases}$

Figure 3.20 A Model For Wenckebach. Markov Chains.



The covariances for the P, R, and T waves are all identically zero.

The X_i waves where $X \in \{P,R,T\}$ and $i \in \{0,1,2,3\}$ are defined to be

$$X_i(n)=X(n-i).$$

Figure 3.20 Continued.
Signatures.

P wave Number	Time (sec.)	P-R Interval (sec.)	Time Since Last P wave (sec.)
0	0	.13	
1	.99	.21	.99
2	2.07	.27	1.08
3	3.03	dropped	.96
4	3.94	.12	.91
5	4.89	.23	.95
6	5.96	.24	1.07
7	7.01	.33	1.05
8	8.05	dropped	1.04
9	8.99	.16	.94
10	9.99	.22	1
11	10.99	.23	1
12	11.91	dropped	.92
13	12.89	.18	.98
14	13.94	.22	1.05
15	14.98	.24	1.04
16	16.02	.28	1.04
17	16.96	dropped	.94
18	17.95	.09	.99
19	18.93	.26	.98
20	19.96	.3	1.03
21	20.97	.34	1.01
22	21.92	dropped	.95
23	22.88	.13	.96
24	23.96	.25	1.08
25	25.06	.29	1.1
26	26.04	.35	.98
27	27.08	dropped	1.04
28	28.13	.16	1.05
29	29.12	.27	.99
30	30.17	.27	1.05
31	31.11	dropped	.94
32	32.12	.17	1.01
33	33.2	.21	1.08
34	34.26	.28	1.06
35	35.32	.35	1.06
36	36.43	dropped	1.11
37	37.49	.16	1.06
38	38.41	.23	.92
39	39.46	.32	1.05

Table 3.1 Wenckebach: P-P and P-R Intervals.
The simulated ECG from which this interval data was computed is shown in Figure 3.21.

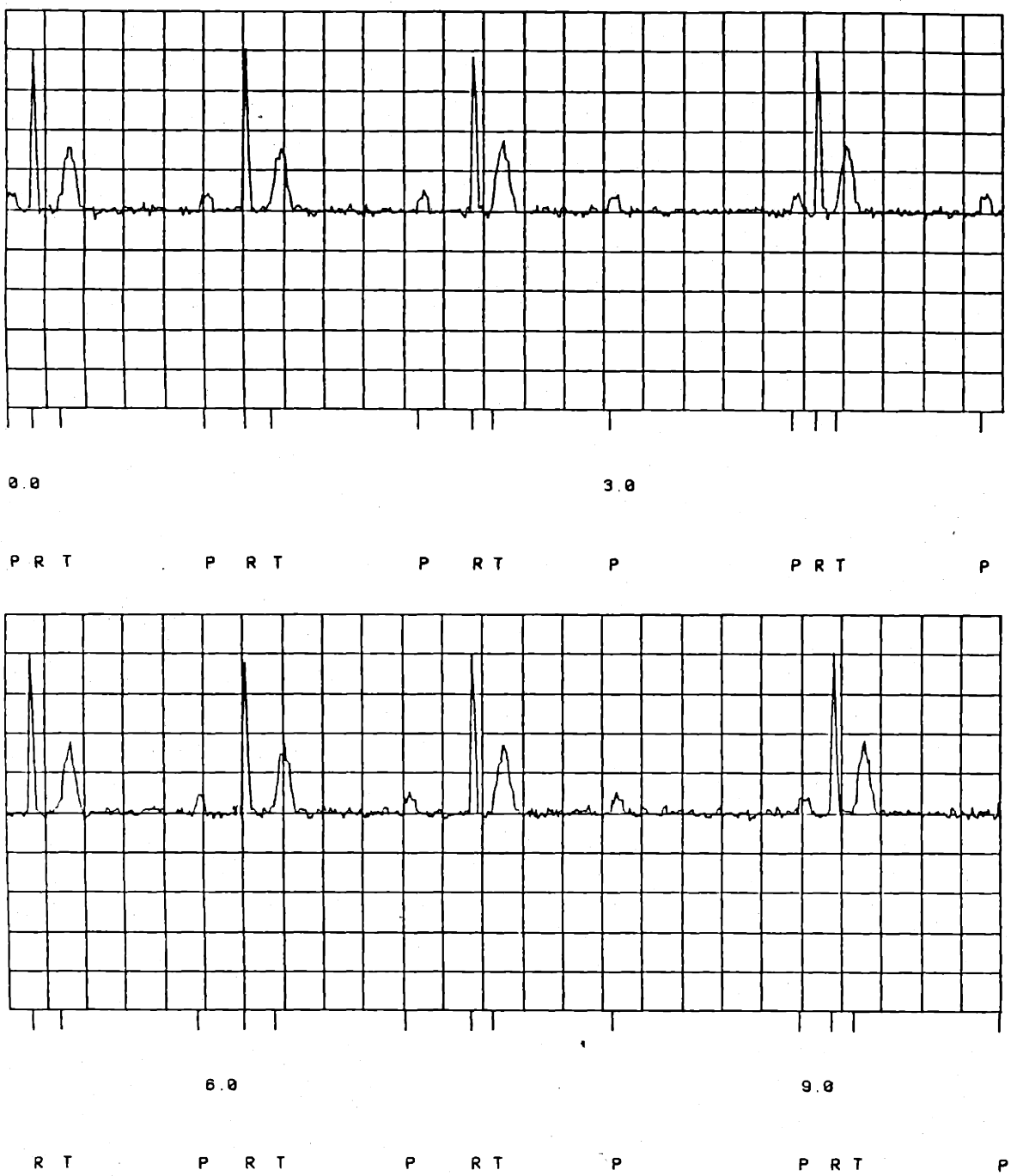


Figure 3.21 Wenckebach: ECG.
The interval data for this simulated ECG is tabulated in Table 3.1.

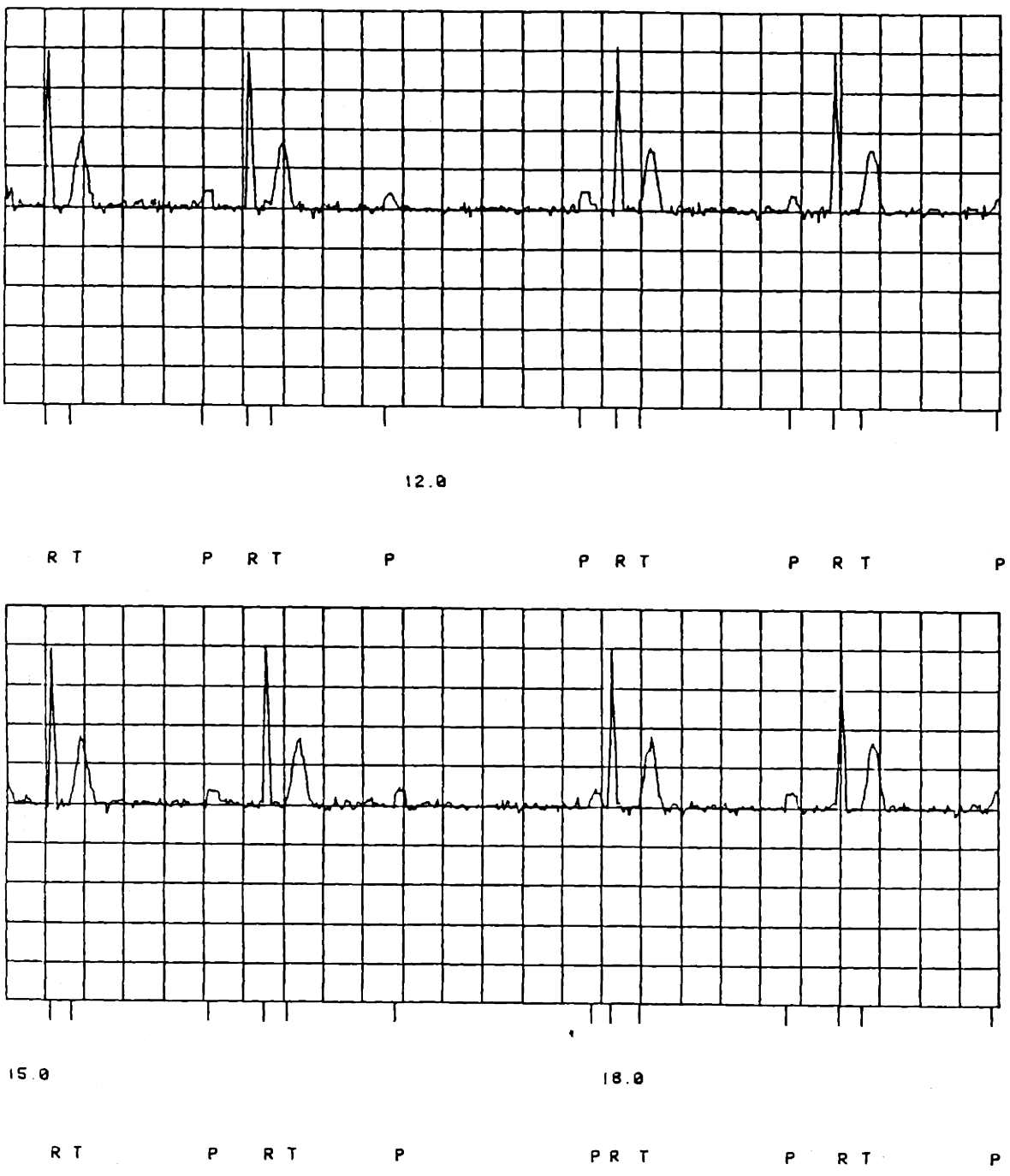
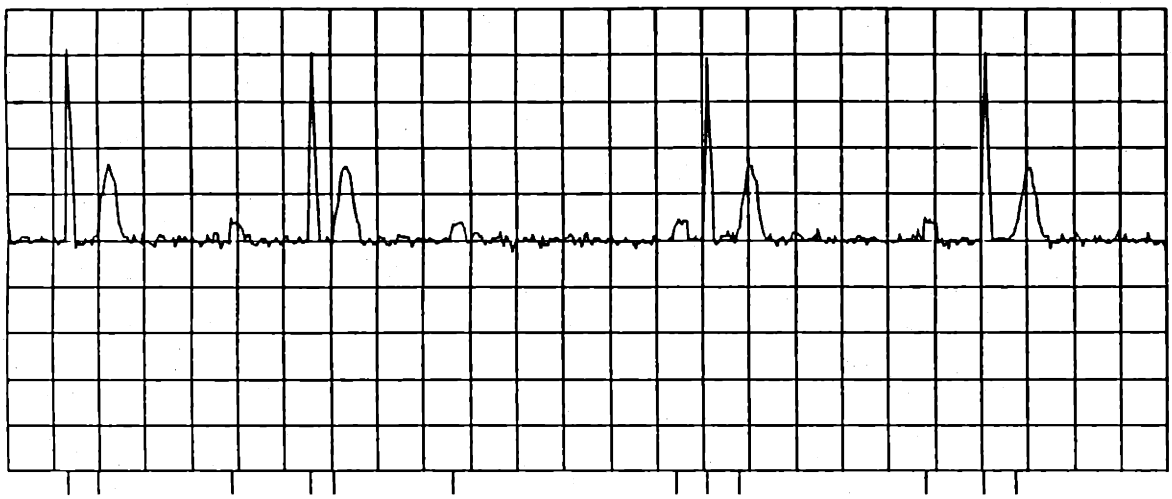
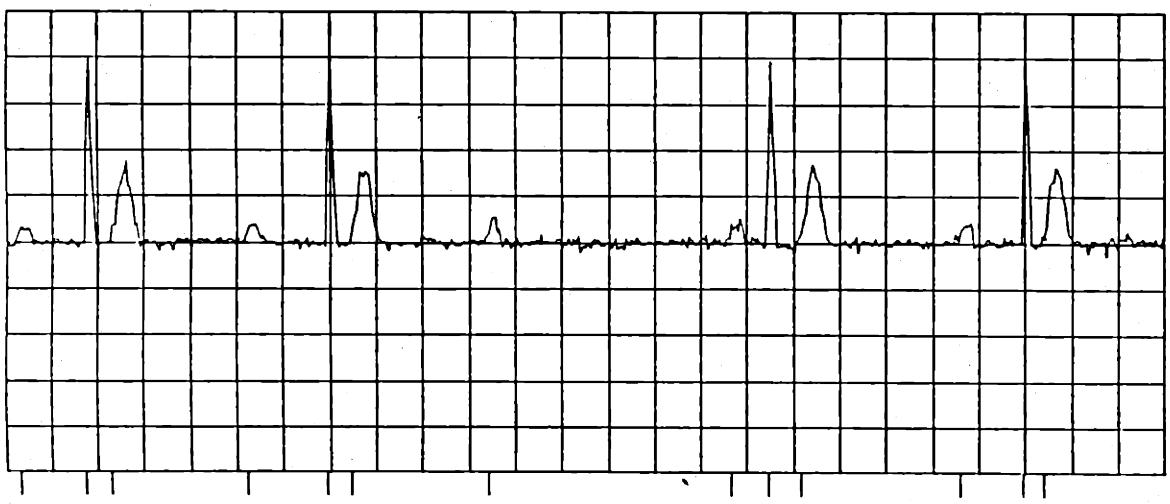


Figure 3.21 Continued.



21.0 24.0

R T P R T P P R T P R T



27.0

P R T P R T P P R T P R T

Figure 3.21 Continued.

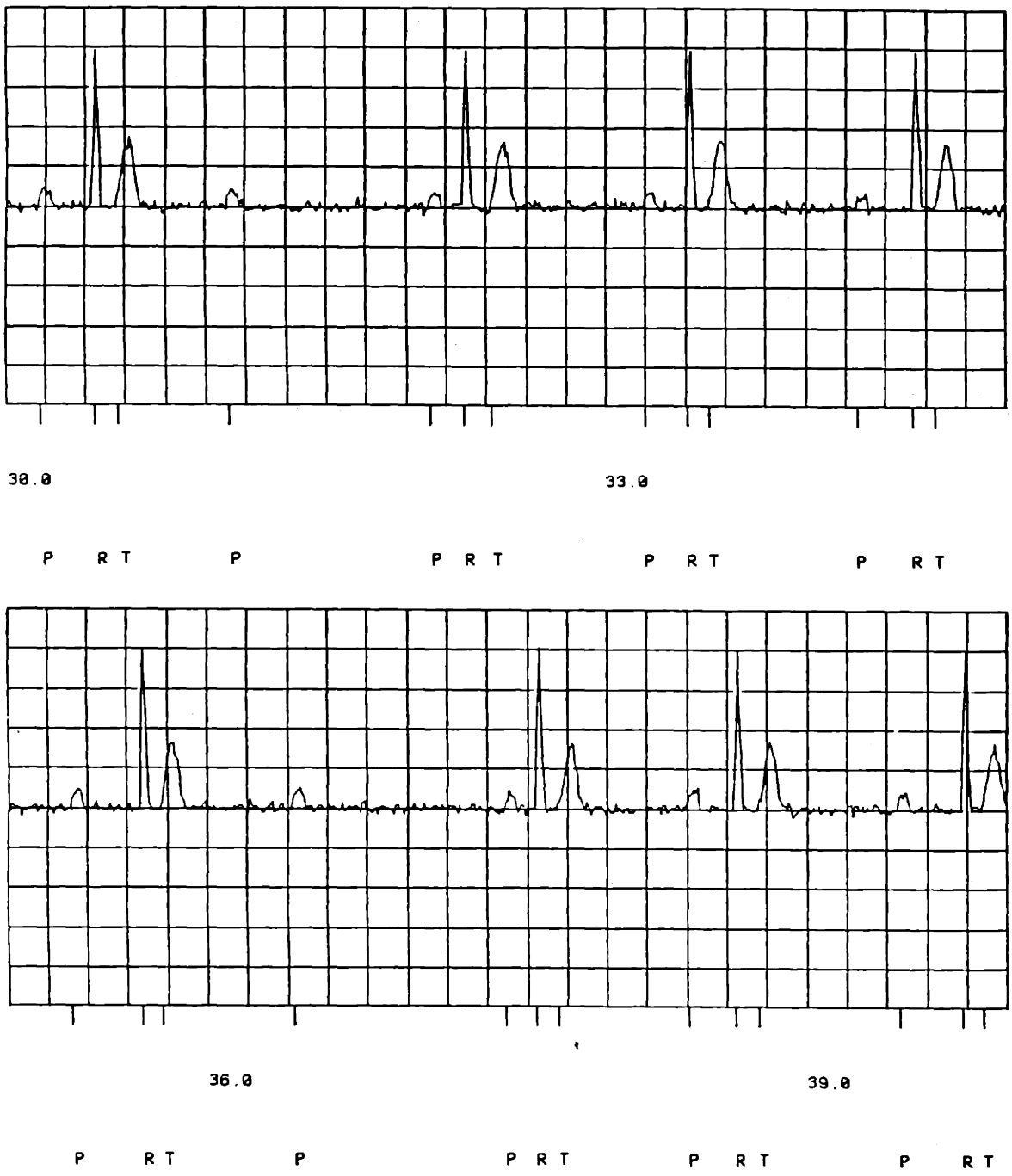


Figure 3.21 Continued.

3.9. Summary of Chapter 3

In this chapter we have described a modeling methodology. From the point of view of the ECG application, it emphasizes the event-oriented nature of the arrhythmia problem, the importance of causality, and the construction of models for complex rhythms out of interactions between a few simple building blocks. From the point of view of signal processing, it emphasizes decompositions (spatial, time-scale, and hierarchical) and control of the level of detail included in the model. This methodology may be applicable to other event-oriented systems but that possibility is not explored.

4. Signal Processing I: Mathematical Model and Performance Assessment

In previous chapters we have described the physiological motivation and mathematical formalism of our models. In the remainder of this thesis, we describe the design of signal-processing algorithms which exploit the mathematical properties of these models. In the present chapter we review the formalism of the mathematical model and then present the estimation goals and performance measures used in this thesis. In Chapter 5 we describe our design approach. Finally, in Chapter 6, we discuss the implementation of our design approach and present several case studies.

4.1. Review of the Mathematical Model

In the previous chapter a mathematical model of the ECG was proposed. This section reviews the mathematical structure of the model and fixes notation to be used throughout the remainder of this thesis.

Globally, the model has two types of structure: a spatial decomposition into interacting submodels and a two-level hierarchical decomposition within each submodel. (See Figure 4.1). Collectively, the upper hierarchical level of the submodels is called the physiological model and describes the dynamics. Mathematically, this is a discrete-time finite-state Markov process. The lower hierarchical level of the submodels forms a part of the electromagnetic model, which describes how events in the physiological model give rise to observed waveforms. The electromagnetic model is a linear superposition of signals with shifted origins which is additively observed in white Gaussian noise. The unshifted signals are called signatures. Each signature is a shift-

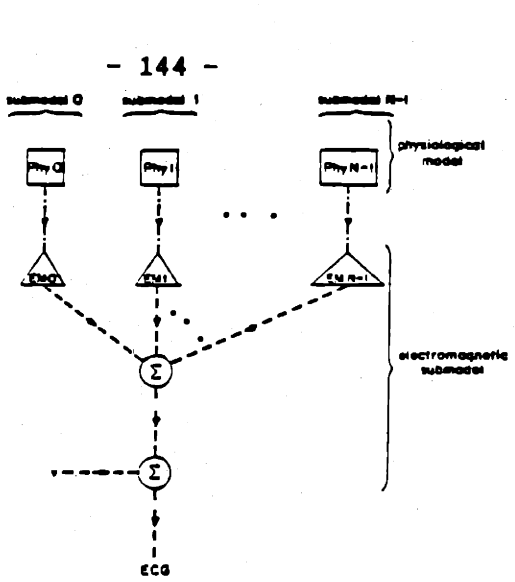


Figure 4.1 Global Structure of the Mathematical Model. Note that the interactions among Phy 0, ... , Phy N-1 are not shown.

invariant finite-durational deterministic function observed in additive white zero-mean Gaussian noise.[1] All parameters throughout the entire model are assumed known.

4.1.1. The Dynamics

The central feature of the physiological model is its spatial decomposition. This decomposition is reflected in the structure of the state space and transition probabilities. The state space, denoted S, is the Cartesian product of N subprocess state spaces, denoted S^i. Each subprocess state space represents one (or an aggregation of several) of the anatomic subunits of the heart. Because of the model's decomposition, the transition probability matrix on S, denoted P, is most easily described in terms of the S^i. The matrix P has a structure (described

[1] The statement that the signature is a deterministic signal observed in additive white zero-mean Gaussian noise is exactly equivalent to the statement that the signature is a white Gaussian signal. That is, the signature could equivalently be described as a shift-invariant finite-durational white Gaussian signal. The reason for using the prior, unwieldy description is that it emphasizes the nonzero mean.

in the following paragraphs) related to the S^i subprocess state spaces because an important function of P is that it captures the interactions between the subprocess state spaces, corresponding to the interactions between the anatomic subunits of the heart.

Let x(n) denote the state at time n, a random variable taking values in S. In terms of the factors of the state space S, we write this as $x(n)=(x^0(n),\ldots,x^{N-1}(n))$, $x^i(n)\in S^i$. For the realizations of these random variables, which are numbers, we use the notation x_n and x_n^i.

The interactions are defined through partitions of the subprocess state spaces. Specifically, for each ordered pair of submodel indices (i,j), there is a mutually exclusive, collectively exhaustive partition of the subprocess state space S^i into $\rho^{i,j}$ subsets. Let the disjoint subsets be denoted $U^{i,j,k}$, k=0, ..., $\rho^{i,j}-1$. By definition, submodel i transmits interaction k to submodel j when $x^i\in U^{i,j,k}$. (Note the directionality of the interaction). The definition of these interactions allows us to model the propagation of a depolarization wave across the boundary between the i^{th} and j^{th} anatomic subunits. Because these interactions are sparse, $\rho^{i,j}$ is typically 1 or 2. (Note that $\rho^{i,j}=1$ corresponds to a constant interaction,, which is equivalent to no interaction at all). In any case, the limited number of types of interactions between submodels implies that $\rho^{i,j}$ is small compared to $\text{card}(S^i)$. The fact that $\rho^{i,j}$ is small in this sense is central to our suboptimal estimators.

The actual interaction which submodel i transmits to submodel j at time n is denoted $h^{i,j}(n)$. This is a deterministic function of the

random variable $x^i(n)$. Its realization, a number, is denoted $h_n^{i,j}$. The above definition of the interactions states that $x_n^i \in U^{i,j,h_n^{i,j}}$.

The structure of P is defined by imposing two assumptions:

(1) given the collective past, the next transition of each submodel is chosen independently of the next transition of all other submodels and

(2) for a particular submodel i, the current impinging interactions and current submodel state form a sufficient summary of the collective past insofar as the single-step evolution of submodel i is concerned.

The above definition of P states that each element satisfies the equation

$$\Pr(x(n+1)=x_{n+1} \mid x(n)=x_n) = \prod_{i=0}^{N-1} \Pr(x^i(n+1)=x_{n+1}^i \mid x(n)=x_n) \quad \text{(assumption 1)}$$

$$= \prod_{i=0}^{N-1} \Pr(x^i(n+1)=x_{n+1}^i \mid x^i(n)=x_n^i, h^{j,i}(n)=h_n^{j,i},\ j=0,\ldots,N-1\ (j\neq i)) \quad \text{(assumption 2)}.$$

Note that if $\rho^{i,j}=1$ for all i,j then P will simply be the Kronecker product of individual transition probability matrices P^i for each submodel. This is the degenerate case in which the submodels do not interact. (This case is not common, though it can occur, as in a model of third degree AV block containing two submodels--one for the SA node and atria proximal to the block and one for the ventricles distal to the

block). Another special case arises in the situation where there exists an i^* such that $\rho^{i^*,j}$ and ρ^{j,i^*} are 1 for all j. This is a situation where one submodel (submodel i^*) neither initiates nor receives interactions. In this case P is the Kronecker product of P^{i^*} with the overall transition probability matrix for all the submodels except i^*. This case occurs in our estimators--see Chapter 6.

In addition to the transition probability matrix P specified above, an initial condition must be specified. Since the model is time-invariant, the time of the initial condition is arbitrarily set to be time -1. This thesis has not been concerned with initial transients due to the initial conditions. Therefore little attention has been paid to their choice. Primarily for convenience, we have assumed that they factor over the individual subprocess state spaces. That is,

$$\Pr(x(-1)=x_{-1}) = \prod_{i=0}^{N-1} \Pr(x^i(-1)=x^i_{-1}).$$

However, this assumption is not crucial to anything that follows.

This completes the description of the discrete-state physiological model.

4.1.2. The Observation Process

The electromagnetic model describes how events in the physiological model give rise to waveforms in the ECG. The submodels of the electromagnetic model are in one-to-one correspondence with the submodels of the physiological model. Furthermore, the state of an electromagnetic

submodel does not affect the state of any other submodel (electromagnetic or physiological) and is affected by only the corresponding physiological submodel's state. Specifically, each electromagnetic submodel is driven by the state transitions of the corresponding physiological submodel. The outputs of the electromagnetic submodels are linearly superposed in additive white Gaussian noise to create the observation.

The output of the i^{th} submodel of the electromagnetic model is a linear superposition of finite-durational signals. Each signal is a different realization of a stochastic process for which the probability law(s) (there can be more than one class of signal per submodel) are time invariant. The stochastic process(es) defined by the probability law(s) are called <u>signatures</u>. More specifically, if the i^{th} submodel of the physiological model undergoes a state transition from j to k at time t, then the contribution at time n to the linear superposition is the signal r(n;i,j,k,t). The signal r(n;i,j,k,t) is a realization of a stochastic process whose probability law is time invariant (i.e. it depends on n and t only through the difference n-t). The probability law is finite-durational deterministic function plus additive white zero-mean Gaussian, which is equivalent (as noted in the previous footnote) to white Gaussian. Furthermore, the signal r(n;i,j,k,t) is independent of the signal r(n;i,j,k,t') for $t \neq t'$. Let N(q,Q)(x) be the Gaussian distribution of mean q and covariance Q evaluated at the point x. The signal r satisfies the following relationships:

(1) finite durational:

$$r(n;i,j,k,t)=0 \quad n\notin\{t,\ldots,t+M^i_{j,k}-1\} \text{ and}$$

(2) time-invariant probability law described previously:

$$p_{r(t;i,j,k,t),\ldots,r(t+M^i_{j,k}-1;i,j,k,t)}(r_t,\ldots,r_{t+M^i_{j,k}-1})$$

$$= \prod_{m=0}^{M^i_{j,k}-1} N(q^i_{j,k}(m),Q^i_{j,k}(m))(r_{t+m})$$

where $q^i_{j,k}(m)$ and $Q^i_{j,k}(m)$ are the known mean and covariance functions of the signature.

Define $M^i = \max_{j,k} M^i_{j,k}$. Let the state trajectory of the i^{th} submodel of the physiological model be $X^i_{0,N} = \{x^i_0,\ldots,x^i_N\}$. Then the output, denoted $e^i(n)$, of the i^{th} submodel of the electromagnetic model at time n is

(3) linear superposition:

$$e^i(n)=\sum_{t=0}^{N} r(n;i,x^i(t-1),x^i(t),t)$$

$$= \sum_{t=n-M^i}^{n} r(n;i,x^i(t-1),x^i(t),t).$$

Finally, the output of the entire electromagnetic model, denoted y, is the linear superposition, with additive stationary white Gaussian noise, of the outputs of the submodels:

(4) linearly superposed in additive noise:

$$y(n)=\sum_i e^i(n)+v(n).$$

The additive white Gaussian noise, denoted v, has mean 0 and covariance R.

Typically, for each submodel of the physiological model, most state transitions (i,j) do not initiate signatures. More precisely, they map to a signature that has zero mean and zero covariance. Such signatures have no effect on the observation and are included for notational convenience only. In the sequel, they are referred to as <u>null</u> signatures.

4.1.3. Comments on the Mathematical Model

Note several aspects of this model:

(1) The dimension of the observation vector y is not specified. Statistical correlations between the components of the added noise v are modeled through the off-diagonal elements of R. Similarly, statistical correlations between the components of r are modeled through the off-diagonal elements of $Q^i_{j,k}(.)$.

(2) The Markov-chain cycle interval need not equal the signature sampling interval. However, the larger must be an integer multiple of the smaller. Note also that signatures can only be initiated at Markov chain transitions. The use of different rates allows the events within the physiological model and the waves in the ECG to be modeled with different levels of temporal resolution.

(3) An individual signature may last many Markov chain cycles.

(4) A single global state transition may initiate as many shifted signatures $r(.;i,j,k,l)$ as there are submodels.

This completes the description of the mathematical model.

4.2. Estimation Goals

The estimation goals of this thesis are to solve an interesting subset of ECG rhythm-analysis problems. Specifically, the only estimation problems considered in this thesis are those that can be reduced to the wave tracking problem. That is, given an ECG and a single completely known model, we seek to determine the number of waves and the time of occurrence and type of each individual wave. We have chosen this problem because it is relevant to the arrhythmia classification problem, it is relatively simple, and it is a necessary component of more complex problems.

The wave tracking problem is clearly not the entire arrhythmia classification problem since the wave tracking problem assumes a specific model for the ECG. As discussed in Chapter 1, the long range goal of the research program of which this thesis is a part is the

classification of ECGs into arrhythmia classes (persistent rhythms) and the location of ectopic events (transient rhythms). As in Gustafson et al. (Gustafson, 1978a, 1978b), the hypothesis test required for persistent rhythm classification involves likelihood ratios. The optimal computation of likelihood ratios involves a set of processors, one for each hypothesis (equivalently, ECG model), where the dimension of each processor's state equals the cardinality of the overall Markov chain state space for that model. Therefore, for computational reasons, the calculation of exact likelihood ratios is not practical. A natural approximation, however, involves the solution of the wave tracking problem considered in this thesis. Specifically, in order to calculate the likelihood of a particular hypothesis, each of the optimal processors performs a recursive summation of conditional likelihoods over all possible state trajectories in the overall Markov chain. A natural approximation is to include only the probabilisticly important trajectories. That is, a standard way in which hypothesis testing is often performed is to design tracking filters based on each of the hypotheses. Then, the relative "likelihoods" of the various hypotheses are deduced by comparing how well each of the tracking filters is performing. In the ECG context such a system would consist of a bank of wave trackers of the type that are developed in this thesis.

Now consider the problem of transient rhythm analysis. There are two approaches to this problem.

(1) As in Gustafson et al., we can model transients as switches from one model to another. The approximate calculation of the likelihood ratios in this approach again involves wave tracking

estimation for the patterns before and after the transient.

(2) Consider transient rhythms which do not correspond to a lasting rhythm change (e.g. a change in heart rate). In this case, assuming that we include the mechanisms for such transients in our model of the identified persistent rhythm class (e.g. inclusion of reentrant pathways or ectopic foci), then detecting such transients is purely a wave tracking problem.

Thus we see that wave tracking is an essential ingredient in rhythm analysis and represents a natural first step in developing this approach to ECG analysis.

4.3. Performance Measures for the Wave Tracking Problem

The performance measures by which estimation algorithms will be evaluated fall into two classes: those relevant to the application and those relevant to the internal workings of the estimator. The later depend on the details of the estimator and are discussed in the sequel. In this section, measures relevant to the application are discussed.

From the arrhythmia classification point of view, the relevant measures are concerned with the type and time of occurrence of the waveforms in the ECG. The type of a waveform is specified by the submodel from which the non-null signature was generated and the specific transition within the submodel that initiated the signature--i.e. the old and the new state in the subprocess state space. The ordered quadruple of submodel, old state, new state, and time is called an annotation.[2]

---

[2] Other applications require different performance measures. For example, the ECG morphology problem is concerned with wave

As described below, all performance measures for the specific estimators we have developed have been evaluated via simulation. Furthermore, all the performance measures calculated are probabilities of particular events. Therefore the following discussion is oriented toward the simulation method for evaluating probabilities.

The evaluation of probabilities by simulation requires four basic steps:

(1) Generation of a sample path of the stochastic process. A realization of the random variables has been chosen using a pseudo-random number generator.

(2) Application of the estimator under evaluation to the simulated data.

(3) Association of annotations in the simulated data with annotations in the estimate. This requires an association rule (described in the following subsections). Since estimators will generally not be perfect, not all events in the simulated data will have corresponding events in the estimate and likewise not all events in the estimate will have corresponding events in the simulated data. Thus the association rule must deal with different numbers of annotations, as well as different types and times.

---

shape and thus with wave shape estimation, for which an appropriate performance measure could be mean squared error.

(4) Counting the number of times certain patterns occur in the associated events. This requires a set of selected patterns (described in Subsection 4.3.4) that represent important characteristics of estimator behavior. Properly normalized, these counts are the desired probabilities.

It is steps (3) and (4), the association rule and the selected patterns, which define the performance criteria. Step (1), the calculation of the sample path, was discussed at length previously (Chapter 3 and Section 4.1) while step (2), the estimator, is discussed in Chapters 5 and 6. Steps (3) and (4) are discussed here.

The development of meaningful association rules and the specification of patterns that are appropriate indicators of estimator performance are not simple problems. The solution is a procedure for the comparison of two sequences of annotations--one for the actual simulated data and one for the estimator. Because there are several possible signature types, each of which may occur a random number of times with each occurrence falling at an uncertain time, we see that a major issue is "lining up" the annotations in the true and estimated sequences and in identifying resulting paired annotations as one of several selected patterns. For example, if a particular estimated annotation corresponding to the start of a QRS complex occurs M time units after the actual QRS complex, and if the duration of the complex is K cycles, then whether we wish to call this an "error" or not depends on whether M is small or large compared to K. Furthermore, an error in the QRS-complex annotation may very well be coupled with an error in the preceding P-wave annotation. Do we classify this as one type of error or do we

classify them individually? We have developed a specific approach to association and have chosen a class of selected patterns on which to base our performance computations that we feel are appropriate for the types of questions one would typically like to answer in evaluating an ECG waveform tracker. In the remainder of this chapter we describe our approach to addressing these issues.

4.3.1. Association Rule

An association rule is an algorithm. Its input is a pair of annotation sequences. One annotation sequence represents the truth corresponding to the observed data, and the other represents the estimate provided by an estimator. The output of the association rule is a sequence of matched annotations. A match consists of one truth annotation, one estimated annotation, and a phase shift between the two, corresponding to the timing error in locating the event.

Consider the motivating examples in Figure 4.2. In this figure, the tick marks indicate the passing of discretized time and the dashed lines connecting tick marks have been included to facilitate visual alignment of the time axes for the simulated data (labeled "Truth") and the output of the estimator (labeled "Estimate"). The truth trajectory consists of five atrial depolarization (P waves) of which four result in ventricular depolarizations (R waves). The remaining atrial depolarization is blocked.

For the (P_1,R_1) beat the desired association is clear since each estimated wave is perfectly aligned with the corresponding truth wave. The situation for the (P_2,R_2) beat is slightly more complex. In this

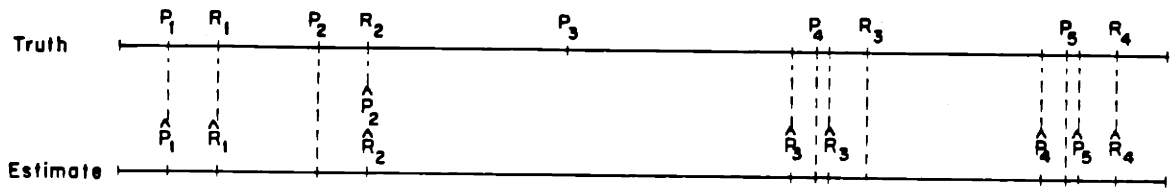


Figure 4.2 Motivation for the Association Rule.

case, the estimator has two annotations at one instant in time, corresponding to estimated events in two submodels, i.e. the occurrence of both a P wave ($\hat{P}_2$) and an R wave ($\hat{R}_2$). Clearly what one would want to occur in this case is for the association algorithm to associate P_2 with $\hat{P}_2$ and R_2 with $\hat{R}_2$. The reason for desiring this association is based on morphological considerations--i.e. the R wave and the P wave look very different, and thus since both annotations are present, it makes little sense to match R_2 with $\hat{P}_2$ and P_2 with $\hat{R}_2$.

Continuing on, P_3 does not match at all because there is no estimated wave sufficiently close (clearly a quantitative meaning must be given to "sufficiently close"). Such an unmatched truth annotation is called a <u>false negative</u>.

For the (P_4,R_3) beat, the decision again should be based on morphology considerations. Based purely on phase shift considerations, P_4

could match with either $\hat{P}_3$ or $\hat{R}_3$. However, it should match with $\hat{P}_3$, the more distant choice, because of the superior morphology match. Having matched P_4 and $\hat{P}_3$, $\hat{R}_3$ cannot also match with P_4 which is the nearest truth annotation (though a poor match from the morphology point of view). Therefore, R_3 matches with $\hat{R}_3$.

In the (P_5,R_4) beat, P_5 matches with $\hat{P}_5$ rather than $\hat{P}_4$ because $\hat{P}_5$ is closer. R_4 matches with $\hat{R}_4$ leaving $\hat{P}_4$ unmatched. The only candidate for $\hat{P}_4$ is P_3 but the phase shift is much to great so they both remain unmatched. Such an unmatched estimated annotation is called a <u>false positive</u>.

Based on this set of examples, it is clear that the desired association rule must have several characteristics:

(1) It must weigh both timing and morphology (wave type).

(2) At some point, a large phase shift must disallow a match.

(3) Between two candidates either of which could be matched to a third signature, a superior morphology match must sometimes be able to override a poorer temporal match--i.e. it must attempt to match wave types within some bounds of reasonableness on relative timing.

We have tried to capture these characteristics in a simple rule. The rule takes the form of an optimization problem. The optimization is over all possible pairings of a truth annotation (or false positive) with an estimated annotation (or false negative). The reward function is an additive function with one term for each matched pair of truth and estimated signatures plus additional terms for each unmatched truth or

estimated signature. The key to the rule is how to jointly weight differences in morphology and phase shifts.

The rule used in this thesis has two subrules:

(1) Matches between nonoverlapping signatures are not allowed. This is the specific interpretation of (2) above. Therefore, the reward for such a match is $-\infty$.

(2) The reward for a match between annotations with overlapping signatures is equal to the crosscorrelation function between the truth and estimated signatures evaluated at their relative phase shift.

In more detail, the reward in subrule (2) for matching a truth signature of type m_1 (which is initiated by transition j_1 to k_1 in submodel i_1) with an estimated signature of type m_2 (which is initiated by transition j_2 to k_2 in submodel i_2) where the relative phase shift is τ signature samples and the signatures overlap (a situation illustrated in Figure 4.3) is

$$\frac{\sum_n q^{i_1}_{j_1,k_1}(n)\hat{q}^{i_2}_{j_2,k_2}(n-\tau)}{\{\sum_n [q^{i_1}_{j_1,k_1}(n)]^2\}^{1/2}\{\sum_n [\hat{q}^{i_2}_{j_2,k_2}(n)]^2\}^{1/2}}.$$

The functions $q^i_{j,k}$ and $\hat{q}^i_{j,k}$ were defined in Section 4.1--they are the mean of the Gaussian signal in the model used to generate the simulated data and in the model used by the estimator respectively. (These two models need <u>not</u> be the same--e.g. the estimator may use simplified templates). The range of each summation is over those n for which the

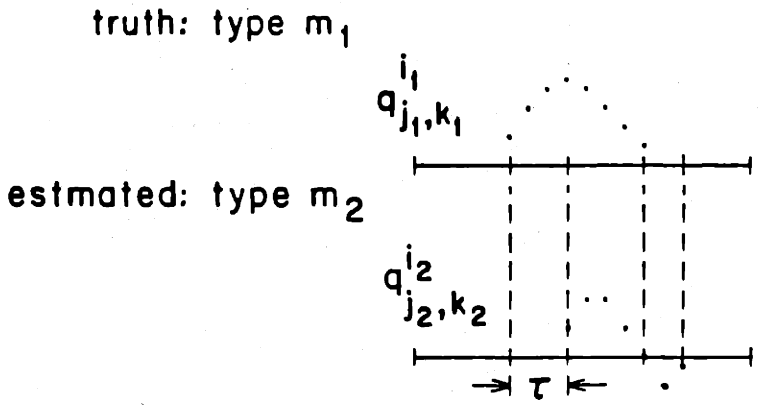


Figure 4.3 Calculation of the Reward.

summand is nonzero.

The first subrule limits the maximum phase shift. The second subrule rewards matches between annotations whose signatures are morphologically similar in spite of the phase shift. Therefore, for some signature morphologies, the penalty for phase shifts increases much more rapidly than for other morphologies.

A complete definition of the association rule is given in Appendix A.

4.3.2. Association Rule--Phase Shift Limits

The patterns used in the final step of the performance calculation, as described in the following subsection, typically depend upon the phase shift between matched annotations. For small phase shifts such that the signatures remain overlapped, the association method of the prior subsection is applicable. Furthermore, larger phase shifts typically seem unphysical--rather than match two such distant waves it seems

more reasonable to label one a false positive and one a false negative. Therefore, the association method of the prior subsection appears generally applicable--it will allow phase shifts up to the overlap limit. However, there is one situation which we have encountered where associating annotations with nonoverlapping signatures is necessary. Specifically, in our development of an estimation methodology we have considered a sequence of cases, each of which is more complex than the one that preceded it. In the first few cases we considered very simplified waveform structures in which "P" and "R" waves are only one sample long. Clearly, if the signatures are only one sample long, then any nonzero phase shift makes them nonoverlapping. However, it seems unreasonable not to allow small phase shifts (i.e. ± 1) since their exclusion is due to the extraordinary idealization of the signatures. In other words, the overlap-based notion of a reasonable phase shift is directly tied to the length of the signatures so that if the signatures are made artificially short, this method of calculating a reasonable phase shift becomes artificial. Two options which recover a reasonable definition of maximum phase shift in these situations while remaining within the framework of the present association method are briefly described in the following.

(1) Pad with zeros the trailing edge of the one sample long signature(s) and proceed with the association method of this subsection.

(2) Add annotations for artificial, zero signatures one sample long to submodel transitions which deterministically precede or follow the transition(s) which initiate(s) nonzero one-sample-long signatures.

Then proceed with the association method of this subsection. By examining the resulting matches (which are all zero-phase shift matches) it is possible to recover the phase shift information--i.e. the matching of one of these artificial annotations with an actual wave would correspond to a phase shift of ± 1.

These two options need not generate precisely the same results. However, we believe that either is an equally satisfactory definition. In all of our computation with signatures that are only one sample long, we have used the second option. Furthermore, when evaluating the effect of signature length on performance, we have maintained consistency by using the second method for all computation if any signature in any case was only one sample long. Note that if nonzero signatures are more than one sample long, the zero-signature method can be ambiguous. For example, suppose a nonzero signature overlaps two different zero signatures (which start at different times), but overlaps no nonzero signature with which it has a positive correlation coefficient. In order to remove the ambiguity, we have required that all matches with artificial annotations be zero-phase-shift matches (which are what naturally occur when all signatures, zero and nonzero, are one sample long). Further discussion of zero signatures and of the exact-match requirement is contained in Appendix B.

4.3.3. Association Rule--Summary

In the prior two subsections an association rule based on an optimization problem is proposed. We also discussed several slight variations in the general framework (e.g. the addition of artificial

annotations). Most of these variations have no effect on the results presented in the Chapter 6, because these results are for models with very few signatures which occur well dispersed in time.

4.3.4. Distinguished Patterns in the Associated Events

This subsection describes the class of associated-event patterns which will be considered in evaluating estimator performance. For all wave tracking problems, the most basic performance measures are the probabilities of matching a truth annotation of type α_1 at time n with an estimated annotation of type α_2 at time $n+\tau$ where τ is the phase shift. These probabilities are called the misclassification probabilities. Therefore the first class of patterns is the class of all possible truth-estimated annotation matches at a given phase shift τ.

More aggregated performance measures can be computed from these. For example, a false positive and false negative probability can be computed for each type of annotation. Consider a specific annotation type, denoted α. Each of these probabilities is calculated as a ratio of event counts. The numerator of the α false negative probability is the number of times an α truth annotation is not matched with an α estimated annotation within a given phase shift. The denominator is the number of α truth annotations. The numerator of the α false positive probability is the number of times an α estimated annotation is not matched with an α truth annotation within a given phase shift. The denominator is the number of α estimated annotations.

These definitions imply characteristics of the false positive and false negative probabilities. First, note the different normalizations.

The false negative normalization is the traditional hypothesis-test normalization. The false positive normalization was chosen differently, as described above, so that the resulting probability reflects the probability of false positives that a further interpretation algorithm operating on these detections would have to deal with in its input.

Second, consider a mismatched truth and estimated annotation pair, that is, a match in which the truth and estimated signatures are not of the same type. The definition of false positive and false negative probabilities splits such a match into a pair of false positive and false negative detections.

When the model used to generate the simulated data contains only one signature type,[3] the false negative and positive probabilities described above are the only performance measures to be considered. When the model used contains two different signature types, we have considered examples that are strongly motivated by the normal ECG. That is, the two signatures are meant to represent the P and R waves. Therefore, the P wave has a lower signal-to-noise ratio (abbreviated SNR) than the R wave, the P wave precedes the R wave, and the P wave is causally related to the R wave.

The basic performance measures for models with both P and R waves

---

[3] Note that in many of the cases we have considered there are several different state transitions within a particular submodel that produce the same signature (which is one method for modeling variability in signature initiation time). Using the definitions we have given, each of these different initiating transitions would produce a distinct annotation. Clearly one would not want to consider a match of two of these annotations as an "error", and our method of event counting takes this into account.

are still the misclassification probabilities described previously. However, in this more complex problem, there are also additional measures of importance. For example, in order to locate the low SNR P waves, the estimator must essentially extrapolate backwards from its estimate of the R wave location in addition to weighing the data around the P wave itself. The extrapolation depends on inverting the causality relationship between the P and the R waves. (Note that this backward extrapolation is exactly what a human does when faced with this task). Because of our interest in this mechanism, two additional performance measures are calculated. They are both misclassification probabilities where the events are sequences of annotations rather than individual annotations. Specifically, estimates are calculated of

(1) Pr(estimated annotation pair is A_1, A_2 | truth annotation pair is P,R)

(2) Pr(estimated annotation matched to truth P is A|

truth annotation pair is P,not R)

where "P, not R" denotes the situation in which there is is a dropped R wave. Here A_1, A_2, and A can be P- or R-wave annotations <u>or</u> they can denote <u>no</u> annotation--i.e. a missed detection. Thus probability (1) tells us something about the correlation in our ability to detect P and R waves that occur in the normal, paired fashion, while probability (2) tells us how well we do in locating P waves that have not led to successful ventricular contractions. The patterns and normalization

factors required to estimate these two probabilities are chosen analogously to the false positive/false negative statistics.

4.3.5. Confidence Limits

When estimating a quantity, such as the performance probabilities described previously, it is also necessary to calculate some measure of confidence in the estimate. Our confidence measure is a quantity denoted fractional standard deviation, which measures the accuracy with which a probability is estimated.

In order to compute confidence limits, we assume a simple model for the occurrences of patterns in subsequences of the matched (i.e. paired actual and estimated) annotation sequence. An example pattern is false positive R waves. Here the corresponding subsequence is that of matched annotations in which the estimated annotation is an R wave. We model the occurrences of the particular pattern in the relevant subsequence and our estimate of its probability of occurrence as follows:

(1) The sequence of occurrences of the specific pattern in the subsequence of matched annotations is modeled as a Bernoulli process with parameter p. That is, at the k^{th} matched annotation in the subsequence we assign a random variable b(k) which takes the value 1 with probability p (corresponding to an occurrence of the pattern) and the value 0 with probability 1-p. Furthermore, b(k) is independent of b(m), $m \neq k$. The realization of b(k) (i.e. the actually observed sequence of occurrences and "non-occurrences" of the pattern) is denoted b_k.

(2) The estimate $\hat{p}$ of p is calculated as

$$\hat{p}=\frac{1}{n}\sum_{k=1}^{n} b_k .$$

Let n_+ be the number of times that b_k takes on the value 1, $k\in\{1,\ldots,n\}$. Then $\hat{p}$ can also be written as

$$\hat{p}=\frac{n_+}{n}.$$

Let E be the expectation operator. It is simple to show that

$$E\hat{p}=p$$

$$E(\hat{p}-p)^2=\frac{1}{n}p(1-p).$$

The fractional standard deviation d is defined as

$$d=\left[E\left[\frac{p-\hat{p}}{p}\right]^2\right]^{1/2}.$$

Therefore,

$$d = \frac{1}{\sqrt{n}} \left[\frac{1-p}{p}\right]^{1/2}.$$

Note that, for our problem, d is an optimistic assessment of the accuracy of p because our observations--i.e. the sequence of occurrences of the pattern--are not strictly independent.

According to the previous formula, the calculation of d requires that we know the value of p. However, we do not know p. Therefore, the best we can do is to compute an estimate of d. The estimate we choose, denoted $\hat{d}$, is the estimate calculated by replacing p by $\hat{p}$ in the previous formula. That is,

$$\hat{d} = \frac{1}{\sqrt{n}} \left[\frac{1-\hat{p}}{\hat{p}}\right]^{1/2}.$$

Using the definition of $\hat{p}$, we can write this as

$$\hat{d} = \left[\frac{1}{n_+} - \frac{1}{n}\right]^{1/2}.$$

Now consider an example. Assume that we are processing data which contains 1000 truth R waves. Assume further that our estimator correctly detected 980 R waves, had 22 additional R wave detections (false positives), and missed 20 R waves (false negatives). Then we would compute the probability of a false positive R wave detection as fpR = 22/(980+22) = .0220 with estimated fractional standard deviation

of $(1/22 + 1/(980+22))^{1/2}=.211$. The probability of a false negative R wave detection would be fnR = 20/(980+20) = .0200 with estimated fractional standard deviation of $(1/20 + 1/(980+20))^{1/2}=.221$. Note, as described in the previous subsection, that different values of n are used in the false positive and false negative calculations. The interpretation of these numbers is that, for instance, fpR lies in the range .0220 $\pm$ (21.1 percent of .0220) with probability .68.[4]

Because our estimators perform well on realistic data, the error probabilities are small numbers and therefore it is difficult to get high accuracy estimates of them by simulation. In order to alleviate this problem we often use models with artificially high noise levels. Furthermore, we emphasize comparison with the performance of the global MAP estimator[5] on the identical sample realization and study of plots of the data with both true and estimated annotations indicated.

4.3.6. Robustness

As we will describe in subsequent chapters, our approach to defining estimators is based on the use of a hypothesized design model--i.e. a specific model of the type described in Chapter 3. We will call such an estimator a matched estimator if the data on which it operates is a realization of the stochastic process defined by the identical design

---

[4] Here we make the further assumption that a central limit theorem holds, and therefore the probability within plus/minus one standard deviation of the mean is .68.

[5] Throughout this thesis, "global MAP estimator" is the wave-tracking estimator in which first the global MAP state-trajectory estimate is computed and then the wave-tracking estimate is computed from the state-trajectory estimate. Note that this estimator is not optimal for the wave-tracking problem.

model on which the estimator was based. In any real world problem, an estimator is never precisely matched because the data is not a realization of any particular design model. This is especially true in the ECG where the mechanisms are complex and poorly understood. Therefore it is crucial to evaluate the effect of mismatch on performance. This issue is termed robustness.

A realistic assessment of robustness depends on the recognition of which features in a design model are heavily relied upon by the estimator. Especially important are situations where such features are clearly idealizations of the ECG. For example, in a normal rhythm model, such an idealized feature might be the assumption that every P wave is followed by an R wave. Given such a special feature, a second design model can be defined in which that feature is modified. Then the altered design model can be used to create simulated data, which is then processed by the estimator designed using the original, idealized model. Comparison of the resulting performance statistics then indicates the robustness of this estimator to this particular idealization.

Generally, recognition of some aspect of a design model as an idealization whose consequences must be investigated also indicates which performance criteria are most important in assessing robustness. In the previous example where the feature is every P wave followed by an R wave, the important statistics are the joint P,R statistics described previously.

5. Signal Processing II: Design Approach

5.1. Estimator Design Philosophy

In this section, we describe the estimator design philosophy of this thesis. First, the possibility of globally optimal designs is considered and rejected. Second, we discuss the philosophy of the suboptimal approach adopted for this thesis. Throughout the discussion of the suboptimal approach, it is important to keep in mind that there are certainly many alternatives to the specific design choices that have been made. Simulation results presented in Chapter 6 illustrate the level of performance that can be achieved with this particular set of choices. For some aspects of the design, simulation results are presented to illustrate an alternative design choice. However, we in no sense consider all reasonable approaches.

5.1.1. Optimal Designs

Optimal solutions to signal processing problems in which the goals are closely related to those in our wave tracking problem and the class of models is identical to the class of our design models are well known. An example is described in the following paragraphs.

The physiological model (i.e. the system's dynamics) is simply a finite-state Markov chain. The electromagnetic model (i.e. the observation process) is simply a superposition of finite-durational Gaussian signals whose parameters depend on the most recent Markov chain transitions. Furthermore, by simple transformations, the design model can be brought to a form where the Markov chain cycle interval and the

signature sampling interval are equal and where each signature (still Gaussian) lasts exactly one sample.

Our wave tracking problem is concerned with the tracking of annotations. A closely related problem is tracking states of the overall Markov chain. Given an estimate of the state trajectory, it is straightforward to generate a solution to the wave tracking problem, since the annotations are deterministic functions of the state transitions. One approach to specifying a state trajectory estimate is by imposing a minimum probability-of-error optimization criterion.[1]

The combination of a minimum probability-of-error optimization criterion applied to state trajectories and a model in which the most recent transition of a Markov chain determines the observation probability distribution function (pdf) at the next sample (Gaussian pdfs are not required) is a solved problem. Furthermore, the solution, a MAP estimator, has an efficient implementation--the Viterbi Algorithm (Forney, 1973).

We do not pursue such a global optimization approach because, as argued in Section 6.24, the computational burden is prohibitive. However, in spite of the facts that

---

[1] Note the difference between the minimum probability-of-error criterion applied to state trajectories versus the same criterion applied to annotation sequences. An estimator that optimizes the criterion for one case generally does not for the other.

(1) we do not wish to pursue a global optimization approach and

(2) the minimum-probability-of-error state-trajectory problem described above is not the same as our wave tracking problem,

we frequently compute the answer to the minimum-probability-of-error state-trajectory problem in order to have an absolute (rather than a relative) benchmark for our suboptimal algorithms.[2]

5.1.2. Suboptimal Designs

A successful suboptimal design methodology must exploit the structure in the design model. The most pervasive structure in the design model is the subdivision into interacting submodels. Therefore, this thesis is focused on exploiting the interacting submodel structure. The approach is to mirror the submodel structure of the design model in the architecture of the estimator. More concretely, the estimator is partioned into subestimators where each subestimator corresponds to, and is primarily concerned with, a single submodel of the design model and thus a single anatomic subunit of the heart. These estimators interact, and finally their individual results are combined to give a global solution.

Each such subestimator is called a local estimator (LE). Its total a priori knowledge is called the local estimator's model (LE's model). The definition of the information exchanged between LEs is called the intercommunication structure (ICS). Each LE is concerned with a

[2] It is the annotation sequence estimator in which the estimate is the annotations implied by the minimum-probability-of-error state-trajectory estimate that we call the global MAP estimate. Note that it is not MAP for the annotation sequence.

particular submodel of the design model. The additional submodels of the design model are referred to as the remainder model of that LE. Note that if a design model has n distinct submodels, then it also has n distinct remainder models, one for each LE.

There are several intertwined issues to be considered.

(1) Local Estimator Models: What is a local estimator's model? Our philosophy suggests that it includes the corresponding submodel, but it may also include a portion of the LE's remainder model. Each LE's model can have input channels which, when the estimator operates, receive observations and/or receive values generated by the ICS.

(2) Intercommunication Structure: What is the intercommunication structure? The ICS has input channels which, when the estimator operates, receive particular results calculated by particular LEs. The ICS itself conveys these values to the input channels of other LEs.

The reason that aspects of the remainder model may need to be included comes directly from the presence of interactions between the submodel with which an LE is primarily concerned and the remaining submodels. However, as we have indicated, a fundamental premise in our work is that the state of the overall design model is too large to be estimated all at once. Therefore it follows that the LE must use an aggregate version of its remainder model, capturing those aspects of importance in tracking the submodel with which it is primarily concerned. In this case the input channels to the LE contain information

from other LEs concerning interactions impinging on the first LE's submodel and, perhaps, information concerning signature initiations elsewhere in the model. Thus the purpose of the ICS is to provide such coordinating information to each LE from the other LEs.

We now continue with the list of issues.

(3) Local Estimator Algorithm: Given the a priori knowledge of (1) and (2), the observations, and the information provided by the ICS, what does each LE do? That is, what estimation problem does each LE solve? In particular, how does it incorporate both the actual measurement (from the electromagnetic model) and the information provided by the other LEs through the ICS?

(4) Global Estimate Reconstruction: How should the outputs of the individual LEs be combined into a solution of the wave tracking problem?

The focus of this thesis is organizational issues--the local estimator models, the intercommunication structure, and the global estimate reconstruction. The local estimator algorithm is designed by traditional means (i.e. to optimize a traditional performance criterion) but it is limited to the scope of the LE's model. In other words, we are in essence viewing classical estimation algorithms such as the Viterbi algorithm as primitives. Our estimators are constructed by interconnecting such primitives (applied to manageable-size pieces of the overall problem), and the critical issues to be investigated involve the nature of how we interconnect these primitives. The remainder of this subsection provides an overview of the LE's model, the ICS, and the

global estimate reconstruction procedure. The design problem on which the local estimator's algorithm is based is discussed in the next section.

The following paragraphs give an overview of a LE's actions as a first step in describing its model and the ICS. The basic notion is iterative refinement of an initial estimate--i.e. each LE will process data several times, where each successive processing stage uses information provided from previous processing by other LEs. The iterations are referred to as passes. The number of passes for a given LE is fixed but the number need not be the same for all LEs.

Each LE makes an initial pass based only on the ECG and a priori knowledge from the design model.[3] That is, the LEs do not communicate.

Following the initial pass, each LE then makes further passes based on the knowledge used for the first pass together with additional knowledge of the following types:

(1) an a priori description of the nature of the information to be received from the ICS (e.g. knowledge that this information will provide estimates of interactions with other submodels), including measures of its quality, and

---

[3] More precisely, the starting point in the design of all LEs is a design in which there is a nontrivial initial pass for each LE Later steps in the design procedure are a set of rational guidelines for the removal of selected passes and the rearrangement of the ICS, and in these steps the initial pass of some of the LEs is often deleted.

(2) the actual communicated values of the information.

The sources of the information transmitted by the ICS are the previous passes of particular LEs. (Information can be transmitted from the previous pass to the present pass of the same LE. This pathway is lumped into the ICS even though it is intra- not inter-LE). Note that the a priori description of information to be communicated through the ICS is a description of the performance of the prior LE passes which provide the information. The basic restriction on the ICS that we have considered is that a given pass of a given LE may not send information to an earlier pass of any LE.

The following paragraphs give an overview of the LE's model and the ICS. Because of the multiple-pass structure of the LE, there are really multiple models, one model for each pass. As mentioned above, communication between sequential passes of the same LE is treated as a part of the ICS. The following describes a single model for a particular pass of a particular LE. Such a model is called a local estimator pass model (LEPM).

Recall that LEs are in one-to-one correspondence with submodels of the design model. The LEPMs of a LE corresponding to a given submodel need not be provided with the exact submodel and/or the exact remainder model. Rather, each may be given approximations, denoted as the local submodel and local remainder model respectively (as we have argued previously the local remainder model will typically be a drastically aggregated version of the full remainder model).

The situation with respect to the local submodel is simple.

Because the LE is focused on its submodel, it is reasonable to provide each LEPM with the most detailed available a priori information concerning its submodel. Therefore, all estimators described in this thesis have local submodels that are identical to the corresponding submodels of the design model.

As we have discussed previously, in general a non-trivial remainder model is required. The requirement stems from the fact that the local submodel is not complete--descriptions of the impinging interactions and the contributions to the observation from the other submodels in the design model are not included. The representation of these effects are the first and second tasks of the local remainder model. Note that the representation can include information received from the ICS (e.g. concerning times at which interactions have occurred).

An important point to note is that there is valuable information in the knowledge not only of interactions impinging on a particular submodel but also of interactions <u>initiated</u> by the particular submodel (e.g. the contraction of the ventricles in a normal heart provides information about atrial behavior prior to the contraction). Specifically, from the point of view of the LE corresponding to the initiating submodel, the initiated interactions are essentially an observation of the state of its local submodel. Therefore, when available, the ICS communicates such information. The third task of the local remainder model is to represent <u>this</u> information.

The mathematical structure of the LEPMs making up the LE's model has not been discussed. Because the local submodel is identical to the submodel itself, one part of the model has the mathematical structure

familiar from the design model. However, the local remainder model could have a radically different structure. In the work reported in this thesis, this possibility has not been exploited, as a small generalization of the design model's structure has been sufficient to create local remainder models with the desired characteristics.

The generalization is that in addition to channels on which Gaussian signatures are observed, there are also separate channels on which integer-valued random variables, whose pmf depends on the prior state transition, are observed. This generalization, introduced so that the LEs can communicate, is discussed in detail in Chapter 6. The design of specific local remainder models is highly problem-dependent, and discussion is deferred to Sections 5.5 and 5.6 and to Chapter 6.

As described previously, the global reconstruction combines the outputs of the LEs into a single global estimate. All estimators discussed in this thesis use the same method. For each LE there is a distinguished pass, usually its final pass, from which the global reconstruction collects information. The information is a state-trajectory estimate for the local submodel of the LEPM. [Recall that the local submodel of each LEPM of each LE is identical to the submodel (in the design model) to which the LE corresponds]. The collection of these trajectory estimates is a state trajectory estimate for the complete design model. This estimate is then used to determine the annotations which are the solution to the wave tracking problem.

This completes the overview of our approach to suboptimal estimators. Sections 5.2 and 5.3 discuss local estimation and global estimate reconstruction in detail in a general context. We follow this with a

discussion of the reduction of design models to the models used by LEs. This discussion proceeds by example. The examples are drawn from three classes of design models which are introduced in Section 5.4. The local estimator models and the intercommunication structure, which are strongly dependent on the particular design model, are introduced in Sections 5.5 and 5.6. Details and illustrative case studies are given in Chapter 6.

5.2. Local Estimation

In this section we describe the estimation problem solved by each LE. From the previous section, the LE's model for each pass is a particular LEPM, which is a Markov chain with transition-dependent observation pdfs. The final output of each pass of each LE is a state trajectory estimate for the local submodel. Mathematically, this problem is very similar to the solved problem posed in Section 5.1.1. The difference is that here we are focusing on the state of a reduced model rather than the state of the entire design model. Since the size of the LE's model is significantly less than that of the complete design model, the implementation of optimal estimators for these submodels is within reason, and this is essentially what we will do. Following the approach of Section 5.1.1, the minimum probability of error optimization criterion can be imposed on the LEPM state-trajectory estimate, implying the use of a MAP estimator with known implementations (in particular the Viterbi algorithm). While nothing in the previous discussion requires it, all passes of all LEs, not just the final pass, compute the MAP estimate of the LEPM state trajectory. The purposes to which the results of the early passes are put is complex (since they are the input to the ICS

which in turn provides information to future passes of the LEs) and dependent on the design model. The ICS, along with the structure of the LEPMs, is described in Sections 5.5 and 5.6 and in Chapter 6.

Thus, at each pass of each LE, an optimal signal processing problem is solved. That optimal signal processing problem is defined by the choice of the LEPM and the ICS. All algorithms for MAP estimation, of course, yield identical results. However, the choice of algorithm for a particular pass of a particular LE may influence the information provided to the ICS from that pass, because certain quantities are natural byproducts of particular algorithms, but not of others. Therefore, the choice of algorithm may, by influencing the information available from the ICS, affect the choice of LEPM in subsequent passes.

The estimators described in this thesis all employ the Viterbi Algorithm (VA) (Forney, 1973) implementation of the MAP estimator. Three qualitative characteristics of the VA that are strongly reflected in the LEs are described in the following paragraphs.

The VA does not naturally retain a "second best" trajectory. Specifically, if there are M states in the Markov chain then there are M retained trajectories. Each retained trajectory is the best choice given that it must terminate on a given state. However, the "second best", trajectory terminating on the same state is not retained. The cause for this behavior in the Viterbi algorithm is that no future information could possibly make the "second best" trajectory more likely than the "first best", and the justification of this stems directly from the notion of state in a Markov chain. However, in our case, each LE in essence has only a "local state" and thus it is not true that future

information (available at the next pass) could not reverse the positions of first and second best. Thus there are reasons for expecting that one could improve performance by exchanging a second best trajectory or a trajectory that is most likely among those that differ "radically" (in the sense of the interaction sequences they specify) from the most likely. For example, such information might be of value in avoiding catastrophic error propagation from one LE to the next. While such local processing and information exchange might be of value, we have not considered them but rather have relied solely on Viterbi algorithms for the local processing. Thus, since "second best" trajectories are not available, all algorithms described in this thesis exchange information extracted from the local MAP trajectory only.

The VA does not naturally provide a measure of confidence in its estimate. Therefore, when exchanging information, LEs send only the value, not the value and an on-line confidence measure. However, as we will see in our case studies the local remainder models may contain a priori information concerning the accuracy of communicated information.

The VA requires an observation model in which observations at different times are conditionally independent given their corresponding state transitions. Therefore, multicycle signatures are not permitted. As discussed previously (Section 5.1.1) this is not a fundamental problem since the state space can be augmented in many ways to overcome it. Examples of this are discussed in Chapter 6.

5.3. Global Estimate Reconstruction

This section discusses the creation of the final annotation trajectory estimate from the results of the individual LEs. Recall from Section 5.1 that the global estimate reconstruction provides an annotation trajectory estimate by collecting each submodel state-trajectory estimate from the corresponding LE and then applying deterministic operations to the state trajectory to get an annotation trajectory. Thus all the work of reconciling local information is done by communication between the LEs and nothing is left for the global estimate reconstruction step.

An important implication of this reconstruction method is the following. Based on the examples in Section 3.8 and in Chapter 6, it is generally the case that many global state trajectories are impossible (e.g. in a normal heart in which each ventricular contraction is caused by an atrial contraction, the P and R waves cannot occur simultaneously). However, there is no reason to expect that locally produced estimates of parts of the global state will be compatible in the sense that once adjoined, the resulting global trajectory has nonzero probability. This is the issue of consistency.

There are three levels of inconsistency.

(1) No inconsistency: the global state-trajectory estimate produced by combining local state-trajectory estimates has positive probability.

(2) Partial inconsistency: each estimated submodel trajectory has positive probability for one or more possible trajectories of interactions impinging on that submodel. However, for at least one submodel there is an inconsistency in that the state-trajectory estimate for this local submodel is not possible for the set of interactions specified by the actual state-trajectory estimates of the other submodels.

(3) Total inconsistency: one or more estimated submodel trajectories has zero probability for all trajectories of interactions impinging on that submodel.

In the estimators studied in this thesis, total inconsistency can never occur because in all LEs the local submodel is identical to the submodel in the design model. However, because the global reconstruction procedure is trivial, it does not exclude or correct any possible partial inconsistency. Therefore, whether the estimate is consistent or only partially consistent depends entirely on the LEs, specifically on communication between the LEs during intermediate stages of the estimation procedure.

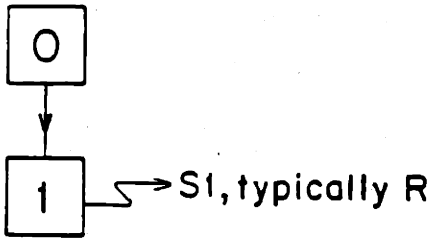
The LE intercommunication described in this thesis is not sufficiently sophisticated to guarantee a consistent global estimate. Furthermore, we do not understand how to achieve such communication without essentially using a global optimization solution. Therefore, we are forced to accept estimates that may be partially inconsistent. A partially inconsistent (or even totally inconsistent) estimate can still have long consistent segments. Furthermore, note that an inconsistent state trajectory estimate may still imply an wave tracking estimate that

has positive probability, i.e. the estimate is consistent from the point of view of the annotations rather than the state. Even if it is inconsistent, the performance of such an estimator may still be quite good, where by "performance" we refer to the probabilistic measures (false positive/false negative, etc.) described in Chapter 4. Since such measures are our fundamental concern we will focus on them rather than on the issue of consistency.

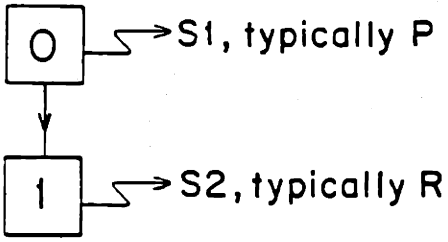
5.4. Design Model Classes

The discussion of the construction of the LE structure starting from a given design model proceeds by example. This section introduces and motivates (from an ECG point of view) the three classes of design model from which the examples used in this and the next chapter are drawn. The three classes also represent a logical sequence of models with increasing complexity which successively introduce new issues to be considered in the estimation process.

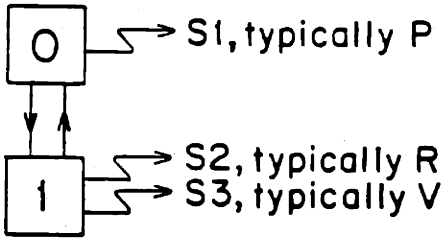
Figure 5.1 shows the three classes of design models. The block diagram notation used in Figure 5.1 was used previously in Chapter 3. In the sequel, these classes are referred to as DM1, DM2, and DM3. Because the focus of this thesis is on organizational issues, no design models with only one submodel are discussed. Furthermore, design models with three or more submodels are also not discussed because all phenomena with which this thesis is concerned can be illustrated in less complex situations. The following discussion provides an ECG motivation for each of these classes.



(a) Design-Model Class DM1



(b) Design-Model Class DM2



(c) Design-Model Class DM3

Figure 5.1 Design-Model Classes.

An ECG motivation for DM1 is normal rhythm with P waves disregarded. Furthermore, by including a second signature type in C1, simple models of some ventricular arrhythmias, such as ventricular oscillation, fall into this class. A motivation for DM2 is normal rhythm complete

with P waves. Furthermore, some AV-nodal rhythms, such as first degree AV block, fall into this class. Similar to DM1, the addition of a second signature from C1 allows the modeling of some ventricular arrhythmias. Finally, a motivation for DM3 is any ventricular arrhythmia which can propagate retrograde.

5.5. The LEs' Models and the ICS

The design of an estimator's architecture requires the transformation of a design model into a set of LEs and an ICS, and a specification of the set of LEPMs which make up each LE's model.

The basic approach is to provide a generic set of components from which local remainder models can be built and a generic set of interconnections for the ICS. The designer, keeping in mind the properties of the design model, can propose an initial estimator architecture using these building blocks. The initial estimator can then be refined by experiment.

In spite of the greater simplicity of the DM1 and DM2 design-model classes, the discussion begins with the DM3 design-model class, with the further assumption that both signatures have equal SNRs. The reason for starting with this case is that it is the simplest case in which the submodels of the design model are symmetric. Therefore it is the simplest case in which one can reasonably seek estimator architectures with symmetry between the LEs. After sketching the estimator architecture in the symmetric case, the specializations to the asymmetric cases will be sketched.

Focus on DM3. As described in Section 5.1.2, each LE makes an initial pass (pass 0) based only on a priori information and the ECG. That is, no portion of the ICS conveys information to a pass 0. Due to the symmetry, only LE 0 is described. A block diagram of the complete LE 0 pass 0 LEPM is given in Figure 5.2.

As shown in Figure 5.2 and discussed in Section 5.1.2, the local submodel for the LE 0 pass 0 LEPM is submodel 0 from the design model. By this choice the state space of C0, the number of transition probability matrices (tpms) for C0 and their values, and the number of interactions P0 initiates with C0 (which equals the number of tpms for C0) are all specified.

The local remainder model for the LE 0 pass 0 LEPM must therefore model

(1) any interactions impinging on submodel 0 and

(2) any contribution of submodel 1 in the design model to the ECG (i.e. signature S1).

The ECG contributions due to submodel 1 are modeled (that is, the LE is made aware of the other signature type) in order to keep the LE from confusing occurrences of the signature S0, with which it is primarily concerned, with occurrences of S1. As shown in Figure 5.2, aspects (1) and (2) are dealt with by different chains: the impinging interactions are modeled by chain P0 and the contribution to the ECG is modeled by chain P4. The state spaces and tpms for P0 and P4 and the interactions initiated by P0 are described later.

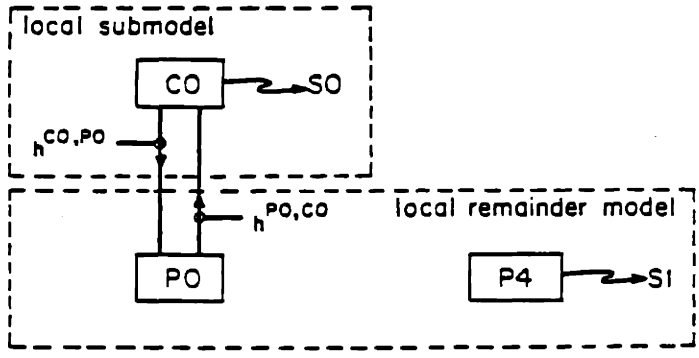


Figure 5.2 Design Model DM3, LEPM for LE 0 Initialization Pass.

As described in Section 5.1.2, the initial pass of each LE is followed by a sequence of refinement passes. For a given LE, all LEPMs for refinement passes have the same structure. However, they have different parameters and are attached to different parts of the ICS. As was done for the initial pass, only the LE 0 LEPM and its portion of the ICS are described for the generic refinement pass. The configuration for LE 1 follows from symmetry.

A block diagram of a refinement pass LEPM for LE 0 is given in Figure 5.3. As in the initial pass LEPM, the local submodel is submodel 0 from the design model. However, the local remainder model is substantially more complex than the local remainder model of the initial pass LEPM.

In a refinement pass, the LE receives a posteriori information through the ICS. As in the initial pass, the local remainder model must model interactions received by submodel 0 and the contribution of submodel 1 in the design model to the observation (i.e. signature S1). Information concerning both of these aspects is available through the ICS. For example, the result of the initial pass of the LE concerned with submodel 1 implies (hopefully accurate) estimates of times at which submodel 1 initiated interactions with submodel 0 and times at which the

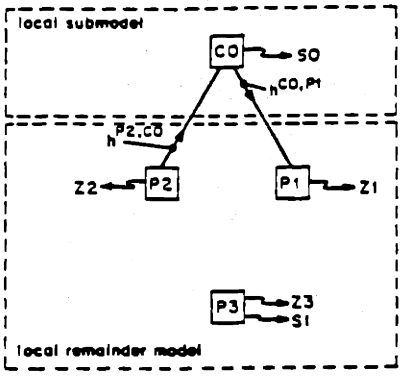


Figure 5.3 Design Model DM3, LEPM for LE 0 Refinement Pass.

signature S1 was initiated. Furthermore, there is also information available through the ICS concerning the interactions initiated by submodel 0. For example, the result of the initial pass for the LE concerned with submodel 0 produces estimates of the times at which submodel 0 initiated interactions with submodel 1. Note that in the initial pass of LE 0, such information on 0-to-1 interactions was not available and thus the LEPM did not account for it. In a refinement pass such information is available and therefore the local remainder model must also include a model of this information. Finally, the LE is also provided with a priori information and with the observations.

In summary, for each refinement pass, the ICS provides

(1) an estimate of the interactions initiated by submodel 0, augmented with an estimate of further information as described below;

(2) an estimate of the interactions received by submodel 0, again augmented; and

(3) an estimate of the times at which the signature S1 was initiated.

The total information in (1) or (2) is referred to as an augmented interaction. The signal in (3) is referred to as a binary annotation

(i.e. a yes or no decision at each time corresponding to whether signature S1 was or was not initiated at that time).

The following paragraphs describe the source of the augmented interaction estimate and the nature of the augmenting information.

While one can imagine very complex ways to extract the information in (1) and (2) from prior passes of multiple LEs, we have always chosen to use a single pass of a single LE, though the pass and LE may differ for (1) in comparison to (2). When possible, the interaction information is taken from the prior pass of the LE responsible for the submodel which initiates the interaction. Therefore, (1) comes from LE 0 while (2) comes from LE 1. The selection of this source for the interaction information was motivated by the method of calculating the interaction estimate. Specifically, the interaction estimate is deterministically calculated from the state estimate for the portion of the LEPM which corresponds to the initiating submodel in the design model. Therefore, the interaction estimates are (presumably) most reliable when the portion of the LEPM involved is the local submodel since it is only in this situation that the chain involved is carried over without severe aggregation from the design model.

The augmenting information referred to in both (1) and (2) is the state of the local submodel in the LEPM of the LE pass providing the interaction information. Therefore, in (1) the augmenting information is the state of the C0 chain in the prior pass of LE 0 while in (2) it is the state of the C1 chain in the prior pass of LE 1. Again, this portion of the LEPM was chosen because it is the only portion carried over unchanged from the design model. A more detailed motivation and

discussion of why this augmentation is performed is presented in Chapter 6 in the course of the case studies.

If one were to implement the full version of our approach--i.e. every LE is implemented for every pass--then it is always possible to take the interaction information from the prior pass of the LE that is responsible for the submodel that initiates the interaction. As we will see, typically there are several LE passes that are eliminated from the design, as they are either redundant or of inconsequential value to overall performance. In these cases, it becomes necessary to determine alternative sources for the interaction information in subsequent passes. In the following paragraphs we describe two examples. These ideas are discussed further in Chapter 6 in which we describe specific designs in great detail.

In all passes of all LEs for the symmetric DM3 case described here, the interaction information can always be taken from the prior pass of the LE responsible for the submodel initiating the interaction. In this case, the state information with which the interaction information is augmented includes the interaction information (as the latter is a deterministic function of the former).

In other situations, the interaction information must be obtained from a different source. For example, consider a DM1 class design model for normal rhythm. Submodel 0 (submodel 1) corresponds to the atria (ventricles). The R wave initiated by the ventricles is modeled (signature S1 from submodel 1) while the P wave is not modeled.

Focus on LE 0 pass 1. Because the local submodel initiates no

signatures, LE 0 pass 1 requires a local remainder model, which conveys information from the ICS concerning the interactions submodel 0 initiates with submodel 1. This interaction information cannot come from the prior pass of the LE that is responsible for the submodel that initiates the interaction (i.e. pass 0 of the atrial LE, LE 0), because LE 0 pass 0 is of no value at all since its LEPM has no observations. (Recall that the P wave is not modeled). Therefore, the interaction estimates must be taken from LE 1 pass 0.

Return now briefly to the general case. When an alternative LE must be used as the source of interaction information, as in this example, the interaction information is still deterministically calculated from the portion of the LEPM which corresponds to the initiating submodel in the design model. Because a different LE is used, this portion of the LEPM is no longer the local submodel but rather is now a part of the local remainder model, and therefore is always highly aggregated. The augmenting information is still the state estimate for the local submodel. Therefore, the interaction and augmenting information are <u>not</u> redundant.

Continuing with the example, consider LE 1 pass 0. The LEPM for this pass requires a remainder model, which is an aggregated model of submodel 0 (corresponding to the atria), in order to provide a mechanism to excite chain C1 (corresponding to the ventricles). The 0-to-1 interaction estimates, which are taken from this pass, are deterministicly computed from the state trajectory estimate through the highly aggregated version of submodel 0. Because submodel 1 is affected by the interactions only when its state occupies a particular state (the

resting state) we are more likely to believe the estimate of 0-to-1 interactions derived from the highly aggregated version of submodel 0 in the local remainder model of LE 1 pass 0 if the estimated excitation attempt occurs at a time when the state of chain C1 is in the resting state. This motivates the communication of both the estimated interaction and the estimated state of chain C1, that is, the communication of the augmented interactions defined previously.

In a similar manner, the binary annotation indicating that annotation S1 did or did not occur is taken from the prior pass of the LE responsible for the submodel initiating S1. Therefore, in (3) above, the signal is taken from the prior pass of LE 1. In the design models considered in this thesis, this is always possible.

The information received through the ICS can always be modeled as a noisy observation of a state transition in some subprocess state space of the LEPM. These non-ECG observations are called pseudo-observations. All of the pseudo-observations are modeled as arising from the local remainder model. In this thesis, they have all been discrete valued but that is not crucial.

Since the LE receives three types of information from the ICS there are correspondingly three pseudo-observations labeled Z1, Z2, and Z3 in Figure 5.3. Pseudo-observation Z1 models the information concerning interactions initiated by submodel 0. This is reflected in the C0-to-P1 direction of the interactions in the LEPM. Likewise, pseudo-observation Z2 models the information concerning interactions impinging on submodel 0. Finally, pseudo-observation Z3 models the information concerning occurrences of S1.

Since Z1, Z2, and Z3 are estimates from the prior pass, a description of the statistics of Z1, Z2, or Z3 is a description (from a particular point of view) of the performance of the prior pass. For example, the tpms on the chain of P1 and the pmfs for Z1 given the last state transition of P1 are a description of the performance of the prior pass with respect to estimating the interactions initiated by submodel 0. The subprocess state spaces and tpms for P1, P2, and P3 and the interactions initiated by P2 are described later.

The estimators described in this thesis all include two refinement passes following the initialization pass. Therefore, for DM3, the complete estimator up to the global reconstruction step is as shown in Figure 5.4. As before, straight solid lines indicate interactions between submodels in the LEPMs and wavy solid lines indicate signatures initiated by the corresponding submodel. The ICS is shown in dashed lines. Note the symmetry in the ICS between passes 0 and 1 and passes 1 and 2. The leading superscripts indicate the identity of the LE and the pass number. For example, ${}^{1,0}\hat{x}_n^{C1}$ is the estimate of the state of chain C1 at time n computed by LE 1 pass 0. Otherwise the h and x notation is as in Section 4.1 with carets added to indicate that these are estimates. The notation Si,$\overline{\text{Si}}$ where i is 0 or 1 is used to emphasize the binary values of the binary annotations. Finally note that the Pi boxes with the same name (e.g. the four instances of P2) are each different chains. The common name refers to the common method used to create these chains from other information and reflects the common role (with respect to the local submodel) played by members of each class of chain. These five methods (for P0, ... , P4) are described later.

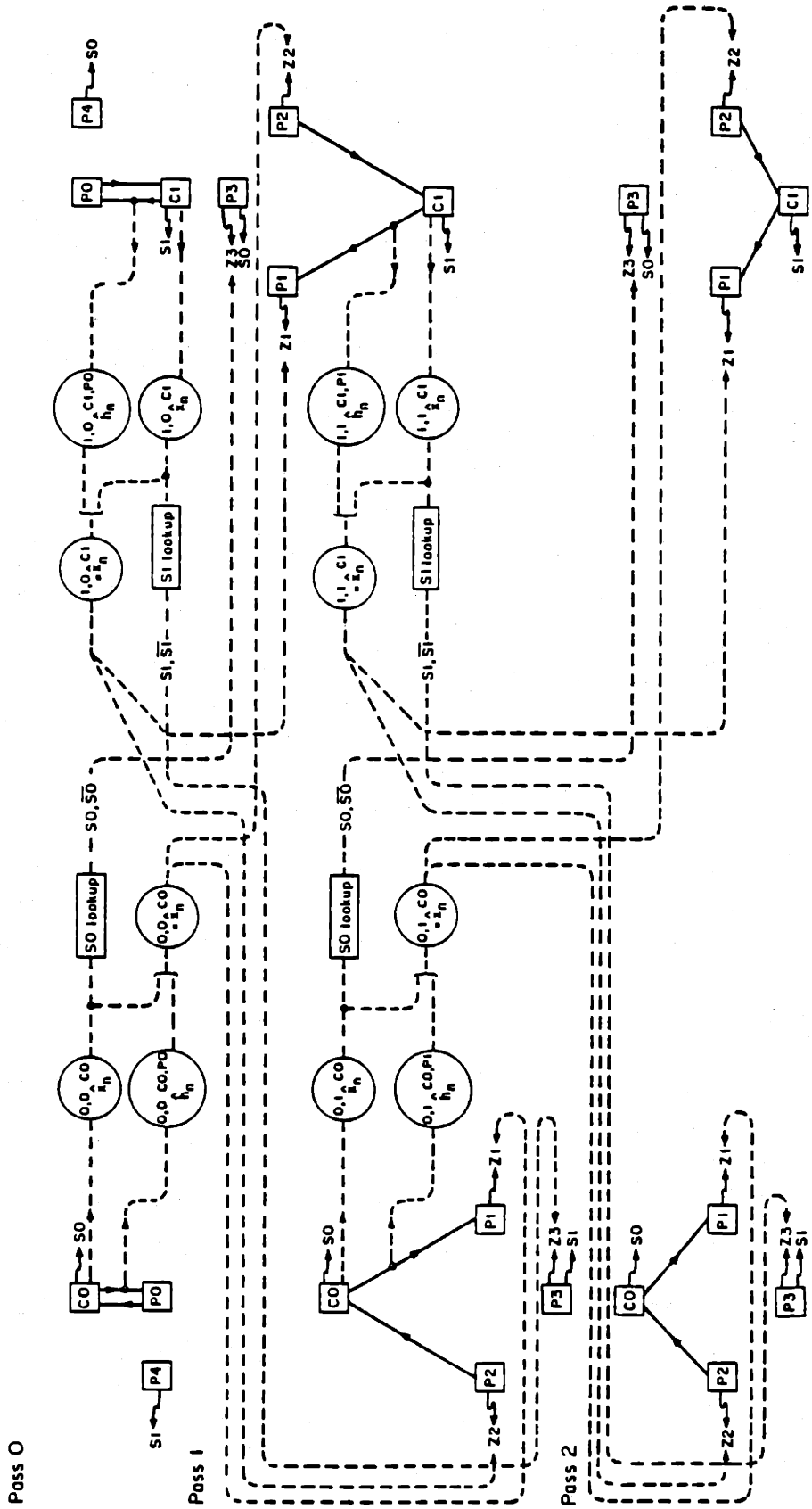


Figure 5.4 DM3 Estimator, Block Diagram.

As described in Section 5.3, the global reconstruction step is extremely simple. First a global state-trajectory estimate is formed by taking the state trajectory estimate for chain C0 from LE 0 pass 2 and for chain C1 from LE 1 pass 2. Then this global state-trajectory estimate is deterministicly transformed into a solution to the wave tracking problem by simply attaching appropriate annotations to the corresponding transitions in the state trajectory.

This concludes the sketch of the architecture of the DM3 estimator. In Chapter 6 we present case studies for particular DM3 design models in which the SNR of the S1 signature is much lower than the SNR of the S2 signature. While we do not describe this specialization of the DM3 estimator, the corresponding specialization of the DM2 estimator is described in the following section, and based on that discussion the DM3 case follows analogously.

5.6. Specializations of the DM3 Architecture to DM2 and to DM1

The estimator architecture proposed for the DM3 design-model class can be specialized to the DM2, and then further to the DM1, design-model classes.

In the DM2 case, the interactions are unidirectional. Therefore certain elements of the LEPMs simply drop out. The resulting estimator up to the global reconstruction step is shown in Figure 5.5. The global reconstruction step is unchanged.

However, even further simplifications, as shown in Figure 5.6, can be made when other aspects of the ECG problem are considered. For

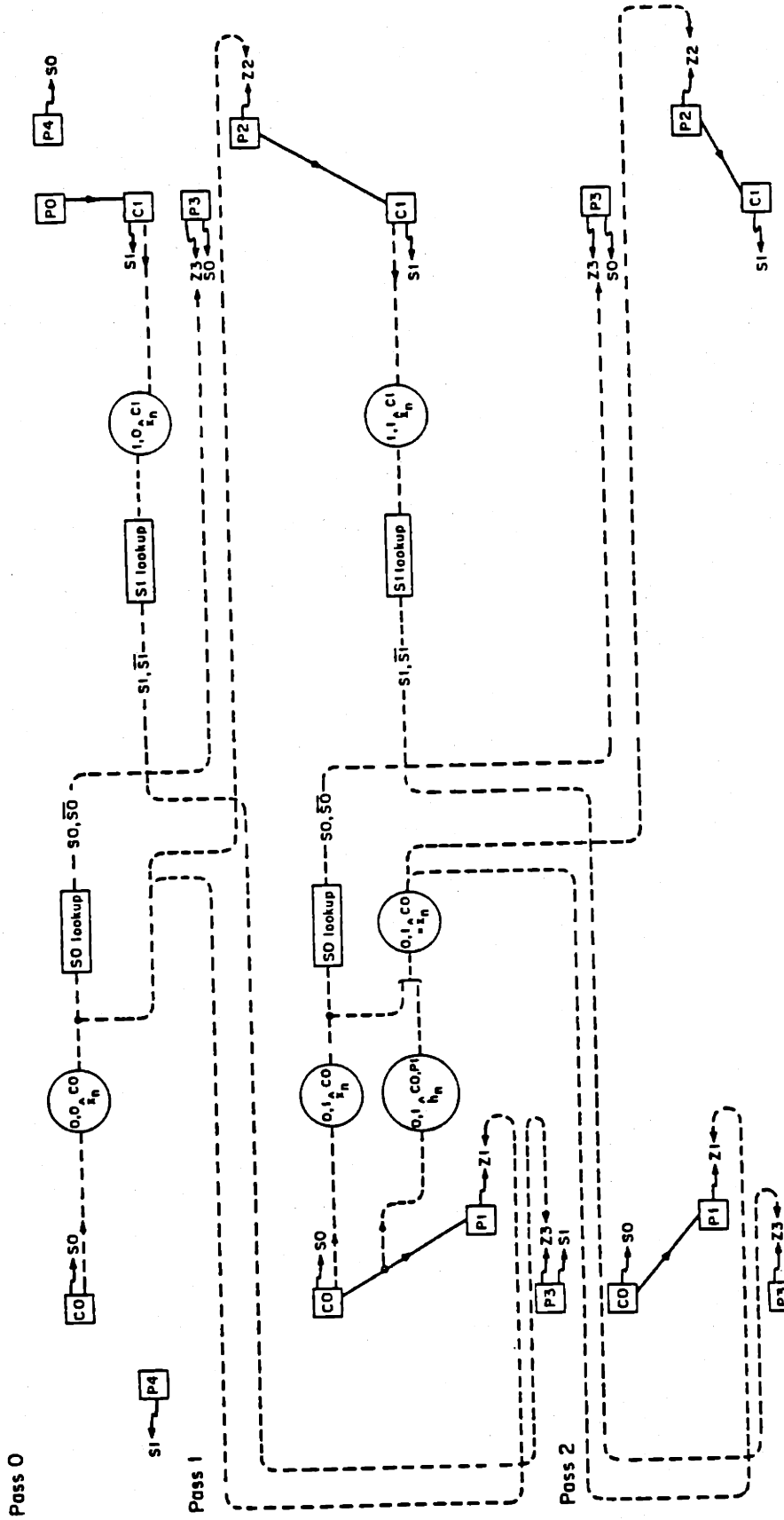


Figure 5.5 DM2 Estimator, Block Diagram.

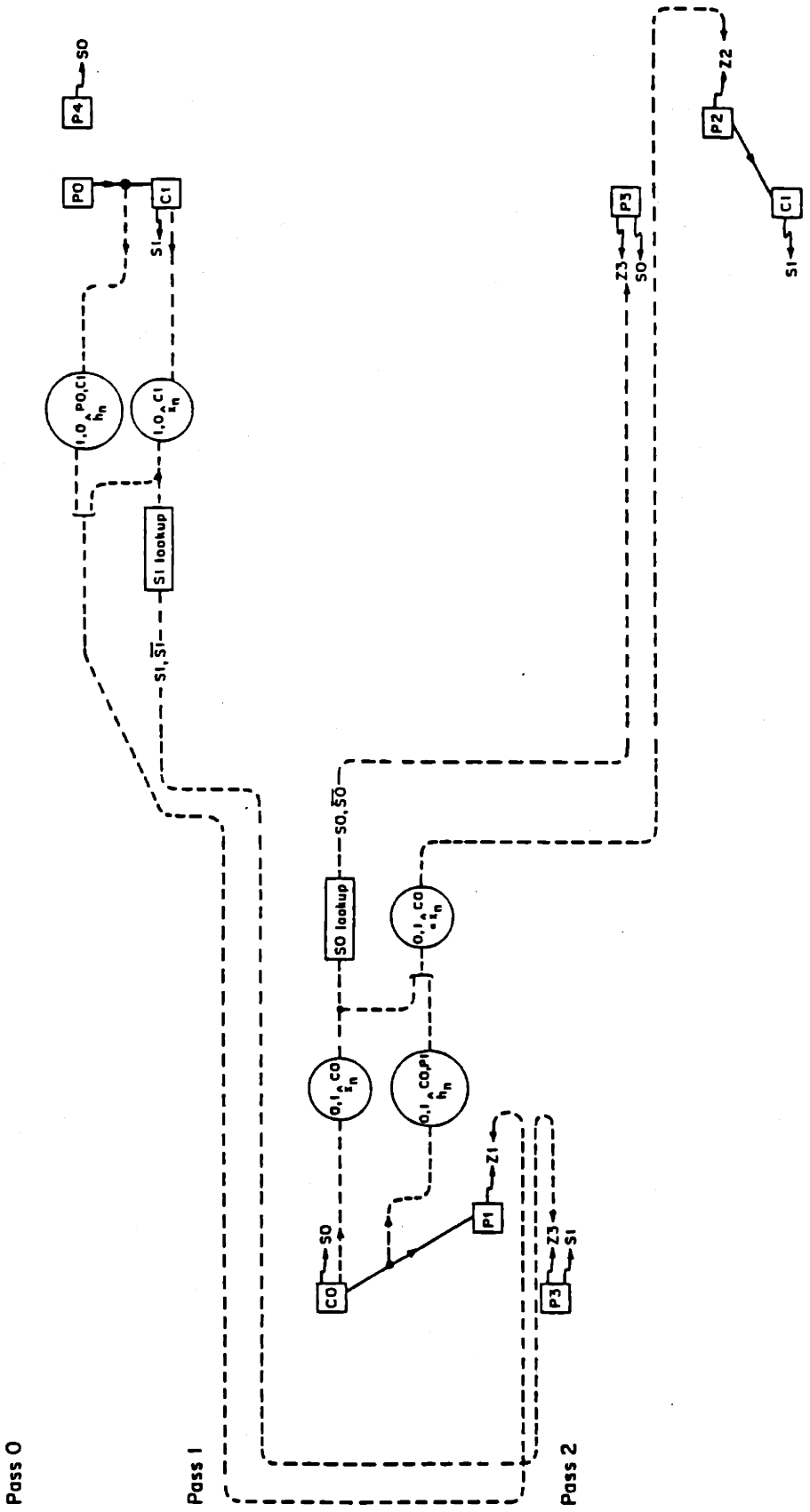


Figure 5.6 DM2 Estimator, Block Diagram.
For this estimator, the mean of the S0 signature is assumed to be substantially less than the mean of the S1 signature.

example, the S0 signature is the atrial activity (P wave) and the S1 signature is the ventricular activity (R wave). Therefore the magnitude of the S0 signature is much lower than the magnitude of the S1 signature. Considering this case, it is clear that LE 0 pass 0 is essentially useless since in that LEPM the low magnitude S0 signature is modeled in detail, but the model for the high magnitude S1 signature is very crude. Specifically, there are no causality constraints relating the S0 and S1 signatures and therefore no reliable mechanism keeping estimated S0 (S1) signatures near true S0 (S1) signatures. Rather, the estimator, which (intuitively) attempts to explain energy in the signal as waves, may use estimated occurrences of the S0 signature to account for the random portion of true S1 signature occurrences. Therefore LE 0 pass 0 should be deleted.

The deletion of LE 0 pass 0 forces changes in the ICS, as was noted in the previous section with respect to DM1 design models. Specifically, in LE 0 pass 1, the information concerning the interactions initiated by submodel 0 can no longer be obtained from LE 0 pass 0. Therefore, it must be obtained from LE 1 pass 0. Thus the transmitted information, formerly the C0-to-C1 interaction estimate[4] $^{0,0}\hat{h}_n^{C0,C1}$ augmented by the C0 state estimate $^{0,0}\hat{x}_n^{C0}$, is now $^{0,1}\hat{h}_n^{P0,C1}$ and $^{0,1}\hat{x}_n^{C1}$. Note that with this new interaction information, the augmenting information $^{0,1}\hat{x}_n^{C1}$ does not include the interaction estimate $^{0,1}\hat{h}_n^{P0,C1}$.

---

[4] Note that chain C1 which receives these interactions does not occur in the LEPM for LE 0 pass 0 in Figure 5.5. However, the interactions initiated by C0 and impinging on C1 can still be computed because they only depend on the state of chain C0.

LE 0 pass 2 is also essentially useless. LE 1 pass 1 essentially replicates LE 1 pass 0 (since the deletion of LE 0 pass 0 leaves them with essentially the same inputs). Therefore, LE 0 pass 1 and LE 0 pass 2 also have essentially the same inputs. For this reason, LE 0 pass 1 and 2 should be combined and, since they have identical structure, the combined pass would be the same as LE 0 pass 1. (Here the phrase "essentially the same inputs" applied to two LE passes means that the ECG observations and information transmitted from other LEs via the ICS are approximately the same for both). Therefore LE 0 pass 2 should be deleted. The deletion of LE 0 pass 2 forces no changes in the ICS and also allows the deletion of LE 1 pass 1 (since the results of LE 1 pass 1 are no longer used). Therefore the DM2 estimator up to the global reconstruction step for this SNR situation has the architecture shown in Figure 5.6. The global reconstruction step is slightly changed. Since LE 0 pass 2 no longer exists, the state-trajectory estimate for C0 is taken from LE 0 pass 1 instead.

Specializing to the DM1 case, the SNR of signature S0 goes to zero. Therefore, chain P4 of LE 1 pass 0, chain P3 of LE 0 pass 1, and chain P3 of LE 1 pass 2 drop out; and chain C0 of LE 0 pass 1 no longer initiates a signature. Note that LE 0 pass 1 has only pseudo-observations. The architecture of the resulting estimator up to the global reconstruction step is shown in Figure 5.7. The global reconstruction step is unchanged from the prior case.

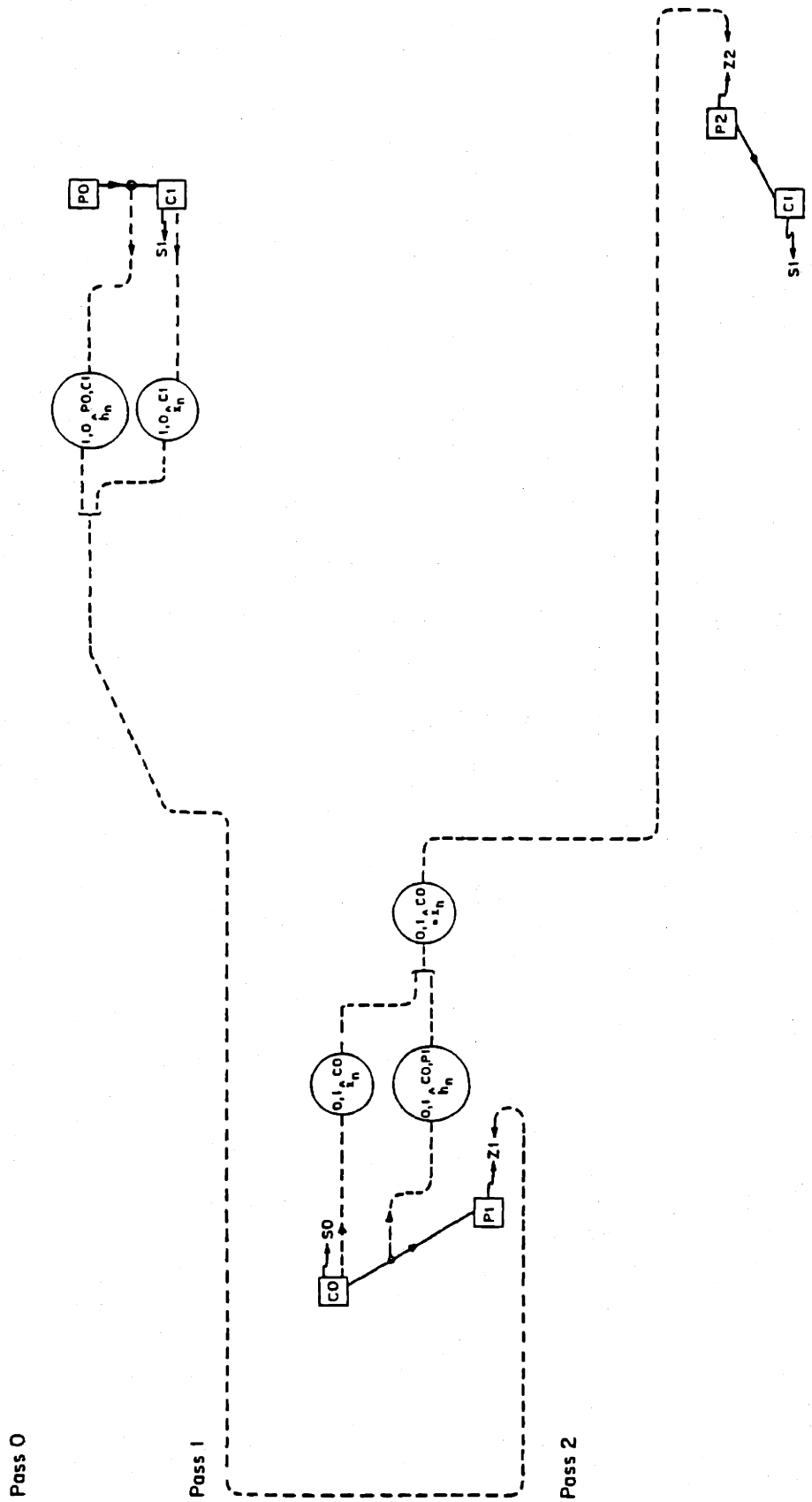


Figure 5.7 DM1 Estimator, Block Diagram.

Section 5.6.

5.7. Summary of Chapter 5

This chapter has proposed a general approach to the design of estimator architectures, and a specific architecture for design-model class DM3 in the case where the SNRs of the two signatures are approximately equal. This architecture was then specialized to the DM2 case, the DM2 case with unequal signature SNRs (motivated by the ECG problem), and the DM1 case. The discussion began with the complex, symmetric DM3 case because of the symmetry of the design model. The symmetry was reproduced in the architecture of the estimator. In the asymmetric DM3, DM2 and DM1 cases, the estimators also lack symmetry.

The latter discussions have shown how a complex architecture can be simplified using judiciously chosen rules of thumb. While this seems to be true for a number of design models relevant to the ECG problem, and may be applicable in other areas, this demonstration by example is as formal as the general design process will be taken in this thesis.

6. Signal Processing III: Implementation and Case Studies

Turning from the general considerations of Chapter 5, this chapter illustrates their application in specific ECG case studies. The structure of the generic P0-P4 chains and the organization of the ICS is discussed in detail and exemplified in a series of case studies, beginning with design models drawn from the simplest class of design models: DM1. Each of the P0-P4 chains is introduced when the increasing complexity of the design model first requires it, along with the related ICS issues. Finally, the last sections of the chapter consider several miscellaneous issues.

6.1. The P0 Class of Chains

The creation of a P0 chain is described in the context of the initial pass of LE 1. This is the case occurring in the DM1 estimator of Figure 6.1.[1] The P0 chain models the interactions impinging on submodel 1 in the case when no information is available through the ICS. As such, it is only used in the initial pass of a LE.

The subprocess state space of the P0 chain was taken to be the coarsest possible aggregation of the subprocess state space of submodel 0 consistent with the full range of possible interactions initiated by submodel 0 and impinging on submodel 1. Thus, all states of submodel 0 in $U^{0,1,i}$ are coalesced into state i of the P0 subprocess state space. The notion of a map T from the states in the submodel 0

[1] In each section of Chapter 6, all figures appear at the end of the section.

subprocess state space to states in the P0 subprocess state space is useful. The fact that the map represents an aggregation is reflected in the fact that it is onto but not one-to-one. This thesis has only considered design models in which the interaction submodel 0 initiates with submodel 1 takes two values, so that the subprocess state space of P0 has two states. Intermediate levels of aggregation, which would require additional states in the P0 subprocess state space, are possible but have not been considered in this thesis.

The interactions which P0 initiates with C1 follow from the definition of the P0 subprocess state space. Let $x^{P0}(n)$ be the state in the P0 subprocess state space and $h^{P0,C1}(n)$ be the interaction P0 initiates with C1. For submodel 0 the interactions $h^{0,1}(n)$ are defined by

$$x^{0}(n) \in U^{0,1,h^{0,1}(n)}.$$

Correspondingly, the interactions $h^{P0,C1}(n)$ transmitted by P0 are defined by

$$x^{P0}(n) = h^{P0,C1}(n).$$

That is, in submodel 0, the same interaction can occur when the submodel 0 subprocess state occupies a number of different states. Thus we use the values of the interactions $h^{0,1}(n)$ to index a <u>partition</u> of the subprocess state space of submodel 0, specifically the sets $U^{0,1,i}$ as the interaction i ranges over all possible values. In the aggregated P0 subprocess state space, each state corresponds to a <u>different</u> interaction. Finally, P0 does not initiate interactions with any other chain

in the LEPM.

The interactions received by P0 are identical to the interactions received by submodel 0. Two methods for choosing a tpm (or a set of tpms) for the subprocess state space of P0 have been considered. They are both based on the notion of preserving selected statistics between the aggregated and unaggregated chains, that is, under the action of the map T. More concretely, the value of some statistic which can be described in terms of the $U^{0,1,i}$ sets in the submodel 0 subprocess state space is calculated. Then the tpm on the subprocess state space of P0 is chosen so that the corresponding statistic has the same value.

The first method we have considered preserves the aggregated version of

$$\lim_{n\to\infty} Pr(x^0(n)\in U^{0,1,j} | x^0(n-1)\in U^{0,1,i}).$$

That is, the tpm on the subprocess state space of P0 is chosen so that

$$Pr(x^{P0}(n)=j | x^{P0}(n-1)=i) = \lim_{n\to\infty} Pr(x^0(n)\in U^{0,1,j} | x^0(n-1)\in U^{0,1,i}).$$

For each i and j, the left hand side of this equation is exactly the i,j element of the tpm for the P0 subprocess.

The second method preserves the aggregated version of

$$\lim_{n\to\infty} \Pr(x^0(n) \in U^{0,1,i})$$

in a similar way. Preservation of this statistic does not provide sufficient equations to determine a unique tpm for the P0 subprocess. Therefore, we have also imposed the side condition that all rows of the tpm are equal. This choice was made primarily because it is simple and provided reasonable performance, which is our chief criterion. Like all of the aggregation ideas, this particular aspect is an area open for future research. Finally, note that in both cases, if submodel 0 receives interactions from submodel 1 (e.g. in DM3), the aggregation is performed independently for each of the tpms of submodel 0.

For all design models discussed in this thesis, preservation of statistic (1) always yields a tpm with the structure shown in Figure 6.2. The reason is that all submodels discussed in this thesis which are transformed into P0 type chains initiate two interactions. Therefore, there are two states in the P0 chain. Furthermore, one of the interactions can never occur twice in a row. Therefore, there is an on-diagonal zero in the tpm. In the tpm of Figure 6.2, there is a single parameter p. Method (1) described above will produce a specific value of p for each impinging interaction. In addition, in our estimation case studies, we have also considered p as an independent design parameter in order to determine how its value can influence estimator performance. Thus our third method is optimization of the parameter p.

Finally, a fourth method is needed for cases discussed in later sections where the first two are clearly not appropriate. The problem

is in the steady-state assumption built into the definition of the statistics. Specifically, consider a submodel which both receives and initiates interactions. Furthermore, assume that the received interactions are constrained (by their initiating submodel) such that a certain tpm is applied rarely and only for brief intervals. In this case the assumption that the submodel reaches steady state under this tpm (an implicit assumption in <u>both</u> of our analytical approaches to computing tpms for P0) is clearly inappropriate. These situations are considered on a case-by-case basis when they arise.

Since the P0 chain does not initiate any signatures, its structure is completely determined at this point.

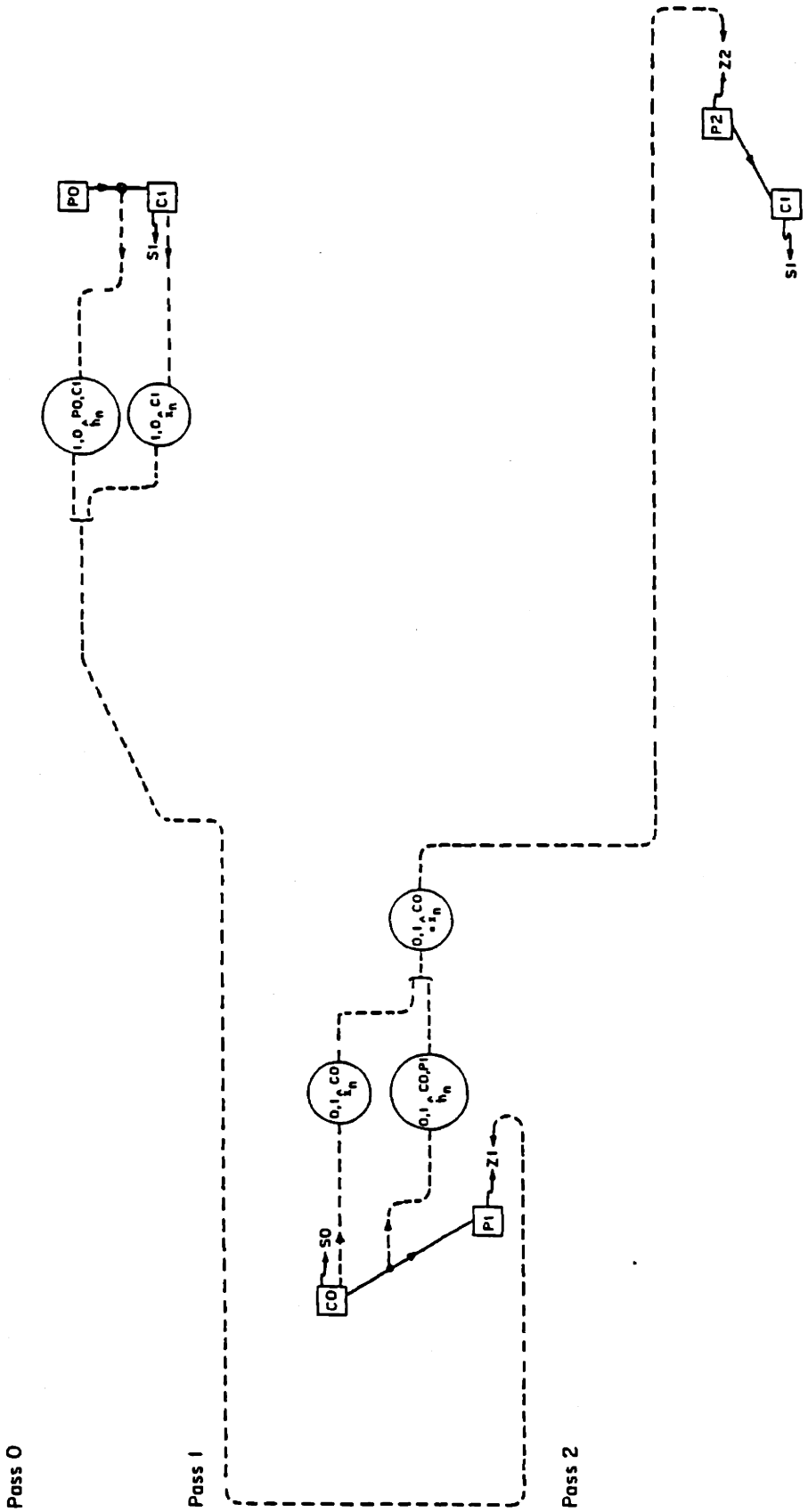


Figure 6.1 DM1 Estimator, Block Diagram.
This figure duplicates Figure 5.7.

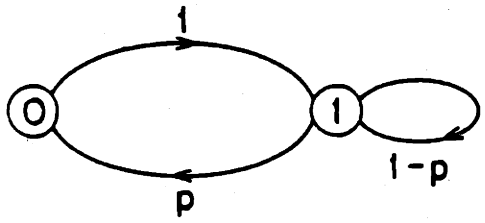


Figure 6.2 P0 Tpm When Preserving Statistic (1).

6.2. The P1 Class of Chains

The creation of a P1 chain is discussed in the context of pass 1 of LE 0. This is the case occurring in the DM1 estimator of Figure 6.1. From the initial pass of LE 1 it is possible to extract information relevant to estimating the interactions submodel 0 initiates with submodel 1. (As discussed previously, in more complex estimators such as the DM3 estimator of Figure 5.4, this estimate may come from the initial pass of LE 0). The purpose of the P1 chain is to use this information, received through the ICS, to assist pass 1 of LE 0.

The information provided by the ICS is an estimate of the augmented interactions described in Section 5.5. Furthermore, since the information provided by the ICS is based on a prior pass through the observed data, there is no reason that the ICS must provide this information in a causal manner. That is, at time n, the ICS may present the estimated augmented interaction at time n+m, m a constant integer not restricted to negative values. The implications of m nonzero will be discussed

later.

The P1 chain has a single pseudo-observation denoted Z1 which is obtained from the information received from the ICS. There are no Gaussian observations initiated by P1. The P1 chain receives interactions from the C0 chain but it does not initiate any interactions.

The following formalism applies to pseudo-observations initiated by the P1, P2, or P3 type of chains. At each time sample, the pseudo-observation is a single, scalar, integer-valued random variable whose pmf is determined by the previous state transition in the subprocess state space of the chain which initiates the observation. The pmf is given explicitly by specifying its value for each permissible value of the integer-valued random variable.

The tpms of P1 and the pmfs of the pseudo-observation Z1 together model the information provided by the ICS. The information is used as an observation of the interactions submodel 0 initiates with submodel 1. Since this information is just an estimate created by a previous pass, the P1 chain is actually modeling the performance, in a specific sense, of all the LE passes which contribute to the estimate.

An important point to note is the following. The information we are trying to capture consists of estimates of the interactions initiated by submodel 0. Since these interactions are completely determined by the state of submodel 0, one might think that the interaction estimates from the prior pass could be successfully modeled as <u>direct</u> observations on the C0 chain. This, however, neglects the major aspects of the <u>dynamics</u> of the estimation procedure used to obtain these

estimates. In particular, without including some additional memory (which we accomplish by introducing the P1 chain), we cannot model phase shifts between actual and estimated interactions. By capturing such phase shifts we obtain a far more accurate and useful characterization of the performance of the prior pass.

As we have just indicated, the purpose of the P1 chain is to introduce memory into the model. To do this, we proceed as follows. The interactions which C0 initiates with P1 are identical to the interactions which submodel 0 initiates with submode 1 in the design model. The subprocess state space and tpms of P1 are chosen so that P1 acts as a shift register for these interactions. That is, at each transition P1 throws out knowledge of the oldest interaction and adds in knowledge of the newest interaction. The length of the shift register plus one is the duration of the model's memory. The additional one unit of memory is due to the fact that the observation pmf for the pseudo-observation Z1 is dependent on the most recent state <u>transition</u>, not on the present state.

Consistent with our previous notation, let Z1(n) be the random variable corresponding to the information received through the ICS at time n and let $Z1_n$ be its realization. Let K be the length of the shift register. In light of the previous description of the states of P1, the observation pmf at time n has the form

$$\Pr(Z1(n)=Z1_n \mid h_n^{0,1}, h_{n-1}^{0,1}, \ldots, h_{n-K}^{0,1}).$$

This pmf can be directly interpreted as a quantification of the dynamic

performance of the prior pass. In all the work reported on in this thesis, this pmf is evaluated via simulation.

The one remaining piece of information we must specify is the nature of Z1(n). The most obvious choice is to let Z1(n) be the estimate of $h^{0,1}(n)$ from the prior pass. However, as we have indicated, we have the ability to use the ICS information in a noncausal manner. Furthermore, since the estimation algorithm we will use (the Viterbi algorithm) is a smoothing algorithm, it is natural to characterize our performance in estimating $h^{0,1}(n)$ as a function of the actual interactions at times <u>before</u> <u>and</u> <u>after</u> n. Thus, we may more generally wish to take Z1(n) as the estimate of $h^{0,1}(n-J)$ for some $J>0$. In the work reported in this thesis, K is two and the estimated augmented interaction is delayed by one (J=1) so it is centered in the window of three interactions represented by each transition in the P1 chain.

6.3. The P2 Class of Chains

The creation of a P2 chain is discussed in the context of pass 2 of LE 1. This is the case occurring in the DM1 estimator of Figure 6.1. From pass 1 of LE 0 it is possible to extract information relevant to estimating the interactions which submodel 0 initiates with submodel 1. The purpose of the P2 chain is to use this information, received through the ICS, to assist pass 2 of LE 1.

The P2 chain draws on ideas already seen in the P0 and P1 chain cases. Specifically, like P0, the purpose of P2 is to model received interactions. However, unlike P0, P2 has an observation. Like P1, the observation is a pseudo-observation supplied by the ICS. The

information received is also analogous to the P1 case. Finally, like P1, the P2 chain acts as a shift register memory, storing a sequence of 0-to-1 interactions. However, unlike P1, the interactions P2 stores are generated internally.

Exactly as in the P1 case, the information provided by the ICS is an estimate of the augmented interactions (possibly shifted in time as with Z1(n)) described in Section 5.5. The only difference is in the source of the estimate. The pseudo-observation (described above) from the P2 chain is denoted Z2. The P2 chain has no Gaussian observations. It initiates interactions with C1 but does not receive interactions from any submodel.

The single tpm of P2 and the pmf of the pseudo-observation Z2 model the information provided by the ICS. The information is used as an observation of the interactions submodel 0 initiates with submodel 1. Since the information is just an estimate created by a previous pass, the P2 chain (like the P1 chain) is actually a model of the performance, in a specific sense, of all the LE passes which contribute to the estimate.

In the P1 case, as discussed previously, if a zero-memory model of the estimates received through the ICS were sufficient, then Z1 could be directly initiated by C0. In the P2 case, even if a zero-memory model is sufficient, it never makes sense to initiate Z2 directly by C1. The reason is that the 0-to-1 interactions that the Z2 pseudo-observation represents are <u>not</u> functions of the state transitions in the C1 subprocess state space. Rather, in the P2 case, the natural zero-memory model is a model with a time-varying tpm for C1 corresponding to the estimated

0-to-1 interactions. That is, the tpm at time n would be determined by the estimated augmented interaction at time n. This is directly motivated by how the true interaction at time n determines the tpm of submodel 1 at time n--i.e. truth is simply replaced by an estimate. However, since the estimates from prior passes typically contain phase shifts, a zero-memory model typically yields poor performance.

The purpose of the P2 chain is to introduce memory into the model. Specifically, each state of the P2 subprocess state space represents a particular state of a shift register memory. The contents of the shift register memory are the interactions which submodel 0 initiates with submodel 1. This is quite similar to the P1 case. However, in the P1 case the interactions were received from an external source and simply stored. In the P2 case there is no external source and the chain must generate the interaction sequences based on a priori information. Let K be the length of the shift register. Let $b_0^i,\ldots,b_{K-1}^i$ be the shift register contents represented by the state labeled "i" in P2. The i,j element $q_{i,j}$ of the tpm for P2 is calculated as

$$q_{i,j} = \lim_{n\to\infty} \Pr(h^{0,1}(n+1-k)=b_{K-1-k}^j,\ k=0,\ldots,K-1 \mid$$

$$h^{0,1}(n-k)=b_{K-1-k}^i,\ k=0,\ldots,K-1).$$

Note that for the transition from i to j to have positive probability, it must satisfy $b_{k-1}^j=b_k^i$, $k\in\{1,\ldots,K-1\}$. In this case,

$$q_{i,j} = \lim_{n\to\infty} \Pr(h^{0,1}(n+1)=b^{j}_{K-1} \,|\, h^{0,1}(n-k)=b^{i}_{K-1-k},\ k=0,\ldots,K-1).$$

Autonomous transitions under this tpm generate a sequence of interactions based on a priori information.

Consider next the way in which P2 initiates interactions with C1. To allow for noncausality in the use of information provided by the ICS we proceed as follows. Let K^* be a distinguished element (usually centered) in the shift register memory, so $K^* \in \{0,\ldots,K-1\}$. When the state of P2 is i, the interaction P2 initiates with C1 is defined to be $b^{i}_{K^*}$. Therefore, the sets $U^{P2,C1,k}$ in the subprocess state space of P2, which define the P2-to-C1 interactions, are

$$U^{P2,C1,k} = \{i \in S^{P2} : b^{i}_{K^*} = k\}.$$

As in previous notation, let Z2(n) be the random variable corresponding to the information received through the ICS at time n and let $Z2_n$ be its realization, an integer. In light of the previous description of the subprocess state space of P2, the observation pmf (for the transition from state i to state j) has the form

$$\Pr(Z2(n)=Z2_n \,|\, j,i) = \Pr(Z2(n)=Z2_n \,|\, b^{j}_{K-1}, b^{i}_{K-1}, b^{i}_{K-2}, \ldots, b^{i}_{0}).$$

(Note that on the right-hand side we have used the fact that $b^{j}_{k-1}=b^{i}_{k}$, $k=1,\ldots,K-1$ in order for the transition to have positive probability). Note, as in the P1 case, that while the length of the shift register is

K, the duration of the model's memory is K+1. In all the work reported on in this thesis, this set of pmfs is calculated by simulation.

Because of the freedom to choose K^*, the ability to have the ICS present time-shifted estimated augmented interactions is less important in the P2 than in the P1 case (i.e. by using $K^* < K-1$ we have already included some noncausality in our model). In the work reported on in this thesis, K is two, K^* is one, and the augmented interaction is delayed by one (J=1). Therefore, as in the P1 case, the observation is centered in a window of three interactions represented by transitions in the P2 chain.

6.4. Numerical Results for Design-Model Class DM1--Introduction

In this and the following sections we present and discuss numerical results for several design models belonging to class DM1. A historical approach is followed in describing numerical results for all design-model classes. This approach was chosen because it clearly presents the reasons behind the evolution of the estimators.

In the DM1 estimator of Figure 6.1 there are five design choices that require experimental justification. They are

(1) P0, the a priori model of interactions impinging on submodel 1 (used in LE 1 pass 0);

(2) P1, the a posteriori model of interactions initiated by submodel 0 (used in LE 0 pass 1);

(3) P2, the a posteriori model of interactions impinging on submodel 1 (used in LE 1 pass 2);

(4) the precise definition of the augmented interaction information transmitted by the ICS from LE 1 pass 0 to LE 0 pass 1; and

(5) the precise definition of the augmented interaction information transmitted by the ICS from LE 0 pass 1 to LE 1 pass 2.

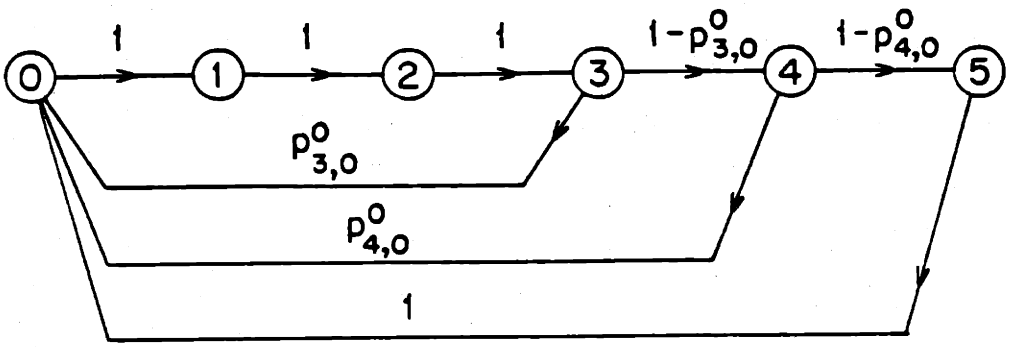
Historically, (5) evolved based on DM2-class design models and is discussed in a later section. In the following sections, (1)-(4) are discussed. The discussion of (1) is placed after the discussion of (2)-(4) because our discussion of (1) focuses on its affect on overall global estimator performance. We also include discussions of the variation of performance (for a fixed estimator structure) as the design model changes and of some robustness issues.

The experiments for DM1-class design models are based on two design models. Figure 6.3 shows a very simple DM1-class design model for normal rhythm. Like many of the design models to be described, this design model has a number of free parameters. Specifically, there are seven free parameters: $p^0_{3,0}$, $p^0_{4,0}$, $p^1_{0,2}$, $p^1_{3,0}$, m_R, σ_R, and σ_{obs}. The probabilities $p^0_{3,0}$ and $p^0_{4,0}$ determine the pmf on the period between SA-nodal depolarizations, the probability $p^1_{0,2}$ sets the AV-junctional delay-time pmf, and the probability $p^1_{3,0}$ determines the ventricular refractory-period pmf. The mean m_R and standard deviation σ_R describe the R wave, which is one sample long. The standard deviation σ_{obs} describes the observation noise (which has zero mean). The sets of values for these seven parameters that were used in our experiments are listed in

Table 6.1.

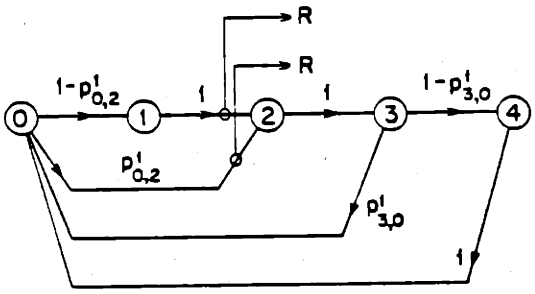
Figure 6.4 shows a second DM1-class design model for normal rhythm. This design model is essentially the same as the design model of Figure 6.3 except that there are approximately twice as many Markov chain cycles between occurrences of an R wave signature. The increased temporal resolution makes it possible to consider phase shifts and multiple Markov-chain cycle signatures in a more realistic situation. For example, in this design model the R wave signature is two Markov-chain cycles long. Only one set of parameters are considered for this design model, and they are indicated in Figure 6.4.

Submodel 0:

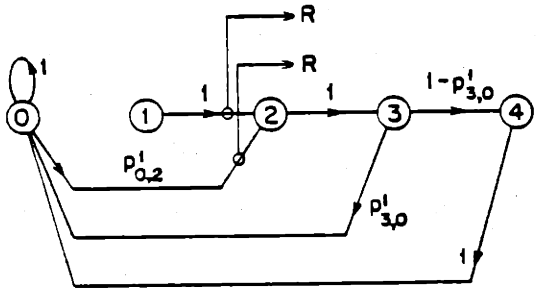


Submodel 1:

if $x^0=0$:



if $x^0\neq 0$:



The Markov chain interval is 160 msec.

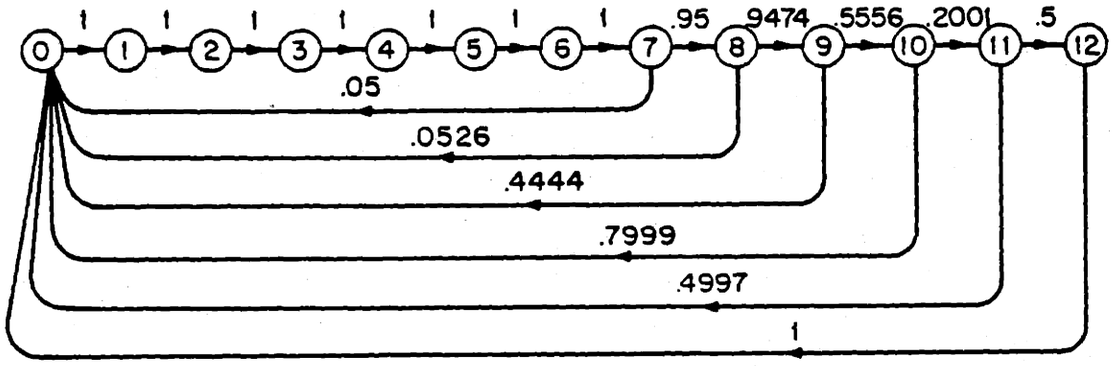
Submodel 0 has no Gaussian observations. There is a Gaussian observation from submodel 1, modeling the R wave. The signature sampling interval is 160 msec. The R wave signature is one sample long, with mean m_R and standard deviation σ_R. The Gaussian observation noise has mean 0 and standard deviation σ_{obs}.

Figure 6.3 A Normal Rhythm Design Model.

		Model Number							
		1	2	3	4	5	6	7	8
parameter	$p^0_{3,0}$	.25	.25	.25	.25	.25	.25	.25	.25
	$p^0_{4,0}$	.6667	.6667	.6667	.6667	.6667	.6667	.6667	.6667
	$p^1_{0,2}$	.1	.1	.1	.1	.1	0	0	0
	$p^1_{3,0}$	.5	.5	.5	.5	.5	.5	.5	.5
	m_R	10	5	4	3	2	5	3	2
	σ_R	$\sqrt{.5}$	$\sqrt{.5}$	$\sqrt{.5}$	$\sqrt{.5}$	$\sqrt{.5}$	$\sqrt{.5}$	$\sqrt{.5}$	$\sqrt{.5}$
	σ_{obs}	1	1	1	1	1	1	1	1

Table 6.1 Parameters for the Design Model of Figure 6.3.

Submodel 0:



Submodel 1:

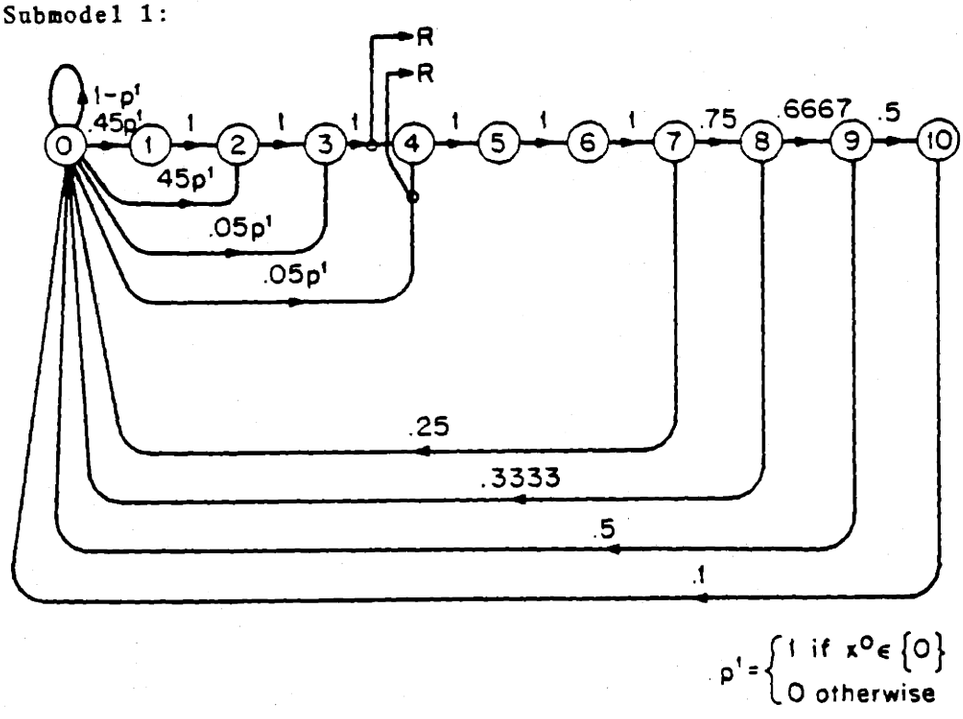


Figure 6.4 A Normal Rhythm Design Model with Twice the Temporal Resolution of the Design Model of Figure 6.3.

Markov chain cycle interval = signature sampling interval = 1 (normalized time).

R: 2 samples long, each sample having mean 2 and standard deviation $\sqrt{.5}$.

Observation noise: mean 0 and standard deviation 1.

Figure 6.4 Continued.

6.5. The P1 Chain

In the estimator of Figure 6.1, the purpose of LE 1 pass 0 is to calculate estimates of the augmented interactions impinging on submodel 1, i.e. to estimate when the atria (submodel 0) have excited the ventricles (submodel 1). The estimates are the realization of the pseudo-observations in the LE 0 pass 1 LEPM. In our initial attempt to provide useful information to LE 0 pass 1, we used estimates of the 0-to-1 interactions alone. As discussed earlier in this chapter, we found that in order to incorporate these pass 0 interaction estimates into pass 1 it was necessary to model the dynamics of the pass 0 estimation procedure. Specifically, we consistently observed phase-shifted interaction estimates in LE 1 pass 0.

To illustrate the nature of this phase-shift phenomenon, we describe in some detail a particular example. Consider the design model of Figure 6.3 with the parameter values of Table 6.1 column 1. A particular realization of the stochastic process defined by Figure 6.3 and Table 6.1 column 1 is shown in Figure 6.5. We now describe the LE 1 pass 0 estimator for this example. Recall that LE 1 pass 0 has only the

original Gaussian measurements (i.e. there are no pseudo-observations) and the LEPM consists of the chains P0 and C1. Such a LEPM is depicted in Figure 6.6. Chain C1 is carried over directly from the design model. The methodology behind the specification of chain P0 has been generally described previously, and more specifics will be discussed in the sequel.

The performance (with respect to the interactions impinging on submodel 1) of the LE 1 pass 0 of Figure 6.6 on the simulated data from the design model of Figure 6.3 and Table 6.1 column 1 is described in Table 6.2. This table is a matrix each element of which is the number of times a certain event occurred during the simulation. The events are all 2^4 possible ordered quadruples of estimated interactions at time n ($^{1,0}\hat{h}_n^{P0,C1}$, 2 possible choices) and truth interaction at times n-1, n, and n+1 ($h_{n-1}^{0,1}$, $h_n^{0,1}$, $h_{n+1}^{0,1}$; 2^3 possible choices). Recall from the design model definition that

$$h_n^{0,1} = \begin{cases} 0 \text{ if submodel 0 attempts to excite submodel 1} \\ 1 \text{ otherwise} \end{cases}$$

and likewise for $^{1,0}\hat{h}_n^{P0,C1}$. Note that certain truth interaction triplets (namely those with more than one zero interaction) have zero probability of occurring. These appear in the matrix as all zero columns. In subsequent tables such columns are suppressed. In columns (1,1,0), (1,0,1), and (0,1,1) the sum of the elements is not exactly equal from column to column due to edge effects at the beginning and end of the simulation. The choice of a triplet of truth interactions was made because the chains are so small that longer intervals are difficult to

interpret. The desire to include both lead and lag phase errors motivated the use of a symmetric triplet. Tables of this type occur frequently in the sequel.

The (0,(1,1,0)) entry (i.e. the entry with 136 elements) is a particular type of phase error. The estimator claimed that submodel 0 attempted to excite submodel 1 one Markov chain cycle before the true attempt at excitation. Note that the corresponding lag (rather than lead) error, which is (0,(0,1,1)) never occurred. This indicates that the estimator has a tendency to produce an early estimate of excitation time.

The (1,(1,0,1)) entry (i.e. the entry with 338 elements) is due to two mechanisms. Of the 338 occurrences, 136 are due to the type of phase error described in the previous paragraph. More specifically, at the time sample following each (0,(1,1,0)) error, a (1,(1,0,1)) error is guaranteed to occur. This is an example of entrainment of errors and illustrates the need to consider errors in a dynamic fashion. The remaining 338-136 = 202 errors are misclassification errors. That is, the estimator erroneously claimed that submodel 0 did not attempt to excite submodel 1. In terms of the design model, such errors are to be expected because the R wave can actually be dropped. That is, if x^1, the state of submodel 1, takes on one of the values {1,...,4} at the time of the excitation attempt, the R wave is dropped. Such events are poorly observable and therefore the estimates are often in error.

This point is worth some amplification. Specifically, recall that in DM1 there are no P waves. Thus, if an R wave has been dropped, we have only the presence of a long R-R interval and the a priori knowledge

of submodel 0 timing to provide any indication that an excitation attempt was made and failed. However, we have only made use of submodel 0 timing in a very crude way (via P0) and thus, one might expect there to be numerous errors at this stage. Many of these may be corrected in LE 0 pass 1 where we make use of detailed submodel 0 timing. Note that in any event, the presence of such errors--even if they are not corrected in a later pass--is of no real significance. If P waves are present--as in DM2--these errors do take on more significance, but we would also expect better performance since LE 0 pass 1 can use both a priori submodel 0 timing information and the Gaussian observations to estimate times at which P waves occur and submodel 1 excitations are attempted (when one of these occurs, so does the other).

While pure misclassification errors can be modeled by a zero-memory model, the phase-shift errors require a more complex model with memory. Such a model is provided by chains of type P1 which can model arbitrary phase errors up to any pre-specified lead and lag. All estimators described in this thesis use P1 type models.

Recall from Section 6.2 that P1 chains are parameterized by the length of the shift register. In the work reported here, the shift register is two truth interactions long. Therefore, the pseudo-observation pmf, which depends on state transitions, is conditioned on a triplet of truth interactions. This decision is based on four factors.

(1) Given that both leading and lagging errors are to be modeled, a triplet is the smallest possible choice.

(2) The chains considered in this thesis are rarely large enough for longer sequences of truth interactions to make sense. That is, longer sequences would represent a substantial fraction of the R-R interval.

(3) The size of the P1 chain grows as the length of the shift register increases. Therefore, the computational complexity of the Viterbi algorithm, and hence that of the <u>implementation</u> of the estimator, increases. Thus there is strong motivation to minimize the amount of memory in the P1 chain.

(4) The number of conditional observation pmfs for the pseudo-observations of chain P1, namely Pr(pseudo-observation|interactions in the shift register), increases as larger shift registers are considered. Furthermore, the pmfs contain probabilities of progressively rarer events. Since we calculate these probabilities by simulation, this implies that longer simulations are required in order to achieve the same level of confidence. Therefore the computation required to <u>design</u> the estimator increases.

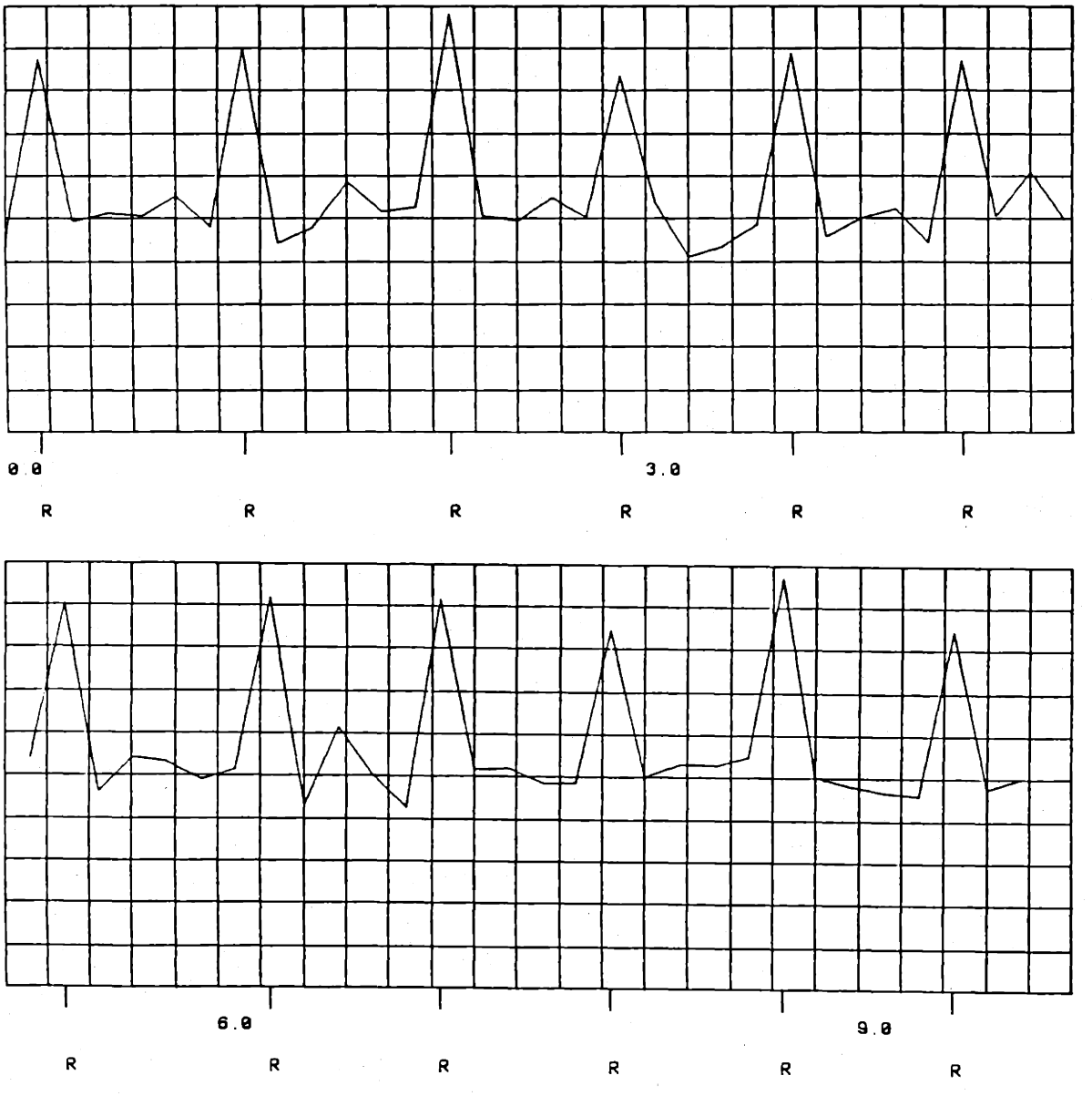
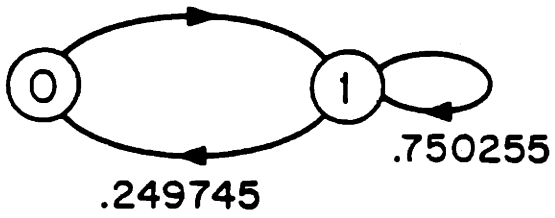


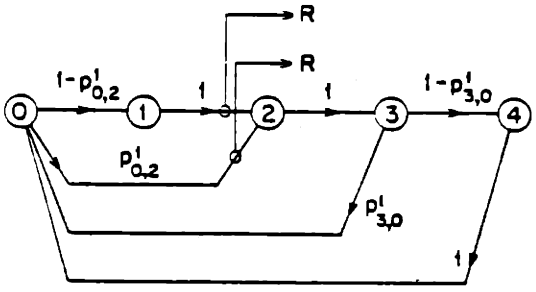
Figure 6.5 A Realization of the Stochastic Process Defined in Figure 6.3 and Table 6.1 Column 1.

Chain P0:

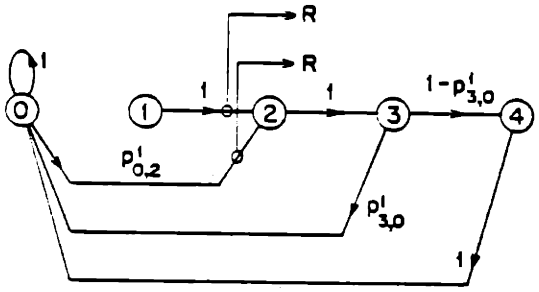


Chain C1:

if $x^{P0}=0$:



if $x^{P0}\neq 0$:



The Markov chain interval is 160 msec.

Chain P0 has no Gaussian observations. There is a Gaussian observation from chain C1, modeling the R wave. The signature sampling interval is 160 msec. The R wave is one sample long, with mean 10 and standard deviation $\sqrt{5}$. The Gaussian observation noise has mean 0 and standard deviation 1.

There are no pseudo-observations from either chain.

Figure 6.6 LE 1 Pass 0 LEPM for the Design Model of Figure 6.3 and Table 6.1 Column 1.

	truth interaction triplet $(h_{n-1}^{0,1}, h_n^{0,1}, h_{n+1}^{0,1})$							
	(0,0,0)	(1,0,0)	(0,1,0)	(1,1,0)	(0,0,1)	(1,0,1)	(0,1,1)	(1,1,1)
estimated interaction $1,0\hat{h}_n^{P0,C1}$ 0	0	0	0	136	0	1654	0	0
1	0	0	0	1856	0	338	1993	4019

Table 6.2 Statistics on Interaction Estimates From LE 1 Pass 0.

6.6. The P2 Chain

In the estimator of Figure 6.1, LE 0 pass 1 has two purposes:

(1) it provides the final state-trajectory estimate for submodel 0 and

(2) it provides estimates of the augmented interactions impinging on submodel 1.

The estimates of the augmented interactions are provided as the realization of the pseudo-observation for LE 1 pass 2. As in the evolution of the modeling of communicated information that led to the P1 chain, our initial efforts involved estimates of interactions only, that is without any augmenting information. In order to create a LEPM for LE 1 pass 2 that allowed it to make effective use of the unaugmented interaction information, we were led to the development of the P2 chain.

As in the P1 case, the reason for using such a complex observation model in the LE 1 pass 2 LEPM is that calculations on simulated data showed a large number of phase-shifted interaction estimates from LE 0 pass 1. To illustrate this, consider again the design model and particular realization used in the discussion of the P1 case. This design model is described in Figure 6.3 and Table 6.1 column 1. The LE 1

pass 0 estimator was discussed previously and is shown in Figure 6.6. The LE 0 pass 1 estimator results from applying the definition of the P1 chain to the design model of Figure 6.3 and Table 6.1 column 1, with pseudo-observation observation pmfs derived from Table 6.2. The resulting LEPM for the LE 0 pass 1 estimator is shown in Figure 6.7.

Several aspects of the LEPM shown in Figure 6.7 require brief comments:

(1) In chain P1, the self loop on state 0 when $x_n^{C0}=0$ is irrelevant (except during an initial transient) because, due to chain C0 constraints, it is impossible to have the state of chain P1 in state 0 and simultaneously receive an excitation attempt from chain C0. In order to create a valid tpm, some definition for transitions out of state 0 had to be made and a self transition was the simpliest choice.

(2) For transitions which have positive transition probabilities but which can never occur (i.e. 0→0 and 1→0 in chain P1), the pseudo-observation observation pmf is arbitrarily chosen to be uniform.

(3) For transitions where the simulation statistics indicate that one or the other possible value of the pseudo-observation can never occur, a zero is not placed in the pseudo-observation observation pmf. Rather, the zero is replaced by a small quantity (in this instance .001).

The conventions on how to define a row of a tpm which will never be used (except perhaps for an initial transient) (i.e. (1) above), the

implications of (1) for the choice of pseudo-observation observation pmfs (i.e. (2) above), and transitions whose measured statistics are singular (i.e. (3) above) have evolved in the course of the thesis. The LEPM in Figure 6.7 is a result of using the final form of the rules. Simulation results presented in this section, however, use an earlier form of these rules. The differences among the various rules result in at most a difference in an initial transient.

In Table 6.3 we display the performance of LE 0 pass 1 in estimating interactions impinging on submodel 1. Note that to create Table 6.3, we have simulated the concatenation of two passes: LE 1 pass 0 and LE 0 pass 1. Table 6.3 has exactly the same format as Table 6.2. The error entries are (0,(1,1,0)), (0,(0,1,1)), (0,(1,1,1)), and (1,(1,0,1)). Based on these results, the need for an observation model with memory to model phase-shifted errors is clear. Such a model is provided by chains of type P2. In the work reported in this thesis, balanced triplets of truth interactions are always used. The reasons for this choice are the same as the reasons for the corresponding choice in chains of type P1.

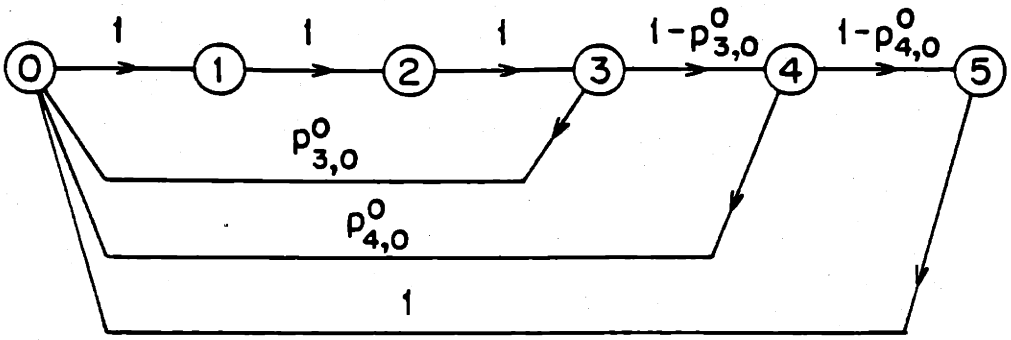
Unlike the P1 case, we have explored an alternative observation model (that is, an alternative to the P2 chain) for the estimated impinging interactions on submodel 1 in LE 1 pass 2. The alternative (called the zero-memory model) is to assume that the estimates are exact. The estimates are used deterministically to switch between the various submodel 1 tpms. This arrangement was discussed in Section 6.3.

For the comparison of the P2 and the zero-memory models in LE 1 pass 2, consider again the design model of Figure 6.3 and Table 6.1

column 1. Both estimators have identical LE 1 pass 0 and LE 0 pass 1 LEPMs which were shown previously. For the estimator using the P2 model, the LEPM for LE 1 pass 2 is shown in Figure 6.8. It is based on the definition of the P2 chain applied to the design model of Figure 6.3 and Table 6.1 column 1, with pseudo-observation observation pmfs derived from Table 6.3. For the zero-memory model, the LE 1 pass 2 LEPM is sub-model 1 of the design model (Figure 6.3 and Table 6.1 column 1) where the choice of tpm is determined by the estimated interactions communicated by the ICS.

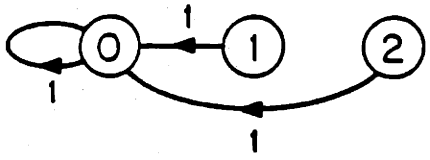
Table 6.4 shows statistics on the misclassification performance when both estimators are applied to the same realization used previously. Note that the signature is only one sample long, so only zero phase-shift matches are allowed. In each table, fp (fn) stands for false positive (false negative) while cs (rs) stands for column (row) sum--the sum of all the elements in that column (row). Furthermore, fpR (fnR) are the false positive (negative) rates for R waves when misclassifications are not allowed. Here, with only one wave type, misclassifications are not possible. The numbers in parentheses are the estimated fractional standard deviations discussed in Section 4.3.5. Finally, teR is the total error rate for R waves, again when misclassifications are not allowed. The superiority of the estimator using the P2 model is obvious.

Chain C0:

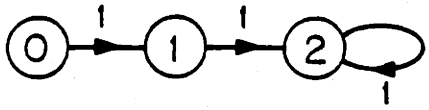


Chain P1:

if $x^{C0}=0$:



if $x^{C0}\neq 0$:



In terms of the interaction doublets (old interaction, new interaction), the states of chain P1 are defined as follows:

state	interaction doublet
0	(1,0)
1	(0,1)
2	(1,1)

where interaction 0 is an excitation attempt.

The Markov chain interval is 160 msec. There are no Gaussian observations for either chain and chain C0 also does not have any pseudo-observations.

Chain P1 initiates a binary pseudo-observation. The observation pmfs of the pseudo-observation, conditioned on the state transition, are:

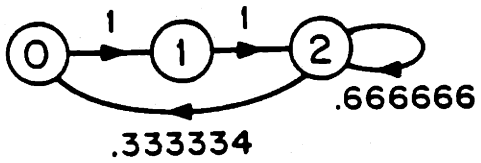
conditioning state transition (old state, new state)		pseudo-observation pmf (value at 0, value at 1)	
0	0	.5	.5
0	1	.830321	.169679
1	0	.5	.5
1	2	.001	.999
2	0	.068273	.931727
2	2	.001	.999

Figure 6.7 LE 0 Pass 1 LEPM for the Design Model of Figure 6.3 and Table 6.1 Column 1.

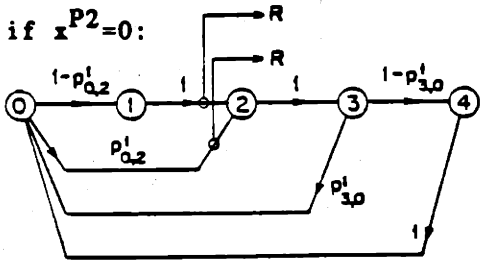
	truth interaction triplet $(h_{n-1}^{0,1}, h_n^{0,1}, h_{n+1}^{0,1})$							
	(0,0,0)	(1,0,0)	(0,1,0)	(1,1,0)	(0,0,1)	(1,0,1)	(0,1,1)	(1,1,1)
estimated interaction 0	0	0	0	107	0	1824	55	1
$^{0,1}\hat{h}_n^{P0,C1}$ 1	0	0	0	1883	0	165	1935	4012

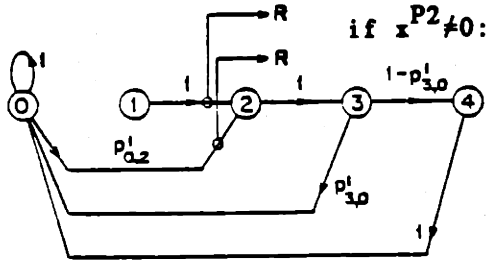
Table 6.3 Statistics on Interaction Estimates From LE 0 Pass 1.

Chain P2:



Chain C1:





In terms of the interaction doublets (old interaction, new interaction), the states of chain P2 are defined as follows:

state	interaction doublet
0	(1,0)
1	(0,1)
2	(1,1)

where interaction 0 is an excitation attempt.

The Markov chain interval is 160 msec.

Chain P2 has no Gaussian observations. There is a Gaussian observation from chain C1, modeling the R wave. The signature sampling interval is 160 msec. The R wave is one sample long, with mean 10 and standard deviation $\sqrt{5}$. The Gaussian observation noise has mean 0 and standard deviation 1.

Chain P2 initiates a binary pseudo-observation. The observation pmfs of the pseudo-observation, conditioned on the state transition, are:

conditioning state transition (old state, new state)		pseudo-observation pmf (value at 0, value at 1)	
0	1	.917044	.082956
1	2	.027638	.972362
2	0	.053769	.946231
2	2	.001	.999

Chain C1 initiates no pseudo-observations.

Figure 6.8 LE 1 Pass 2 LEPM for the Design Model of Figure 6.3 and Table 6.1 Column 1.

P2 Model:

Truth	Estimated R	Estimated fn	Estimated rs
R	1788	0	1788
fp	0		0
cs	1788	0	

fpR = 0/1788 = 0
fnR = 0/1788 = 0
teR = (0+0)/1788 = 0

Zero-Memory Model:

Truth	Estimated R	Estimated fn	Estimated rs
R	1787	1	1788
fp	49		49
cs	1836	1	

fpR = 49/1836 = .0267(.141)
fnR = 1/1788 = .000560(1.00)
teR = (49+1)/1788 = .0280

Table 6.4 Estimator Performance for the Comparison of the P2 and the Zero-Memory Models in the LEPM for LE 1 Pass 2.

6.7. Augmented Interactions--LE 1 Pass 0 to LE 0 Pass 1

Consider the design model of Figure 6.4. A reason for considering augmented interaction estimates from LE 1 pass 0 to LE 0 pass 1 is the observability issue mentioned in Section 5.5. Note first that interactions can only be observed indirectly via submodel 1 since submodel 0 initiates no signatures. Furthermore, when $x_n^1 \neq 0$, submodel 1 is not affected at all by the impinging interactions. Note that when $x_n^1 \neq 0$, the pmfs for the time since x_n^1 was last 0 and the time when x_n^1 is next 0 depend on the present value of x_n^1. Therefore, it is reasonable to anticipate that the quality of the interaction estimate at time n depends on the state of chain C1 at time n.

A second reason for considering augmented interactions is to ameliorate the effects of the aggregation of submodel 0 in LE 1 pass 0. Note first that during a conducted beat in the design model, the states of submodel 0 and submodel 1 are highly correlated. That is, the global state (x_n^0, x_n^1) starts at (0,0) and, until $x_n^0=6$, the global state is with high probability either (i,i) or (i,i+1). Therefore, knowing x_n^1 implies a great deal about x_n^0. Now assume that LE 1 pass 0 gives reasonable

performance in the sense that $\hat{x}_n^{C1}$ is near to x_n^1 most of the time. In this case the estimate $\hat{x}_n^{C1}$ implies a great deal about x_n^0. Consequently it should be possible to use this information to refine the interaction estimate computed from $\hat{x}_n^{P0}$. Specifically, in P0 chains of the form shown in Figure 6.2, the information derived from $\hat{x}_n^{C1}$ can indicate which interactions are due to spurious returns of x_n^{P0} to state 0. As discussed in the next section, spurious returns to state 0 in the P0 chain are common when p (in Figure 6.2) is larger that a certain threshold. In spite of this, such p are important because they contribute to superior performance for other reasons. By providing information useful in isolating the spurious excitation attempts indicated by $\hat{x}_n^{P0}$, $\hat{x}_n^{C1}$ is essentially circumventing the aggregation used to transform x_n^0 into x_n^{P0} by exploiting the strong correlation of x_n^1 with x_n^0 in the time interval including and immediately after a successful excitation.

A particular realization of the stochastic process defined by the design model of Figure 6.4 is shown in Figure 6.9. For this design model, a LE 1 pass 0 estimator was constructed analogous to the estimator shown in Figure 6.6. Statistics on the augmented interaction estimates from this estimator are shown in Table 6.5. When using unaugmented interactions, the relevant statistics are the marginal statistics (where the state estimate has been summed out of the joint statistics of Table 6.5). These marginal statistics are shown in Table 6.6. The increased information in the more detailed statistics is dramatic. For example, if the ICS transmits the information that $^{1,0}\hat{h}_n^{P0,C1}=1$ and $^{1,0}\hat{x}_n^{C1}=6$, then the LE 0 pass 1 estimator is essentially assured that the

truth interaction triplet is (1,1,1). If only $^{1,0}\hat{h}_n^{P0,C1}=1$ is transmitted, the uncertainty is much greater.

While the augmented interaction estimates carry more information, it is not guaranteed that the estimation algorithms we propose are able to use this additional information to achieve a useful gain in performance. In fact, as we will now describe, it is only when the tpm for the P0 chain is simultaneously altered that we achieve decidedly superior performance by including augmented interaction information. Consider the estimators described in block diagram form in Figures 6.10 and 6.11. The only difference in the structure of these two estimators is that in one (Figure 6.10) the ICS communicates augmented estimated interactions from LE 1 pass 0 to LE 0 pass 1, while in the other (Figure 6.11) only the estimated interactions themselves are used. The LEPM for LE 1 pass 0 is the same in both estimators. The parameters of the LEPMs for LE 0 pass 1 and LE 1 pass 2 are different since the parameters depend (through the pseudo-observation observation pmfs) on the statistics of the communicated quantities. Table 6.7 shows the final wave-tracking misclassification results for these two estimators based on the same simulated ECG used to compute the statistics of Tables 6.5 and 6.6. Two sets of statistics for each estimator are shown. Since the signature is only one sample long, in one set (labeled without phase shifts) no phase-shifted matches are allowed. In the second set (labeled with phase shifts), phase shifts of ± 1 Markov chain cycle are allowed. The performance with augmented interactions is disappointing as it is essentially unchanged from the performance with interactions alone.

Changing the P0 tpm (see the following section) obviously changes

the statistics of both the augmented and unaugmented interaction estimates. In the estimators considered so far (i.e. the estimators whose performance is indicated in Tables 6.5, 6.6, and 6.7), the value of p in the LE 1 pass 0 LEPM of Figure 6.6 was p=.105258. In order to demonstrate that augmented interactions can provide significantly improved performance, we now consider a new value of p, specifically p=.399975.

Simulations equivalent to those generating the statistics of Tables 6.5-6.7 were performed for the new value of p. These statistics are shown in Table 6.8 (for the augmented interaction estimates from LE 1 pass 0), Table 6.9 (for the unaugmented interaction estimates from LE 1 pass 0), and Table 6.10 (for the wave-tracking misclassification performance). Tables 6.5 and 6.6 are compared with Tables 6.8 and 6.9 in the following section. In this section we focus on the comparison of Table 6.7 and Table 6.10.[2]

Returning to the suboptimal estimator, note that the performance of the estimator using augmented interaction estimates and the new value of

---

[2] Table 6.10 actually corresponds to only one half of Table 6.7. We were unable to calculate the results for the estimator based on communicating unaugmented interactions because the Viterbi algorithm tree (see Forney, 1973) for LE 0 pass 1 exceeded the memory limitations of our computer. The reason for this difficulty is as follows. For all LE 0 pass 1 estimators in the DM1 class of design model, the pseudo-observation is the only observation. For this specific example the observation pmfs (constructed from the data of Table 6.9 and shown in Table 6.11) for the pseudo-observations are almost independent of the underlying truth interaction triplet. That is, the pmfs conditioned on the pair of triplets (1,1,0) and (1,0,1) are almost identical, as are those for the pair of triplets (0,1,1) and (1,1,1). Furthermore, the differences between the two pairs of pmfs are small. Therefore, the Viterbi algorithm must process a large number of observations before it is able to make a decision, and the resulting tree exceeds the available storage. Said another way, the pseudo-observation has a very low SNR.

p in the LEPM of LE 1 pass 0 is a substantial improvement over the performance of either estimator (i.e. using either augmented or unaugmented interaction estimates) using the old value of p in the LEPM of LE 1 pass 0 (compare Tables 6.7 and 6.10). Therefore, in the sequel, augmented interaction estimates are used because they allow greater freedom in the choice of the tpm for the LEPM of LE 1 pass 0 and this freedom can be exploited to achieve superior performance. This performance improvement can be understood in the following manner.

(1) Looking ahead to the following section, increasing p provides better performance because it decreases the entrainment of P0 and C1.

(2) However, increasing p leads to many spurious excitations. But it is exactly these spurious excitations which augmented interactions are most effective in suppressing.

Thus there is a substantial synergistic effect between increasing p and transmitting augmented interaction estimates.

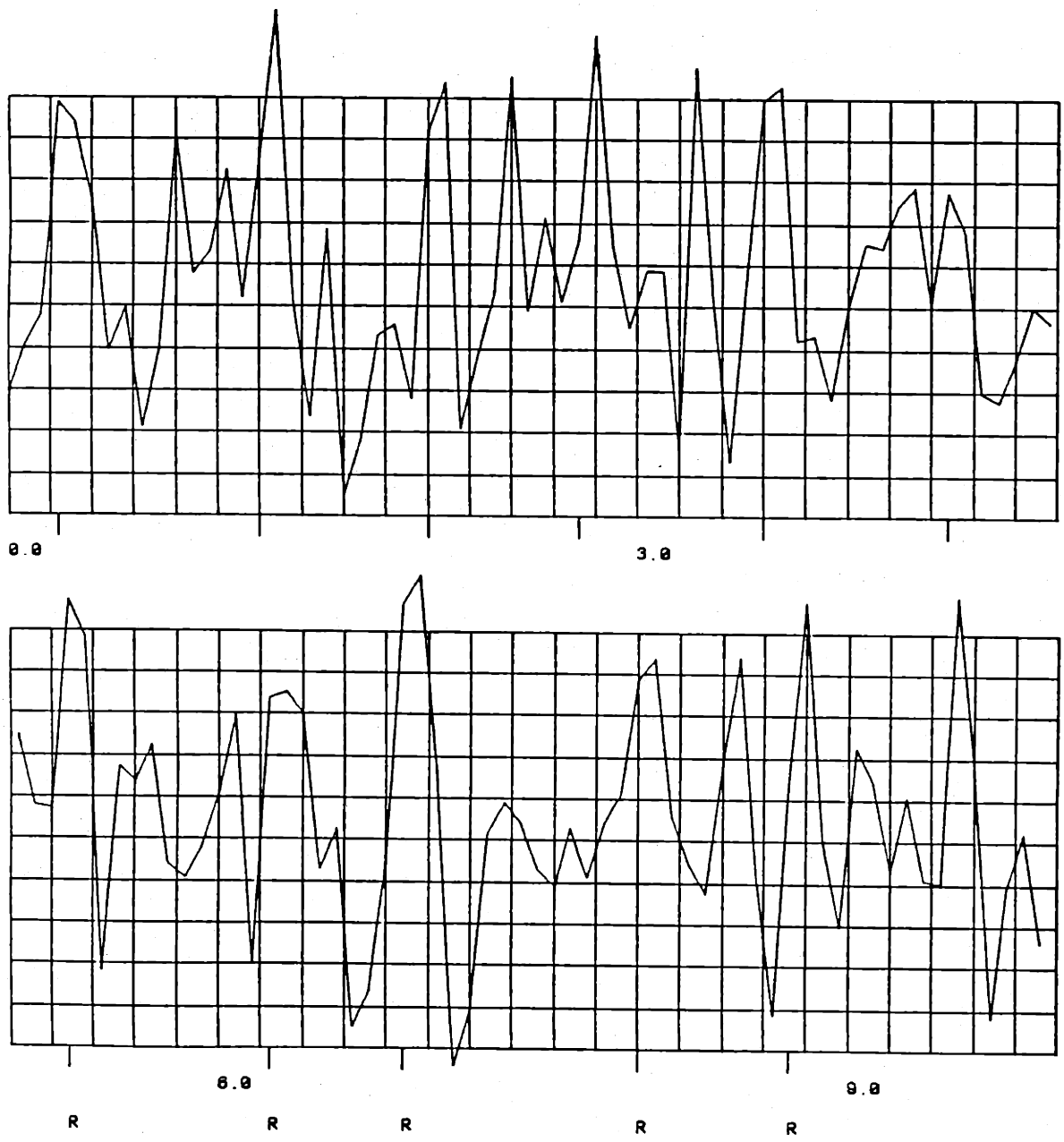


Figure 6.9 A Realization of the Stochastic Process Defined in Figure 6.4.

			truth interaction triplet $(h^{0,1}_{n-1}, h^{0,1}_{n}, h^{0,1}_{n+1})$			
			(1,1,0)	(1,0,1)	(0,1,1)	(1,1,1)
estimated interaction $1,0\hat{h}^{P0,C1}_{n}$	0	estimated state $1,0\hat{x}^{C1}_{n}$ 0	58	348	364	81
		1	0	0	0	0
		2	0	0	0	0
		.	.	.	.	.
		.	.	.	.	.
		.	.	.	.	.
		10	0	0	0	0
	1	0	1277	1255	1029	8056
		1	0	1	2	2
		2	50	57	348	397
		3	3	50	57	742
		4	0	3	50	799
		5	1	0	3	848
		6	0	1	0	851
		7	20	0	1	831
		8	40	13	0	769
		9	171	40	13	598
		10	282	134	36	309

Table 6.5 Statistics on Augmented Interaction Estimates From LE 1 Pass 0.

		truth interaction triplet $(h^{0,1}_{n-1}, h^{0,1}_{n}, h^{0,1}_{n+1})$			
		(1,1,0)	(1,0,1)	(0,1,1)	(1,1,1)
estimated interaction $1,0\hat{h}^{P0,C1}_{n}$	0	58	348	364	81
	1	1844	1554	1539	14202

Table 6.6 Statistics on Interaction Estimates From LE 1 Pass 0 (based on Table 6.5).

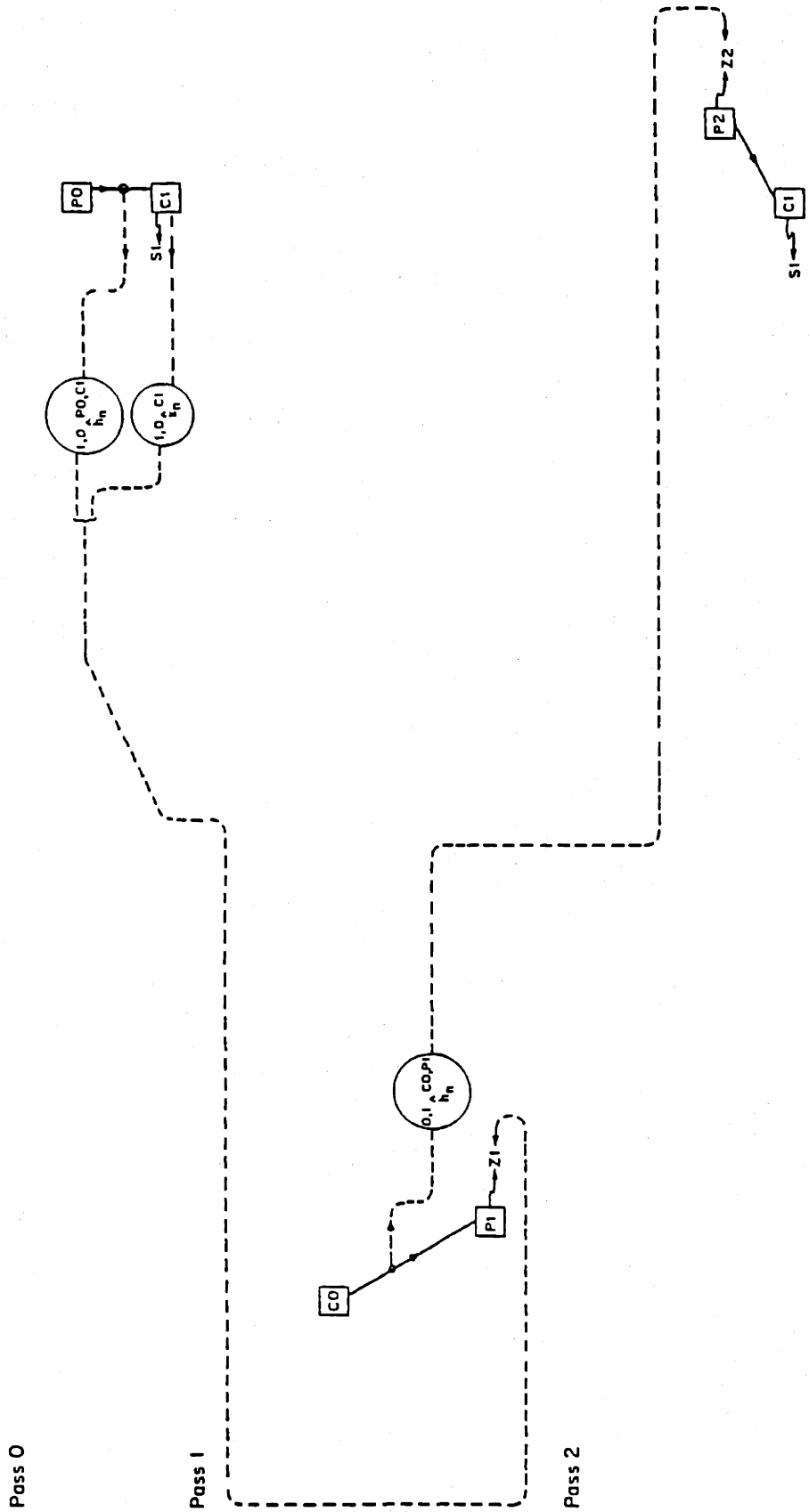


Figure 6.10 Comparison of Augmented and Unaugmented Interactions for Communication Between LE 1 Pass 0 and LE 0 Pass 1. Estimator for the Augmented Case.

Section 6.7.

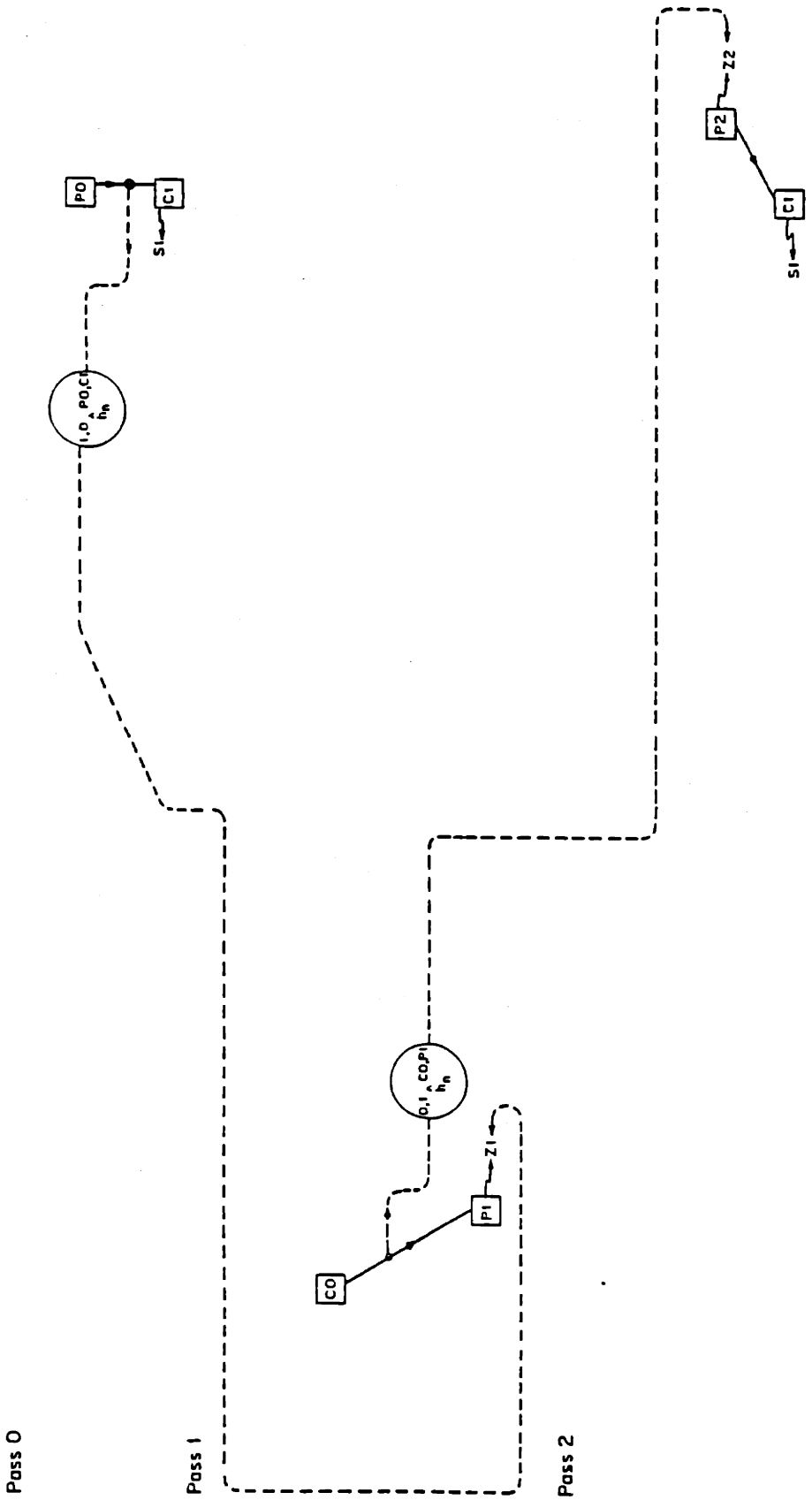


Figure 6.11 Comparison of Augmented and Unaugmented Interactions for Communication Between LE 1 Pass 0 and LE 0 Pass 1. Estimator for the Unaugmented Case.

Estimator of Figure 6.10

With phase shifts

		Estimated R	fn	rs
Truth	R	1238	518	1756
	fp	25		25
	cs	1263	518	

fpR = 25/1263 = .0198(.198)
fnR = 518/1756 = .295(.0369)
teR = (25+518)/1756 = .309

Without phase shifts

		Estimated R	fn	rs
Truth	R	1082	674	1756
	fp	181		181
	cs	1263	674	

fpR = 181/1263 = .143(.0688)
fnR = 674/1756 = .384(.0302)
teR = (181+674)/1756 = .487

Estimator of Figure 6.11

With phase shifts

		Estimated R	fn	rs
Truth	R	1207	545	1752
	fp	20		20
	cs	1227	545	

fpR = 25/1227 = .0163(.222)
fnR = 545/1752 = .311(.0356)
teR = (20+545)/1752 = .322

Without phase shifts

		Estimated R	fn	rs
Truth	R	1062	690	1752
	fp	165		165
	cs	1227	690	

fpR = 165/1227 = .134(.0724)
fnR = 690/1752 = .394(.0296)
teR = (165+690)/1752 = .488

Table 6.7 Comparison of Augmented and Unaugmented Interactions for Communication Between LE 1 Pass 0 and LE 0 Pass 1. Performance.

			truth interaction triplet $(h^{0,1}_{n-1}, h^{0,1}_{n}, h^{0,1}_{n+1})$			
			(1,1,0)	(1,0,1)	(0,1,1)	(1,1,1)
estimated interaction $1,0\hat{h}^{P0,C1}_{n}$	0	estimated state $1,0\hat{x}^{C1}_{n}$ 0	425	606	176	202
		1	0	0	0	0
		2	44	73	390	515
		3	3	28	33	296
		4	5	8	50	985
		5	1	1	3	352
		6	6	5	5	1037
		7	9	0	1	347
		8	47	18	4	911
		9	42	7	9	252
		10	203	79	33	357
	1	estimated state $1,0\hat{x}^{C1}_{n}$ 0	649	488	449	3934
		1	73	390	432	127
		2	34	35	174	143
		3	8	50	75	915
		4	1	3	29	329
		5	5	5	8	1035
		6	0	1	1	355
		7	34	6	5	1008
		8	7	9	0	300
		9	202	47	18	719
		10	103	42	7	160

Table 6.8 Statistics on Augmented Interaction Estimates From LE 1 Pass 0.

		truth interaction triplet $(h^{0,1}_{n-1}, h^{0,1}_{n}, h^{0,1}_{n+1})$			
		(1,1,0)	(1,0,1)	(0,1,1)	(1,1,1)
estimated interaction $1,0\hat{h}^{P0,C1}_{n}$	0	785	825	704	5254
	1	1116	1076	1198	9025

Table 6.9 Statistics on Interaction Estimates From LE 1 Pass 0 (based on Table 6.8).

Estimator of Figure 6.10

With phase shifts

		Estimated		
		R	fn	rs
Truth	R	1364	395	1757
	fp	58		58
	cs	1420	395	

fpR = 58/1420 = .0408(.129)
fnR = 395/1757 = .225(.0443)
teR = (58+395)/1757 = .258

Without phase shifts

		Estimated		
		R	fn	rs
Truth	R	1186	571	1757
	fp	234		234
	cs	1420	571	

fpR = 234/1420 = .165(.0597)
fnR = 571/1757 = .325(.0344)
teR = (234+571)/1757 = .458

Table 6.10 Comparison of Augmented and Unaugmented Interactions for Communication Between LE 1 Pass 0 and LE 0 Pass 1. Performance of the Estimator Communicating Augmented Interactions.

		$Pr(\hat{h}^{1,0}_{n}{}^{C0,P1}_{C0,P1} = h \mid x^{P1}_{n-1}, x^{P1}_{n})$	
		h=0	h=1
state transition $x^{P1}_{n-1} \rightarrow x^{P1}_{n}$	(1,1)→(1,0)	.413	.587
	(1,0)→(0,1)	.434	.566
	(0,1)→(1,1)	.370	.630
	(1,1)→(1,1)	.368	.632

Table 6.11 Observation Pmfs for the Pseudo-Observation of LE 0 Pass 1 When Communicating Unaugmented Interaction Estimates From LE 1 Pass 0.

6.8. The P0 Chain

In Section 6.1, all aspects of the P0 type chains were fixed except for the tpm(s). Recall from Section 6.1 that three algorithms were proposed for choosing these tpms:

(1) preservation of $\lim_{n\to\infty} Pr(x^0(n) \in U^{0,1,i} \mid x^0(n-1) \in U^{0,1,j})$,

(2) preservation of $\lim_{n\to\infty} Pr(x^0(n) \in U^{0,1,i})$ with the additional constraint that the tpm have equal rows, and

(3) optimization over the parameter p in the tpm of Figure 6.2.

Recall also the (1) leads to tpms with the structure shown in Figure 6.2 so that (1) is a specific choice of p in (3).

Initially, we considered methods (1) and (2). Therefore the comparison between these two methods is based on estimators of the type shown in Figure 6.11. That is, the comparison is based on estimators in which the ICS does not communicate augmented interactions. Consider the design models of Figure 6.3 and Table 6.1 columns 2-5. These design models differ from each other solely due to changes in the R wave mean. Realizations for each different R wave mean are shown in Figure 6.12. Table 6.12 lists the performance of the entire estimator for both methods (1) and (2) in terms of fpR, fnR, and teR as the mean of the R wave varies.

On the basis of the results shown in Table 6.12, method (1) is at least as effective as method (2). Therefore, method (2) was not considered further.

At a later point in the investigation we again considered tpms for P0. At this point we considered method (3) and based the comparison on estimators of the type shown in Figure 6.10. That is, the comparison is based on estimators in which the ICS communicates augmented interaction estimates from LE 1 pass 0 to LE 0 pass 1 and unaugmented interaction estimates from LE 0 pass 1 to LE 1 pass 2. Recall from the prior section that communication of augmented interaction estimates from LE 1

pass 0 to LE 0 pass 1 was necessary for certain choices of tpm in the LEPM of LE 1 pass 0 since the unaugmented interaction estimates provided very little information. The design model used is shown in Figure 6.4. A realization is shown in Figure 6.5.

Estimated statistics on the wave-tracking performance of the complete estimator in terms of fpR, fnR, and teR as p varies are shown in Table 6.13 and plotted in Figures 6.13 and 6.14. Results on the global MAP estimator[3] are included for comparison. The case p=.105258 is method (1). In the previous sections estimators using method (1) and using p=.399975 were discussed.

The reason for the unusual choices of p is that these choices give integer (or half integer) increments in the value of $\lim_{n\to\infty} Pr(x^{P0}(n)=0)$ in comparison with the value implied by method (1). The ratio between the probability for some arbitrary p and for the p implied by method (1) is denoted α in Table 6.13 and in the remainder of the thesis. The parameter α therefore compares the a priori probability of an excitation attempt in the design model and in the LE 1 pass 0 LEPM. (Calculations not reported here indicate that the optimal value of α is quite variable for different design models and therefore the discussion in this thesis emphasizes p and not α).

Performance can be dramatically improved by the proper choice of p. In the example for which results are shown in Table 6.13 and

[3] Throughout this thesis, "global MAP estimator" is the wave-tracking estimator in which first the global MAP state-trajectory estimate is computed and then the wave-tracking estimate is computed from the state-trajectory estimate. Note that this estimator is not optimal for the wave-tracking problem.

Figures 6.13 and 6.14, the optimal value of p is near .6. We do not have a method (except simulation) for determining the optimal value of p given a design model and the structure of our estimator. However, in the remainder of this section we point out two qualitative aspects of our estimator that might be relevant to such a method.

The P0 chain (shown in Figure 6.2) has two basic types of behavior depending on the value of p. In the first case, for p small, the chain prefers to remain in state 1. In fact, the a priori maximum probability trajectory is for the state to remain forever in state 1. Therefore, in the LE 1 pass 0 estimator, the observations must outweigh the a priori trajectory information in order for $\hat{x}_n^{P0}$ to ever leave state 1 and attempt to excite chain C1. This explains why the statistics on the estimated augmented interactions for p=.105258 (see Table 6.5) are nonzero only for ${}^{1,0}\hat{x}_n^{C1}=0$.

The second qualitative type of behavior, for p large, is a period-two oscillation between states 0 and 1. In this case the effect of the observations is to counteract this predisposition in order to cause a self-transition at 1 in order to adjust the phase of the next excitation attempt (that is, residency in state 0) or to cause several self-transitions at 1 in a row in order to avoid exciting the lower chain. An example of an instance in which the estimator must make several self-transitions at 1 in a row is a dropped R wave. This explains why the statistics on the estimated augmented interactions for p=.399975 (see Table 6.8) are nonzero for <u>many</u> values of ${}^{1,0}\hat{x}_n^{C1}$.

The break point between the two types of behavior occurs when the

probability of the state trajectory 1→0→1 equals the probability of the state trajectory 1→1→1. This occurs when p is a root of the equation $p=(1-p)^2$. The only root in [0,1] is $\frac{3-\sqrt{5}}{2}\simeq.381966$. This value of p lies between the two values used in the prior section (i.e. .105258 and .399975) and therefore the statistics shown in Tables 6.5 and 6.8 have the properties noted in the previous paragraphs. Since the optimal value of p lies near .6, it seems that the second type of behavior is preferable.

In methods (1) and (2), the tpm for P0 is chosen so that certain properties of submodel 0 are preserved by the P0 chain. Therefore, some properties of the interactions impinging on submodel 1 are preserved. However, the properties chosen may not be the appropriate properties for overall performance considerations. For example, preserving the mean rate of excitation attempts [as is done in method (2)] does not make any statement about the pmf on the interarrival times between excitation attempts. However, it is the almost periodic arrangement of excitation attempts by submodel 0 that makes most excitation attempts successful. In light of these ideas it may make more sense to preserve properties of successful excitation attempts or equivalently (in these design models) R waves. (Note that this is a statistic that is not a function of only the subprocess state space being aggregated).

In the design model of Figure 6.4, the steady state probability of an R wave occurring at any particular time is $\lim_{n\to\infty} \Pr(x^1(n)=4)$, which is approximately .08732. Table 6.13 and Figure 6.15 give the corresponding probability for the LE 1 pass 0 LEPMs derived from this design model under different choices of p. Based on these results, a value of p

somewhere in the range (.23,.40) would give a LE 1 pass 0 LEPM with a steady-state R wave probability equal to the steady-state R wave probability of the design model, that is .08732. However, since the optimal value of p lies near .6, it is not clear that this discussion of alternative statistics for preservation under the submodel 0 to P0 aggregation contributes toward an understanding of the optimal value.

The selection of p, given a design model and an estimator architecture, is still an open question. However, the discussion of this section does serve to illustrate two important points.

(1) The choice of aggregation can have a major effect on signal-processing performance.

(2) The proper (i.e. best signal-processing performance) choice of aggregation may not be an aggregation which preserves an obvious property of the design model.

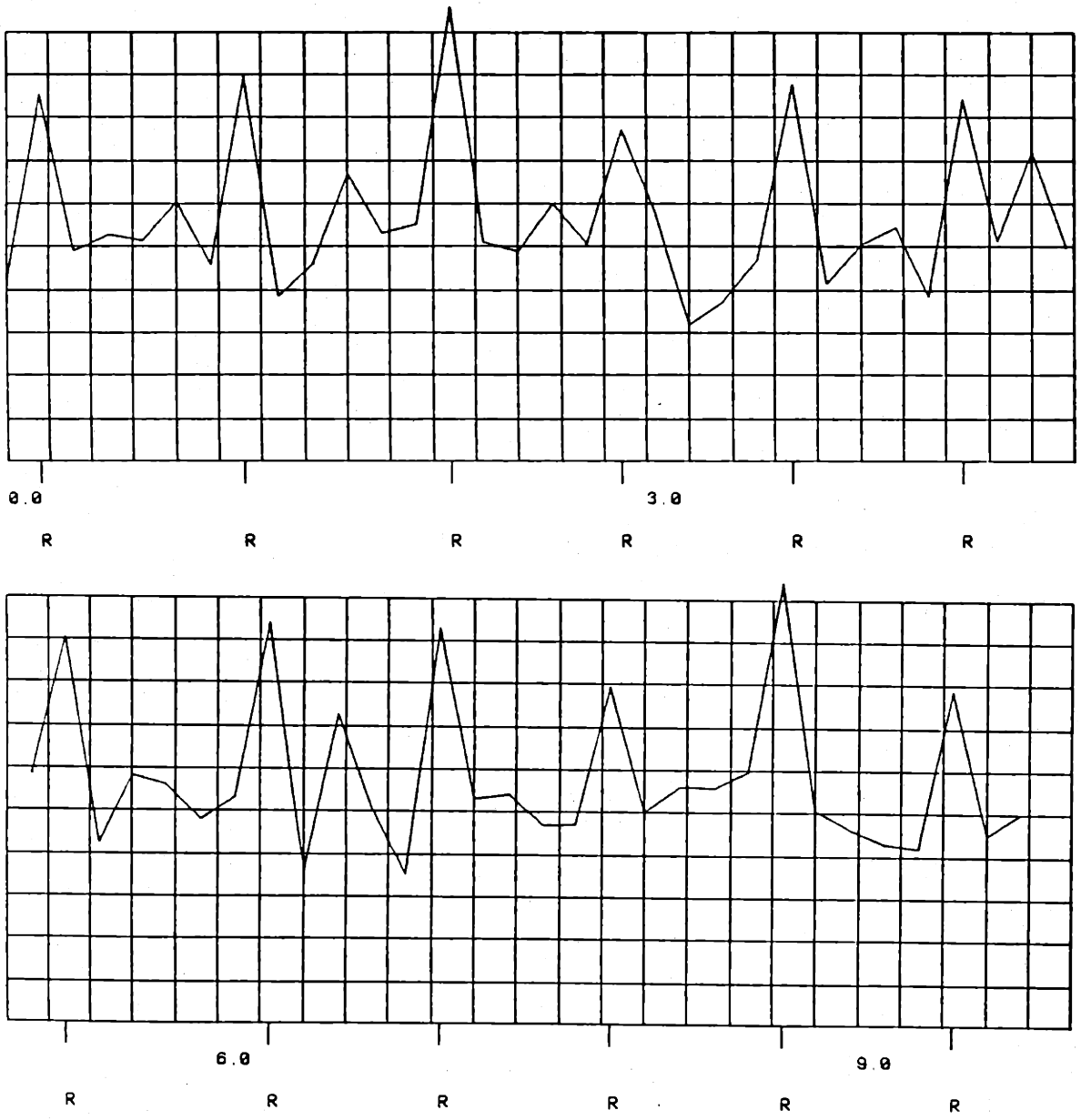


Figure 6.12 A Realization of the Stochastic Process Defined in Figure 6.3 and Table 6.1.
Part (a): Parameters from Column 2.

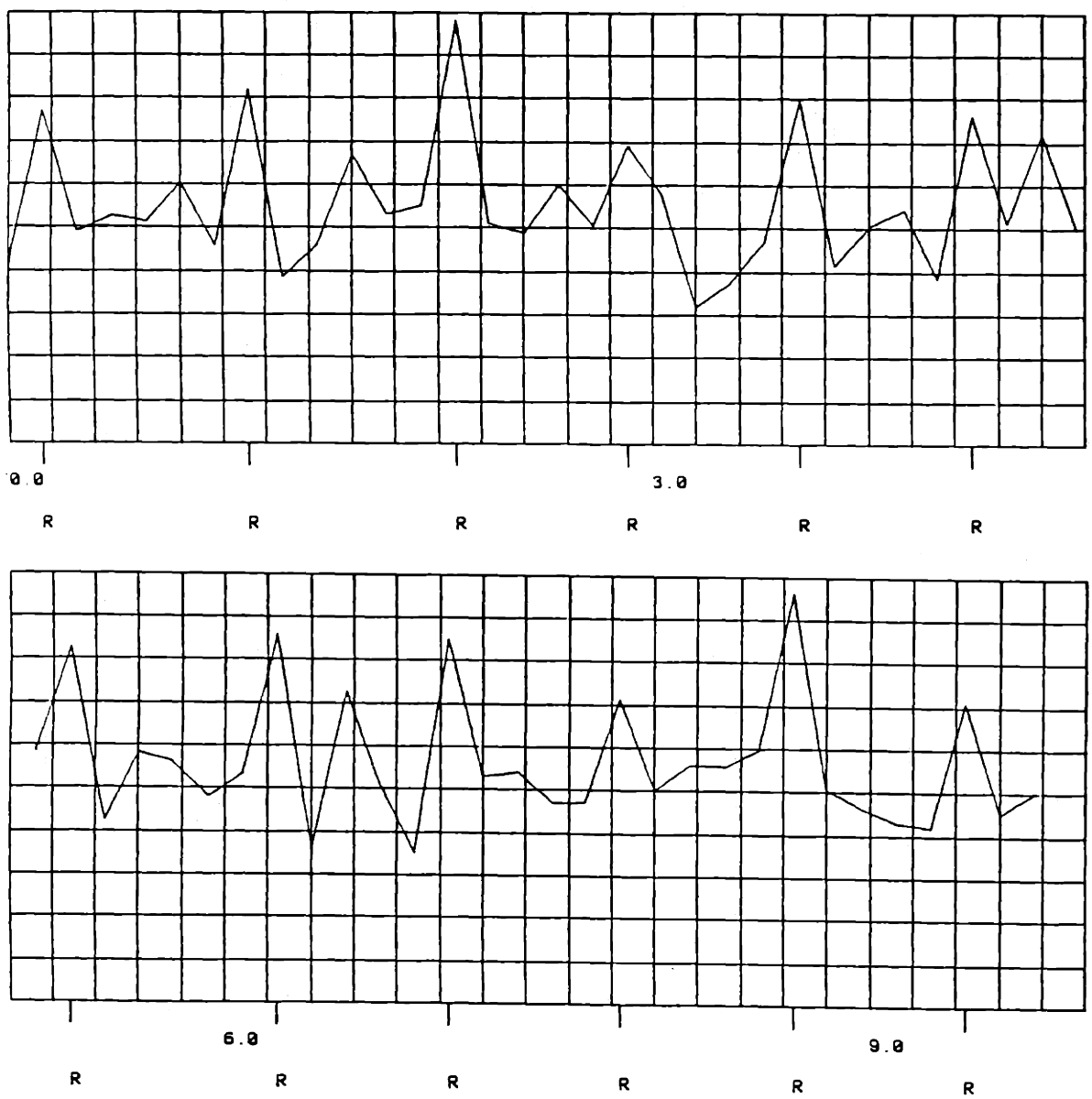


Figure 6.12 Continued.
Part (b): Parameters from Column 3.

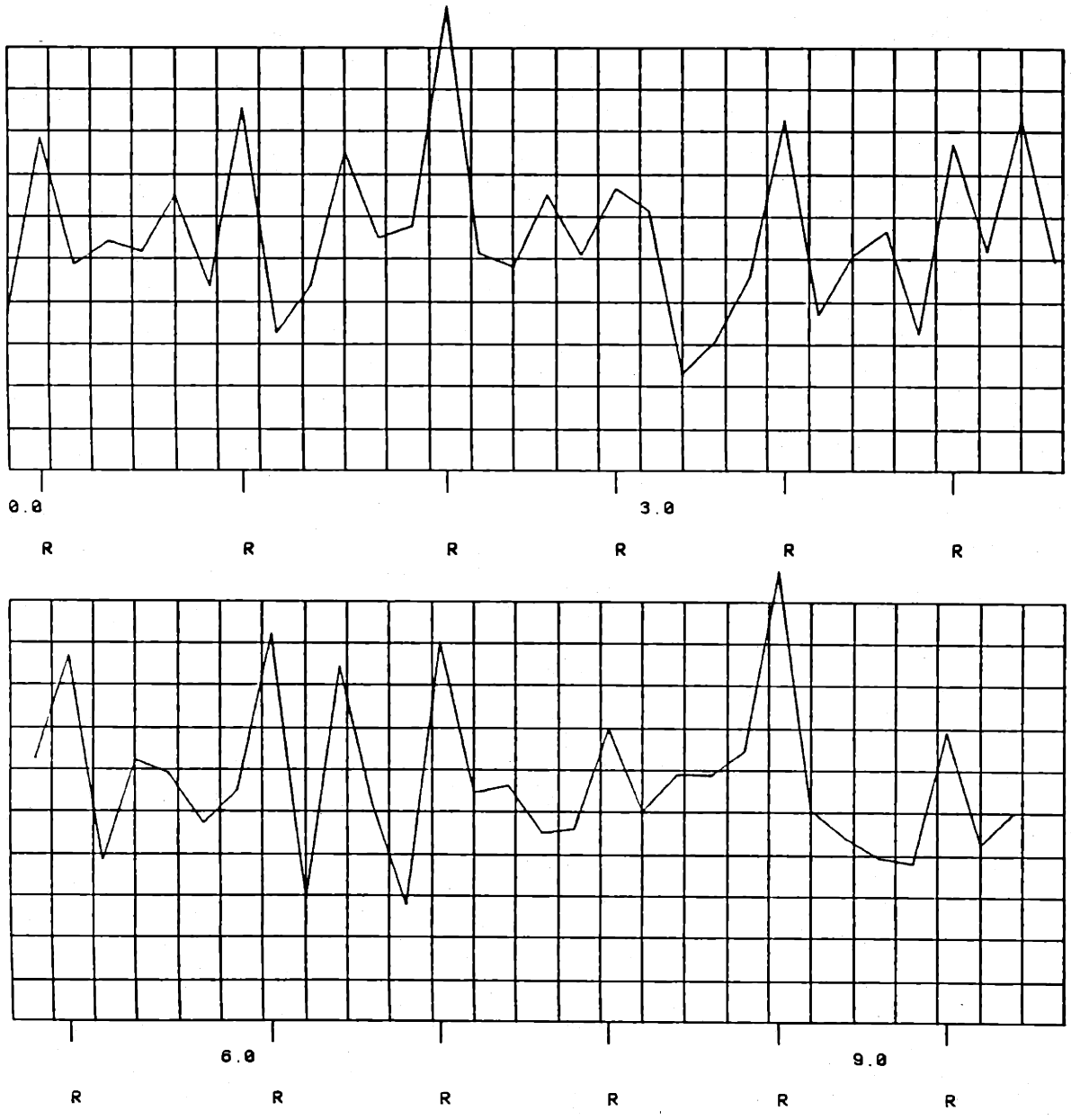


Figure 6.12 Continued.
Part (c): Parameters from Column 4.

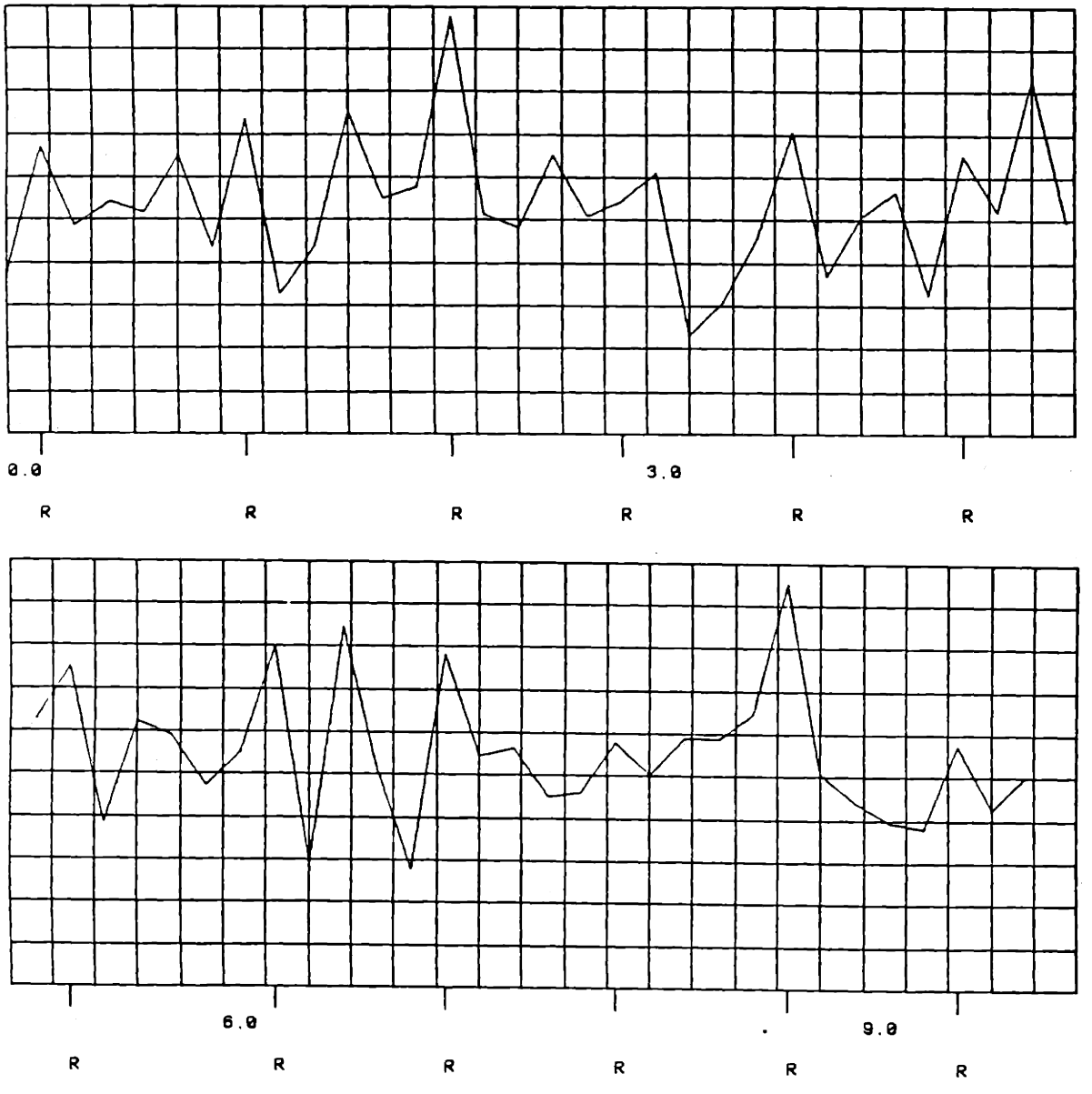


Figure 6.12 Continued.
Part (d): Parameters from Column 5.

		LE 1 pass 0 tpm method 1			LE 1 pass 0 tpm method 2		
		fpR	fnR	teR	fpR	fnR	teR
mean	5	10/1779 =.00562 (.315)	19/1788 =.0106 (.228)	(10+19)/1788 =.0162	10/1799 =.00562 (.315)	19/1788 =.0106 (.228)	(10+19)/1788 =.0106
	4	52/1781 =.0292 (.137)	59/1788 =.0330 (.128)	(52+59)/1788 =.0621	50/1772 =.0282 (.139)	66/1788 =.0369 (.121)	(50+66)/1788 =.0649
	3	135/1759 =.0767 (.0827)	164/1788 =.0917 (.0744)	(135+164)/1788 =.167	156/1770 =.0881 (.0765)	175/1789 =.0978 (.0718)	(156+175)/1789 =.185
	2	387/1698 =.228 (.0447)	475/1786 =.266 (.0393)	(387+475)/1786 =.483	365/1635 =.223 (.0461)	517/1787 =.298 (.0371)	(365+517)/1787 =.494

Table 6.12 Comparison of Methods (1) and (2) for the Choice of a Tpm for Chains of Type P0.

p	α	With phase shifts fpR fnR teR	Without phase shifts fpR fnR teR	$\lim_{n\to\infty} \Pr(x^{C1}(n)=4)$ in LE 1 Pass 0
1	5.250233	.0360(.139) .238(.0427) .267	.165(.0604) .340(.0332) .471	no steady state
.909014	5	.0331(.145) .231(.0437) .257	.159(.0618) .331(.0341) .457	.1066452
.749942	4.5	.0382(.132) .210(.0463) .242	.170(.0582) .319(.0349) .458	.1045658
.615341	4	.0393(.130) .206(.0469) .238	.172(.0577) .315(.0352) .457	.1019672
.399975	3	.0408(.129) .225(.0443) .258	.165(.0597) .325(.0344) .458	.09446312
.235281	2	.0215(.184) .249(.0415) .266	.148(.0654) .346(.0328) .459	.08169435
.105258	1	.0198(.198) .295(.0369) .309	.143(.0688) .384(.0302) .487	.05748526
MAP estimator		.0375(.131) .181(.0507) .213	.174(.0564) .297(.0367) .445	

$$\alpha(p) = \lim_{n\to\infty} \Pr(x^{P0}(n)=0;p) \Big/ \lim_{n\to\infty} \Pr(x^{P0}(n)=0;.105258)$$

Table 6.13 Method (3) for the Choice of a Tpm for Chains of Type P0.

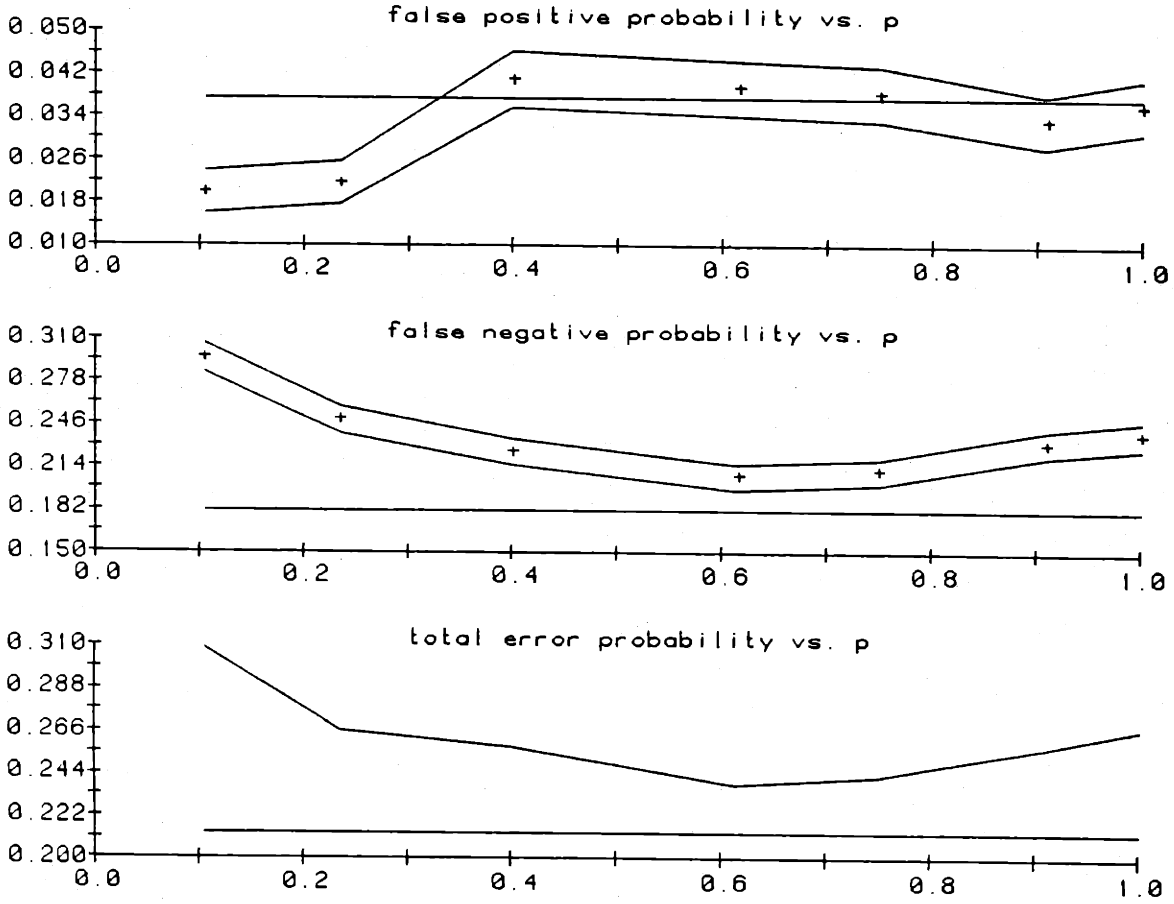


Figure 6.13 Performance as a Function of p when Phase-Shifted Matches are Allowed.
One sigma bounds derived from the estimated fractional standard deviation are included. The horizontal line indicates the performance of the global MAP estimator.

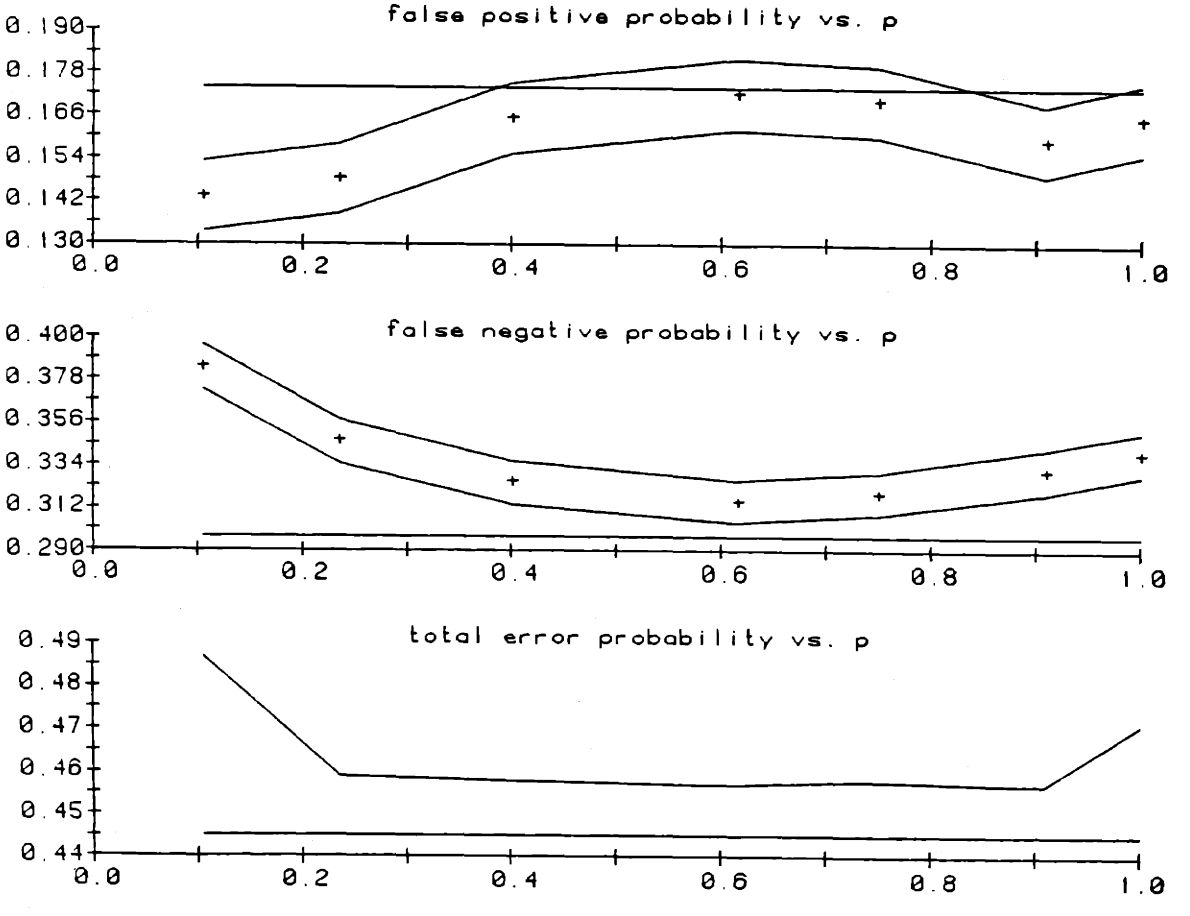


Figure 6.14 Performance as a Function of p when Phase-Shifted Matches are Not Allowed.
One sigma bounds derived from the estimated fractional standard deviation are included. The horizontal line indicates the performance of the global MAP estimator.

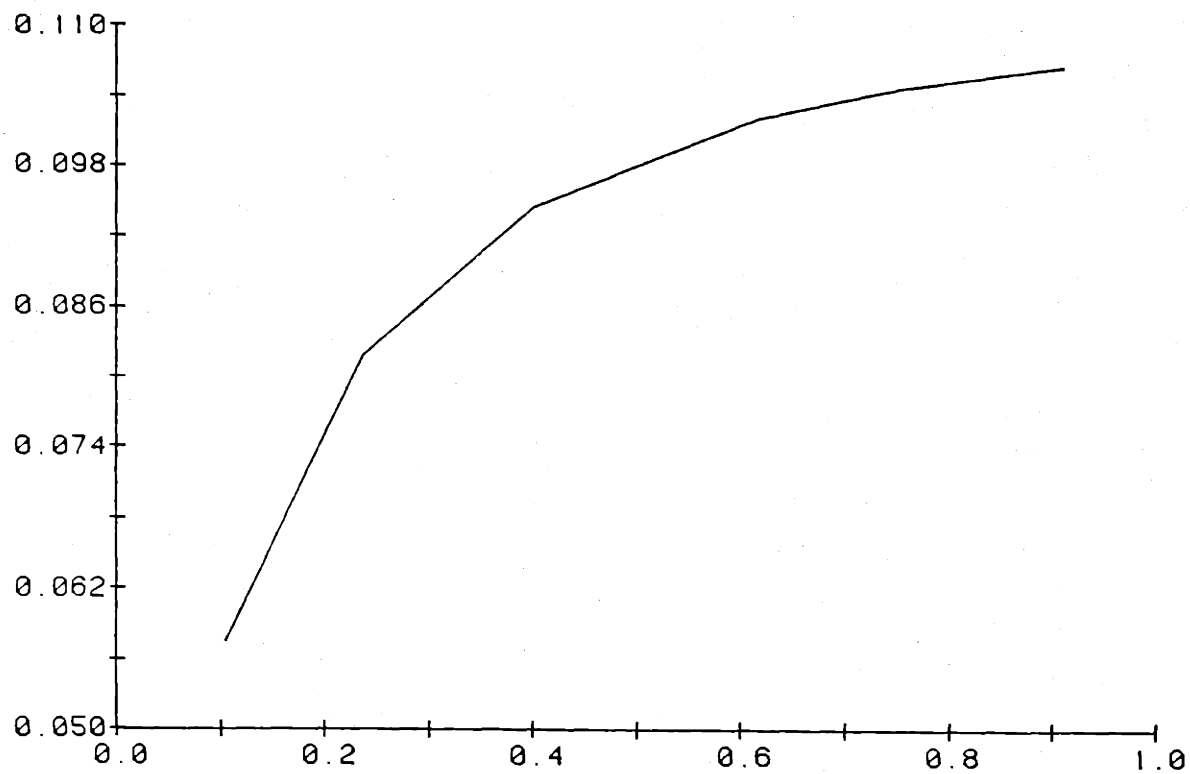


Figure 6.15 $\lim_{n\to\infty} \Pr(x^{C1}(n)=4)$ in the LEPM of LE 1 Pass 0 as a Function of p.

6.9. Robustness to Incorrect Modeling

Analytic ECG models are obviously more constrained than real ECGs. Consequently an important issue is the robustness of an estimation algorithm designed using such an analytic model. This section presents an example in which the suboptimal estimator is shown to be more robust than the global MAP estimator. Specifically, we define two classes of design models. Each design model in the first class is paired with a design model in the second class and visa versa. For each pair, we consider the performance of an estimator designed using the selected design model when the data is a realization of the same design model and when the data is a realization of the other design model of the pair--i.e. the matched and mismatched cases.

One set of design models, called the class of random P-R design models, is the set of design models of Figure 6.3 and Table 6.1 columns 2, 4, and 5. As noted previously, the parameter sets differ only in the choice of the R wave mean. Example realizations for each design model are given in Figure 6.12.

The second set of design models, called the class of deterministic P-R design models, is the set of design models shown in Figure 6.3 and Table 6.1 columns 6-8. The difference between the random P-R and deterministic P-R sets of design models is the delay line in submodel 1 between state 0 and state 2. This delay line represents the excitation-to-R-wave coupling interval. In the random P-R design model the interval is random while in the deterministic P-R design model it is deterministic. Example realizations for each deterministic P-R design model are shown in Figure 6.16.

For the two cases in which the data and estimator are based on the same design model, the performance statistics are shown in Tables 6.14 and 6.15 and Figures 6.17 and 6.18. The performance statistics for the important case in which the data is based on the random P-R design model and the estimator is based on the deterministic P-R design model are shown in Table 6.16 and Figure 6.19. The performance statistics for the reverse case are shown in Table 6.17 and Figure 6.20. In all cases, the suboptimal estimator has the block diagram shown in Figure 6.11.

Figures 6.17-6.20 show that the suboptimal estimator is insensitive to both mismatches. On the other hand (Figures 6.17-6.20), the global MAP estimator is insensitive to the mismatch in which the data is based on the deterministic P-R design model while the estimator is based on the random P-R design model, but it is sensitive to the other mismatch.

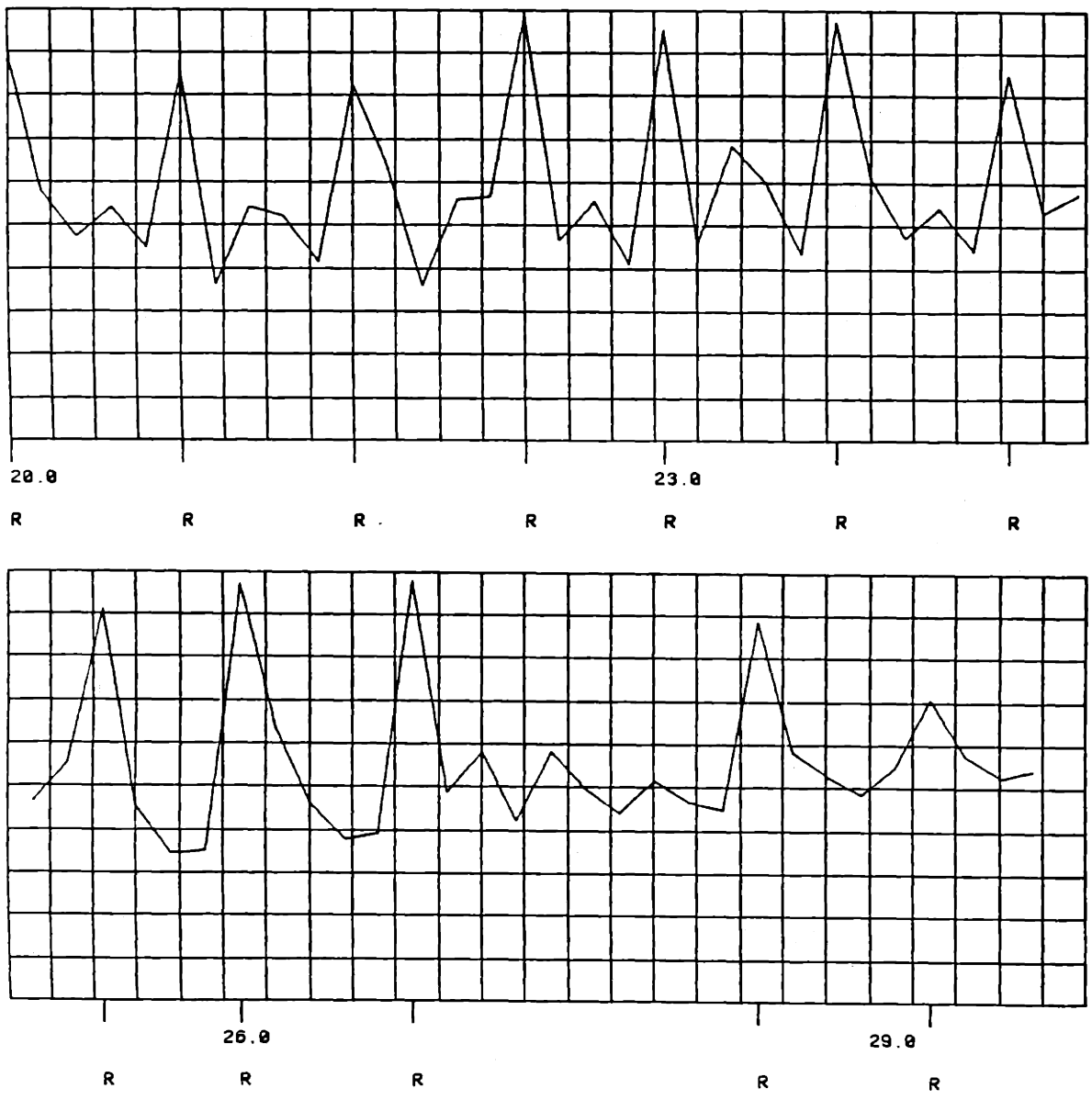
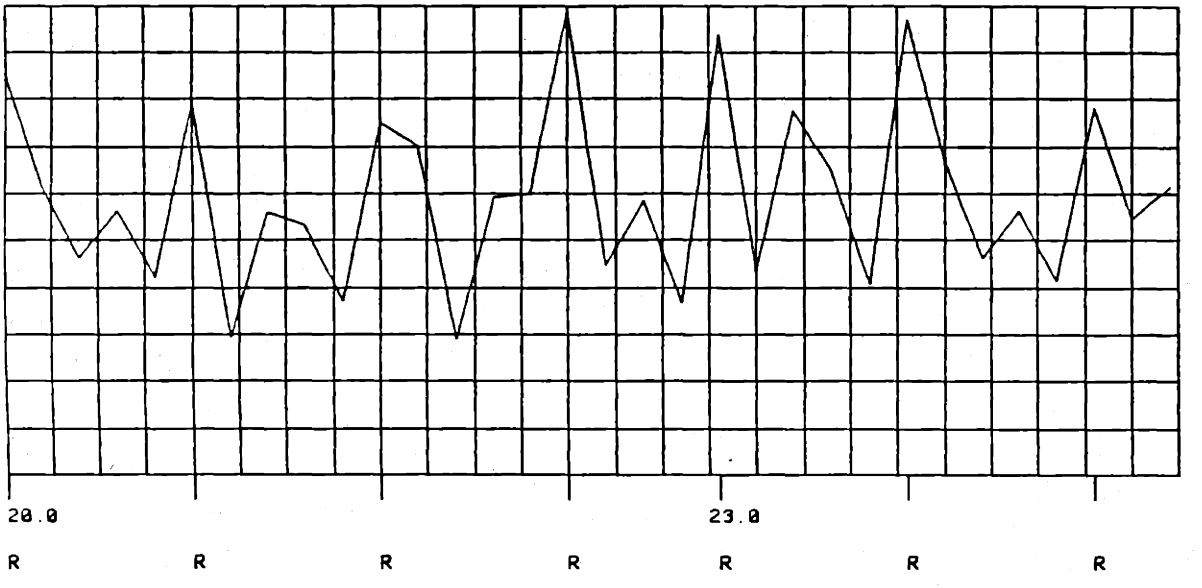


Figure 6.16 A Realization of the Stochastic Process Defined in Figure 6.3 and Table 6.1.
Part (a): Parameters from Column 6.



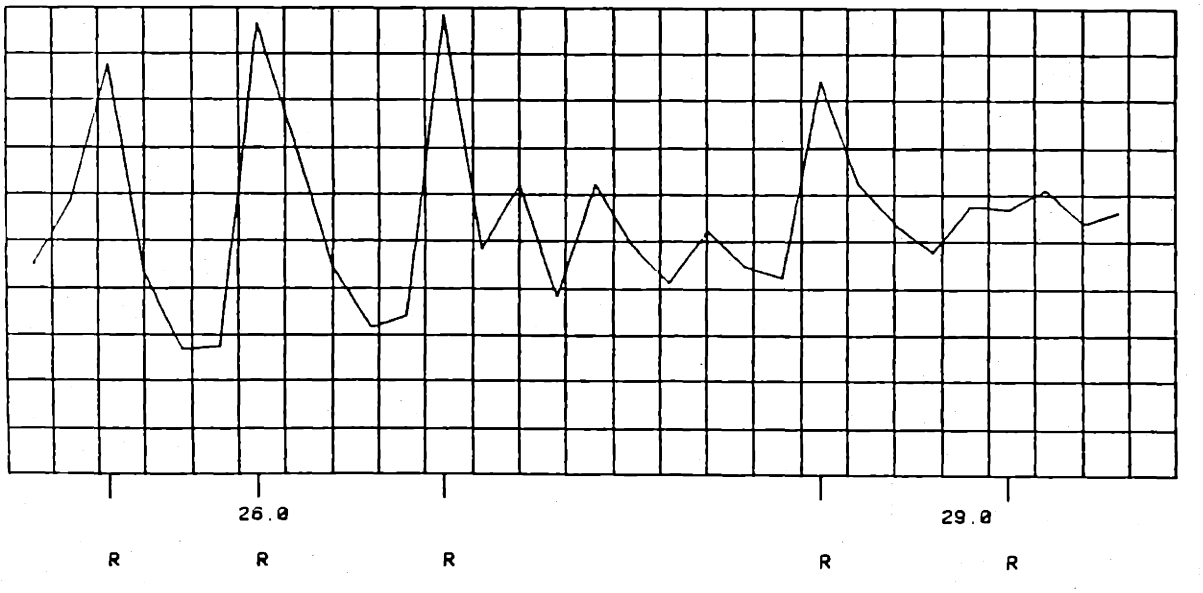


Figure 6.16 Continued.
Part (b): Parameters from Column 7.

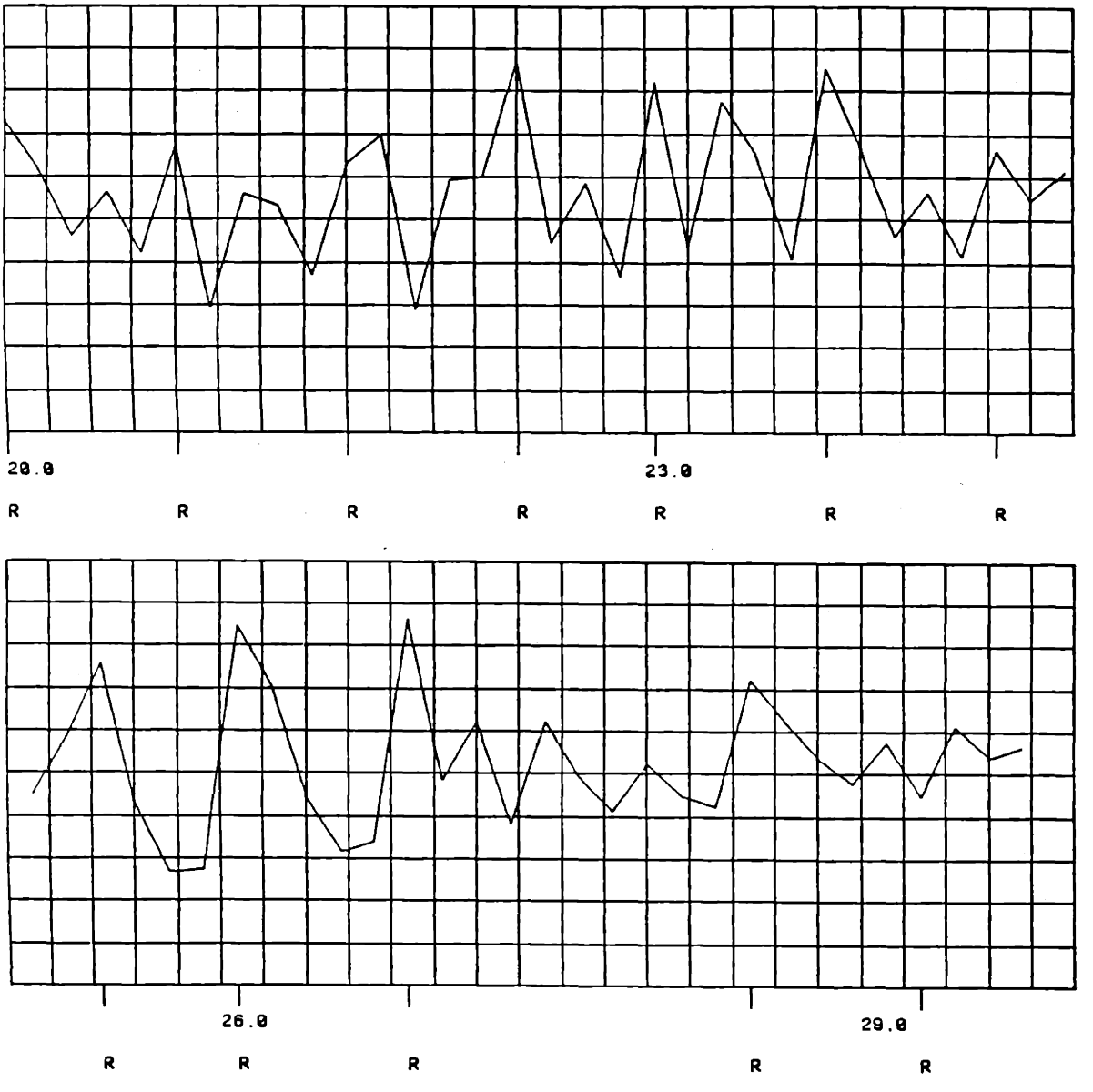


Figure 6.16 Continued.
Part (c): Parameters from Column 8.

		Nonoptimal Estimator			MAP estimator		
		fpR	fnR	teR	fpR	fnR	teR
mean	5	10/1779 =.00562 (.315)	19/1788 =.0106 (.228)	(10+19)/1788 =.0162	18/7159 =.00251 (.235)	62/7203 =.00861 (.126)	(18+62)/7203 .0111
	3	135/1759 =.0767 (.0827)	164/1788 =.0917 (.0744)	(135+164)/1788 =.167	92/1737 =.0530 (.101)	144/1789 .0805 (.0799)	(92+144)/1789 .132
	2	387/1698 =.228 (.0447)	475/1786 =.266 (.0393)	(387+475)/1786 =.483	276/1670 =.165 (.0550)	395/1789 =.221 (.0444)	(276+395)/1789 =.375

Table 6.14 Robustness.
Truth Design Model: Random Excitation to R wave Coupling.
Estimator Design Model: Random Excitation to R wave Coupling.

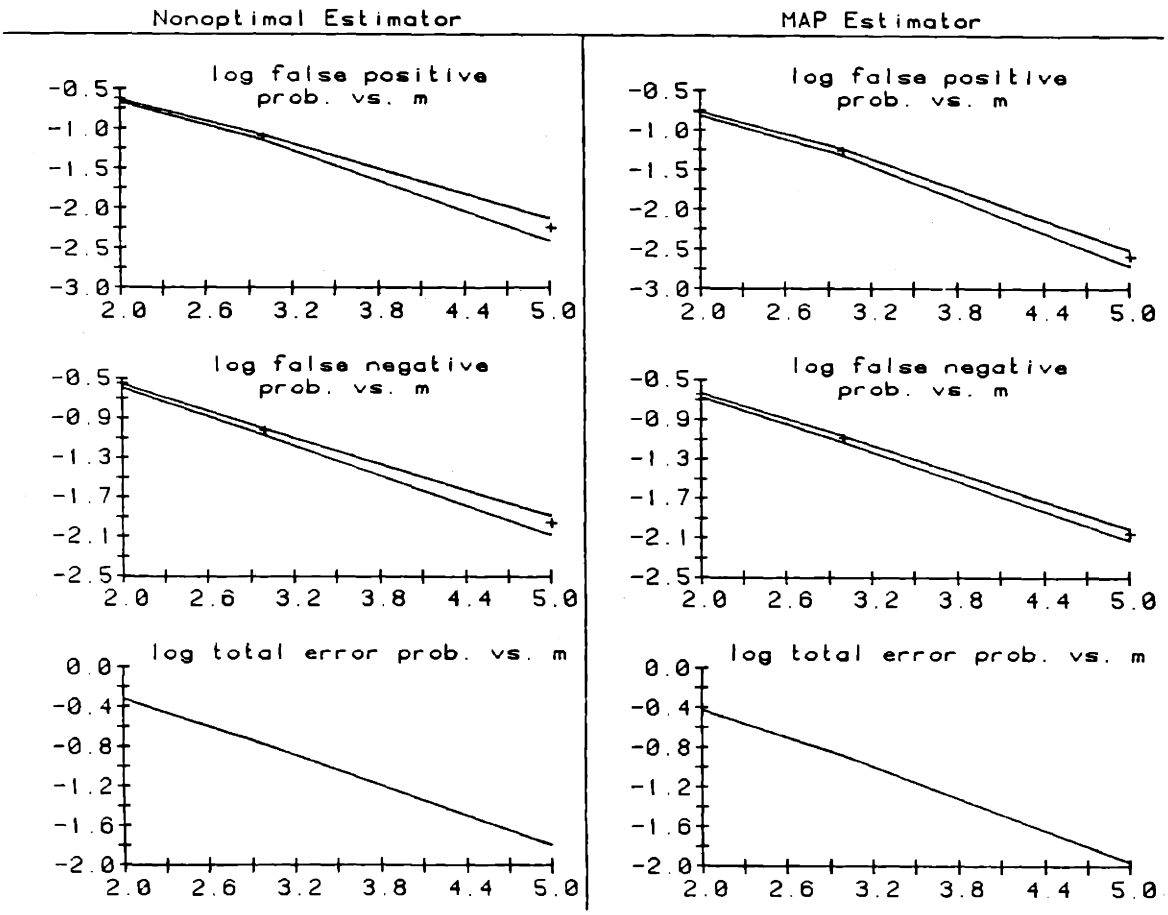


Figure 6.17 Plots of the Statistics in Table 6.14.

		Nonoptimal Estimator			MAP estimator		
		fpR	fnR	teR	fpR	fnR	teR
mean	5	17/1770 =.00960 (.241)	21/1774 =.0118 (.217)	(17+21)/1774 =.0214	18/7084 =.00254 (.235)	47/7111 =.00633 (.149)	(18+47)/7111 =.00914
	3	144/1754 =.0821 (.0798)	164/1774 =.0924 (.0744)	(144+164)/1774 =.174	80/1730 =.0462 (.109)	125/1775 =.0704 (.0862)	(80+125)/1775 =.115
	2	400/1731 =.231 (.0438)	441/1772 =.249 (.0413)	(400+441)/1772 =.475	252/1686 =.149 (.0581)	341/1775 =.192 (.0487)	(252+341)/1775 =.334

Table 6.15 Robustness.
Truth Design Model: Deterministic Excitation to R wave Coupling.
Estimator Design Model: Deterministic Excitation to R wave Coupling.

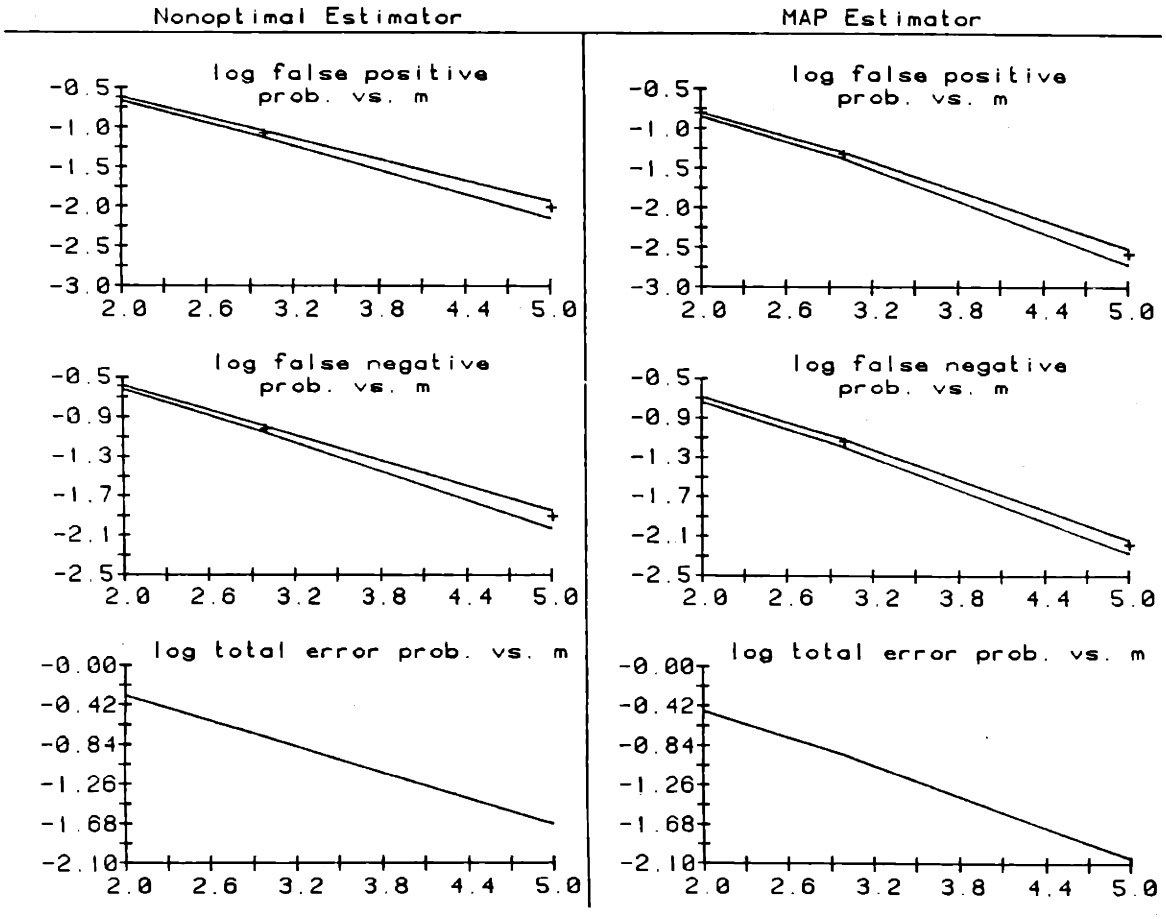


Figure 6.18 Plots of the Statistics in Table 6.15.

		Nonoptimal Estimator			MAP estimator		
		fpR	fnR	teR	fpR	fnR	teR
mean	5	66/7133 =.00925 (.123)	134/7201 =.0186 (.0856)	(66+134)/7201 =.0278	166/7101 =.0234 (.0767)	268/7203 =.0372 (.0599)	(166+268)/7203 =.0603
	3	282/3517 =.0802 (.0571)	361/3596 =.100 (.0499)	(282+361)/3596 =.179	222/3457 =.0642 (.0649)	362/3597 =.101 (.0498)	(222+362)/3597 =.162
	2	360/1677 =.215 (.0467)	465/1782 =.261 (.0399)	(360+465)/1782 =.463	285/1679 =.170 (.0540)	395/1789 =.221 (.0444)	(285+395)/1789 =.380

Table 6.16 Robustness.
Truth Design Model: Random Excitation to R wave Coupling.
Estimator Design Model: Deterministic Excitation to R wave Coupling.

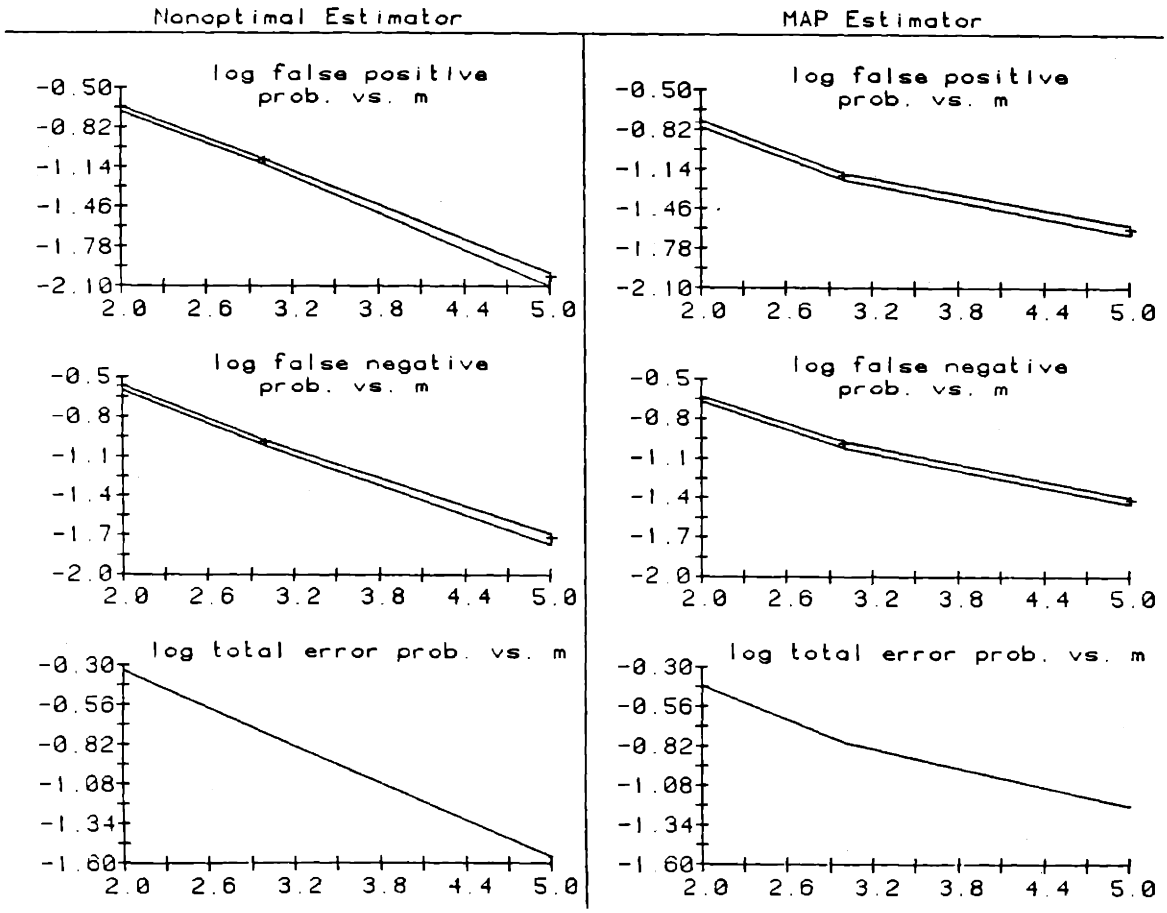


Figure 6.19 Plots of the Statistics in Table 6.16.

		Nonoptimal Estimator			MAP estimator		
		fpR	fnR	teR	fpR	fnR	teR
mean	5	73/7137 =.0102 (.116)	46/7110 =.00647 (.147)	(73+46)/7110 =.0167	19/7080 =.00268 (.229)	50/7111 =.00703 (.141)	(19+50)/7111 =.00970
	3	367/3600 =.102 (.0495)	318/3551 =.0896 (.0535)	(367+318)/3551 =.193	162/3439 =.0471 (.0767)	273/3550 =.0769 (.0581)	(162+273)/3550 =.123
	2	357/1679 =.213 (.0470)	450/1772 =.254 (.0407)	(357+450)/1772 =.455	256/1672 =.153 (.0575)	359/1775 =.202 (.0471)	(256+359)/1775 =.346

Table 6.17 Robustness.
Truth Design Model: Deterministic Excitation to R wave Coupling.
Estimator Design Model: Random Excitation to R wave Coupling.

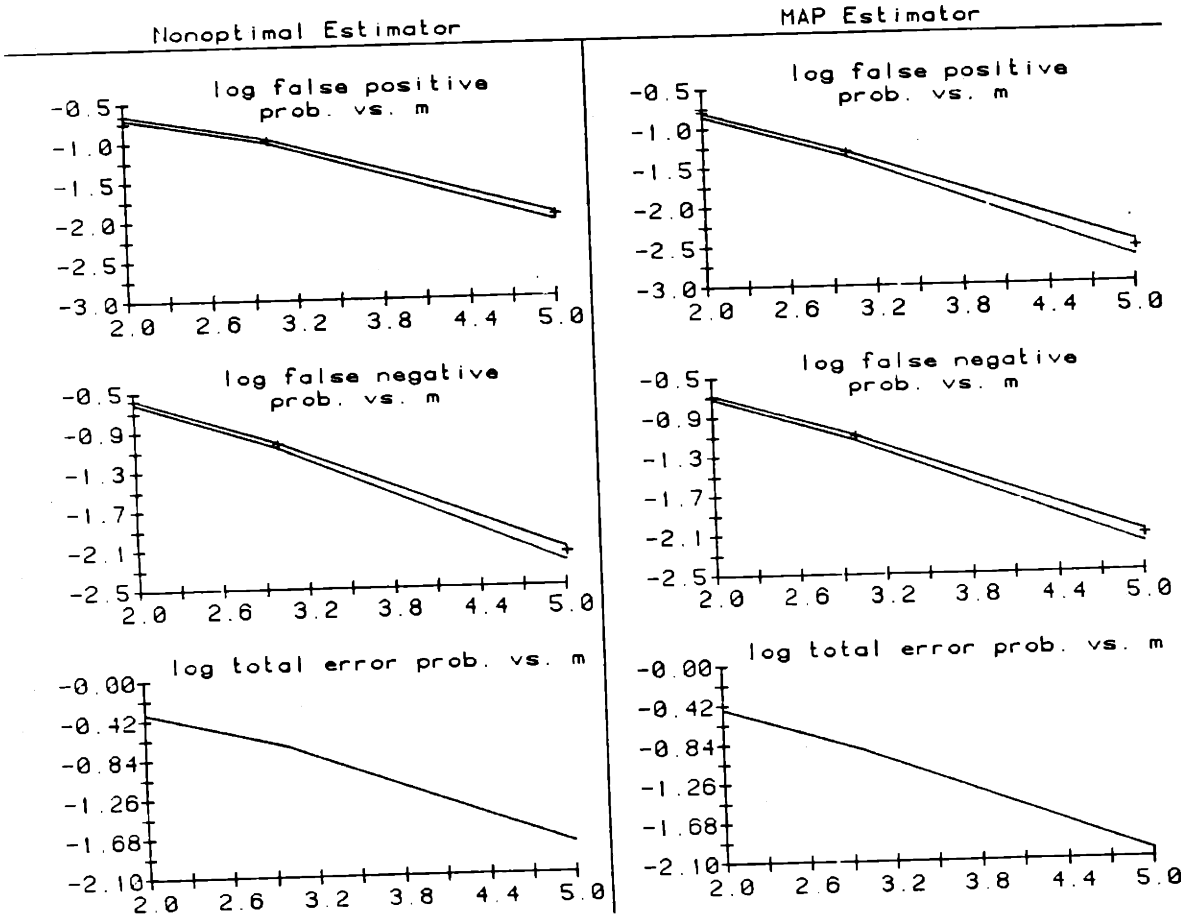


Figure 6.20 Plots of the Statistics in Table 6.17.

6.10. Performance as a Function of Design Model

In this thesis, we present estimation results for just a few design models, though we claim that our estimation approach has general utility. In order to give an idea of the variety of design models considered, we now describe some additional design models that are variants of the design model of Figure 6.3. For each of these design models, the suboptimal estimator of Figure 6.11 has been applied with results similar to those described previously. The four variants are:

(1) Variation in the pmf for the interval between excitations of submodel 1 ($p^0_{3,0}$, $p^0_{4,0}$). (If there were P waves, then this would be the P-P interval pmf).

(2) Variation in the pmf for the interval between excitations of submodel 1 and R waves ($p^1_{0,2}$). (If there were P waves, then this would be the P-R interval pmf).

(3) Variation in the ventricular refractory-period pmf ($p^1_{3,0}$).

(4) Variation in the SNR for all of the above (m_R).

6.11. The P4 Class of Chains

The creation of a P4 chain is described in the context of the initial pass of LE 1. This is the case occurring in the DM2 estimator of Figure 5.6. The case of LE 0 is identical.

The estimation performed by LE 1 would be simplified if we could provide LE 1 with an ECG from which all signatures due to submodel 0 had been removed (i.e. S0). However, due to the stochastic nature of the

signal, this is not possible. Therefore, we introduce the P4 chain. The intended purpose of the P4 chain is to estimate the times when the S0 signature occurs and to use these estimated times in essence to subtract the mean of the S0 signature and adjust the variance of the remaining signal to reflect the additional randomness due to the stochastic portion of the S0 signature. Therefore we call P4 a subtractor chain. (Note that the subtraction and adjustment occur based on an estimated event time, which may be in error).

Note that the P4 chain models the contribution of a single signature to the observation. The P4 chain does not utilize any a posteriori information from the ICS, so it is only used in the initial pass of a LE.

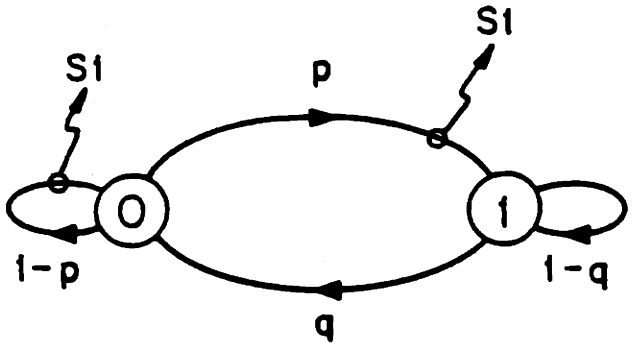
The P4 chain strongly resembles the P0 chain. The difference is that the P4 chain is concerned with the contribution of a signature to the observation while the P0 chain is concerned with the interactions impinging on some second chain. Furthermore, the P4 chain is always autonomous.

In the P0 chain, the subprocess state space was fixed and a variety of schemes for choosing the tpm(s) were proposed. In the P4 case, both the subprocess state space and the tpm are fixed for this thesis. Let $N_{S1}(n)$ be the time of the n^{th} S1 annotation. Let the n^{th} interarrival time $\tau_{S1,S1}(n)$ be defined by $\tau_{S1,S1}(n)=N_{S1}(n+1)-N_{S1}(n)$. Assuming that the design model has a steady-state state distribution, there is a steady-state pmf on $\tau_{S1,S1}$. That is, $\lim_{n\to\infty} Pr(\tau_{S1,S1}(n)=k)$ exists for all k. The basic idea for the P4 chain is to choose the subprocess state space, tpm, and signature-initiating transitions such that the

steady-state expected value of $\tau_{S1,S1}$ in the P4 chain and in the design model are equal.

Implementation of the basic scheme for P4 chains requires a chain with at least two states. Since use of a chain with more that two states requires more computation when implementing the estimator and additional constraints when designing the estimator, we have used exclusively chains with two states.

Even in the two state situation, the ability to independently specify two transition probabilities requires that a second design constraint be imposed. We have chosen to require that the value of $\lim_{n\to\infty} \Pr(\tau_{S1,S1}(n)=1)$ in the P4 chain equal the value in the design model. Therefore, the steady-state distribution on $\tau_{S1,S1}$ is matched exactly at short interarrival times and in an average sense at long interarrival times. The two-state P4 chain which satisfies these constraints is shown in Figure 6.21.



$$N_{S1}(n) = \text{time of } n^{th} \text{ S1 annotation}$$

$$\tau_{S1,S1}(n) = N_{S1}(n+1)-N_{S1}(n)$$

$$p = 1-\lim_{n\to\infty} Pr(\tau_{S1,S1}(n)=1)$$

$$q = \frac{p}{\lim_{n\to\infty} E\tau_{S1,S1}(n)-1}$$

where

$$E\tau_{S1,S1}(n)=\sum_{k=1}^{\infty} kPr(\tau_{S1,S1}(n)=k).$$

Figure 6.21 Definition of the P4 Chain.

6.12. The P3 Class of Chains

The creation of a P3 chain is described in the context of pass 1 of LE 0, as in the DM2 estimator of Figure 5.6. From the initial pass of LE 1, it is possible to extract estimates of the times when S1 annotations occur. The purpose of the P3 chain is to use this information,

received through the ICS, to assist pass 1 of LE 0. Therefore, the P3 and P4 chains share a common purpose. For this reason, the P3 chain is also called a subtractor chain. The only difference between the P3 and P4 chains is that the P3 chain incorporates a posteriori information.

The P3 chain strongly resembles the P2 chain. It is an autonomous chain with a single pseudo-observation, denoted Z4, which conveys the information provided by the ICS. The information received from the ICS is a (perhaps phase-shifted) estimate of the binary annotations for signature S1. Let the binary annotation for signature S1 at time n be denoted $a^{S1}(n)$. Recall from Section 5.5 that this signal is defined as

$$a^{S1}(n)=\begin{cases}1 & \text{if an S1 annotation occurred at time n}\\ 0 & \text{otherwise}\end{cases}.$$

Each state of the P3 subprocess state space represents a particular state of a shift register memory. The contents of the shift register are the binary annotations. Denote the length of the shift register as K. Let $b_0^i,\ldots,b_{K-1}^i$ be the shift register contents represented by state i. Analogous to the P2 case, the i,j element of the tpm for P3 is calculated as

$$\lim_{n\to\infty}\Pr(a^{S1}(n+1)=b_{K-1}^j,\ldots,a^{S1}(n+1-K+1)=b_0^j \mid a^{S1}(n)=b_{K-1}^i,\ldots,a^{S1}(n-K+1)=b_0^i).$$

Denote this probability $p_{i,j}^{P3}$. As in the P1 and P2 cases, a necessary condition for $p_{i,j}^{P3}>0$ is $b_{k-1}^j=b_k^i$, $k\in\{1,\ldots,K-1\}$.

A P3 chain initiates Gaussian signatures which, though they may be initiated by several different transitions, are all identical. In the context of LE 0 pass 1, this signature is the S1 signature. Thus we need only to describe which transitions initiate the signatures. Let $K^* \in \{0,\ldots,K-1\}$ be a distinguished element in the shift register memory. (K^* is introduced to allow for the possibility of noncausality). A transition from i to j initiates a Gaussian signature if $p^{P3}_{i,j}>0$ and $b^i_{K^*}=1$.

As in previous notation, let Z4(n) be the random variable representing the estimated binary annotation received through the ICS at time n, and let $Z4_n$ be its realization. In light of the previous description of the subprocess state space of P3, the observation pmf (for the transition from state i to state j) has the form

$$\Pr(Z4(n)=Z4_n | b^j_{K-1}, b^i_{K-1}, b^i_{K-2}, \ldots, b^i_0).$$

As in the P1 and P2 cases, the duration of the model's memory is K+1 in spite of the fact that the duration of the shift register memory is only K. In all the work reported on in this thesis, this set of pmfs is calculated using simulation.

Throughout the work reported in this thesis, K is two; K^* is one; and when forming Z4(n), the estimated binary interactions are not shifted. That is, $Z4(n)=\hat{a}^{S1}(n)$ where $\hat{a}^{S1}(n)$ is the estimate of $a^{S1}(n)$. Therefore the observation is centered in a three element window of binary annotations. This completes the discussion of the P3 type of chain, the last of the five generic estimator chains (i.e. P0, ..., P4)

to be presented.

6.13. Numerical Results for Design-Model Class DM2--Introduction

Numerical results for design models belonging to class DM2 were used to address three issues:

(1) the definition of the augmentation for interactions transmitted from LE 0 pass 1 to LE 1 pass 2 (recall that this issue, which also applies to DM1 design models, was not discussed in the preceding sections because historically it was a later innovation);

(2) the necessity, in models motivated by the ECG problem, of having subtractor chains (i.e. P3 and P4);

(3) the performance of both the suboptimal and the global MAP estimators for three design models which do not exhibit the determinism of the design models considered previously (e.g. those shown in Figure 6.4); and

(4) the consistency of the global estimate.

The basic design model is shown in Figure 6.22. Several different sets of parameters are shown in Table 6.18. The models of Figure 6.22 and Table 6.18 columns 1-4 are models of normal rhythm, including both the P and the R waves. The models of Figure 6.22 and Table 6.18 columns 4-7 are design models which do not exhibit the determinism of the design models shown in Figure 6.4.

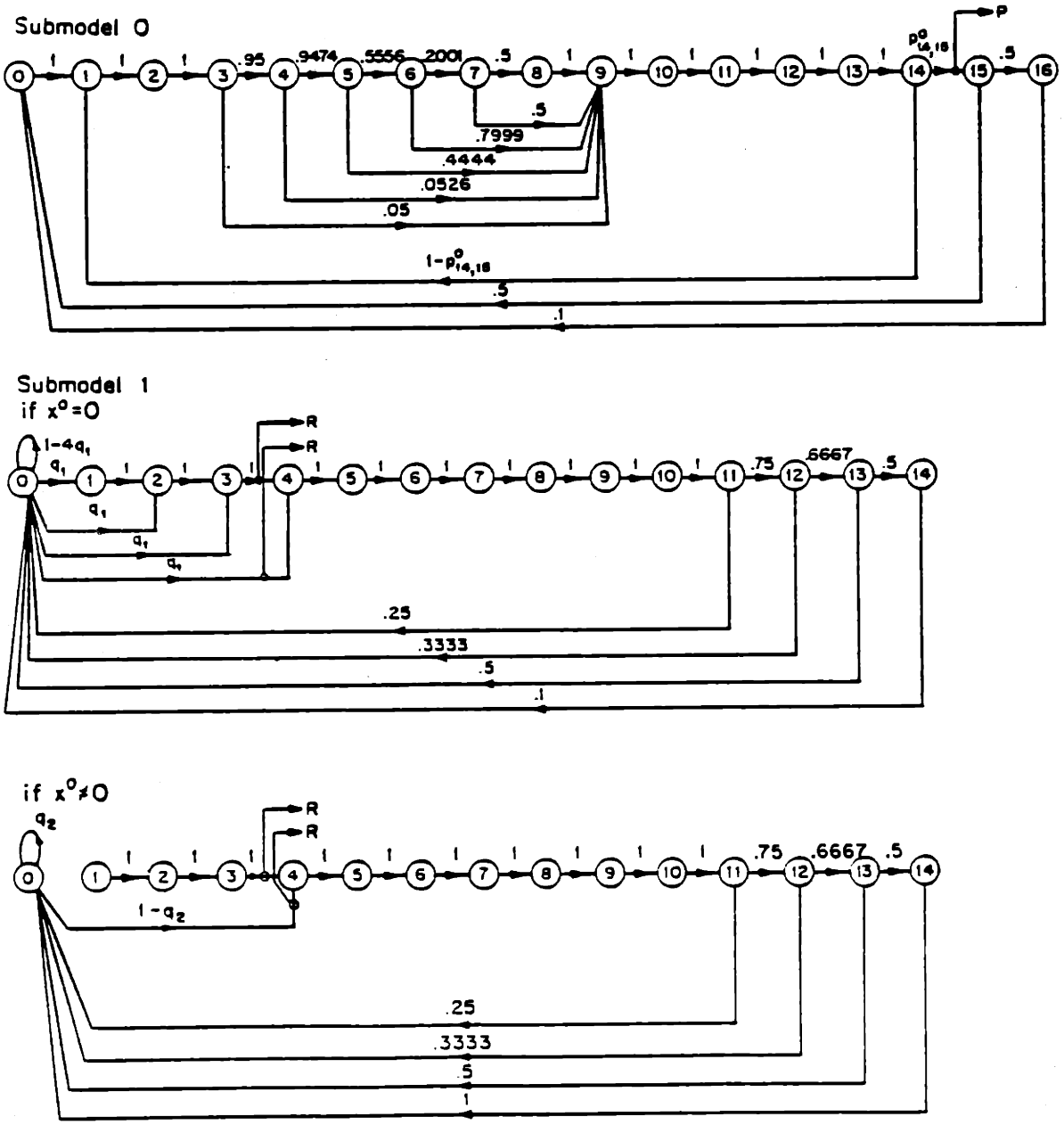


Figure 6.22 A Normal Rhythm Design Model With Both P and R Waves.

Markov chain cycle interval = signature sampling interval = 1 (normalized time).

P: 2 samples long, each sample having mean m_p and standard deviation $\sqrt{.1}$.

R: 3 samples long, with means m_1, m_2, and m_3 and common standard deviation $\sqrt{.5}$.

Observation noise: mean 0 and standard deviation 1.

Figure 6.22 Continued.

		Model Number						
		1	2	3	4	5	6	7
parameter	$p^0_{14,15}$	1	1	1	1	.95	1	1
	q_1	.25	.25	.25	.25	.25	.25	.2375
	q_2	1	1	1	1	1	.95	1
	m_p	.6	.6	.9	.9	.6	.6	.6
	m_1	.5	.75	.5	.75	.75	.75	.75
	m_2	3	4.5	3	4.5	4.5	4.5	4.5
	m_3	.5	.75	.5	.75	.75	.75	.75

Table 6.18 Parameters for the Design Model of Figure 6.22.

6.14. Augmented Interactions--LE 0 Pass 1 to LE 1 Pass 2

Consider the design model of Figure 6.22 and Table 6.18 column 1. In these design models, signature S0 (2 Markov-chain cycles in duration) corresponds to the low SNR P wave while signature S1 (3 Markov-chain cycles in duration) corresponds to the high SNR R wave. A sample of the stochastic process defined by this design model is shown in Figure 6.23. Two different estimators for DM2-class design models are shown in Figures 5.6 and 6.24. The only difference between these two estimators is whether the ICS link between LE 0 pass 1 and LE 1 pass 2 transmits augmented or unaugmented interaction estimates.

Statistics on the augmented interaction estimates for LE 0 pass 1 in the estimator of Figure 5.6 (which are identical to those in the estimator of Figure 6.24 since these two estimators only differ in later passes) are shown in Table 6.19. The corresponding statistics for unaugmented interactions, shown in Table 6.20, were obtained from the statistics in Table 6.19 by summing up the columns. Note first that the

estimated interaction $^{0,1}\hat{h}_n^{C0,P1}$ is redundant if the estimated state $^{0,1}\hat{x}_n^{C0}$ is known. This true because $^{0,1}\hat{h}_n^{C0,P1}$ is defined as

$$^{0,1}\hat{h}_n^{C0,P1} = \begin{cases} 0 \text{ if } {}^{0,1}\hat{x}_n^{C0}=0 \\ 1 \text{ otherwise} \end{cases}.$$

This explains the all-zero rows in Table 6.19.

The reason for considering the augmented interactions is again one of phase errors. Consider the (1,0,1) column in Table 6.19. Assume that the LE 0 pass 1 estimator is doing fairly well. Then most of the time, $^{0,1}\hat{x}_n^{C0}=0$ when truth interaction 0 occurs (i.e. the actual interaction is $h_n^{0,1}=0$). This explains the 611 entry, which is by far the largest entry in the (1,0,1) column. However, sometimes the estimated state leads the true state by one cycle and in this case there is a corresponding lead between $^{0,1}\hat{x}_n^{C0}=0$ and the truth interaction, producing the 297 entry. A lead of two cycles generates the 109 entry. Similarly there can be lag situations, which lead to the 159 and 171 entries. What is uncommon is a large mistake such as $^{0,1}\hat{x}_n^{C0}=9$ when a truth interaction occurs. Correspondingly, the entries for $^{0,1}\hat{h}_n^{C0,P1}=1$, $^{0,1}\hat{x}_n^{C0}\in\{4,\ldots,13\}$ (still in column (1,0,1)) are all very small. When using unaugmented interactions, the probabilities of the relatively common small-phase errors are summed together with the probabilities of the uncommon large errors. This results in the misleading statistics shown in Table 6.20.

The implications of this can be seen from the following example. Suppose that the interaction estimate that is transmitted is a 1. The LE 1 pass 2 estimator must decide which of the four possible actual triplets (1,1,0), (1,0,1), (0,1,1), and (1,1,1) caused this 1. If the estimator is given only the unaugmented interaction, then, using the statistics of Table 6.20, it assigns conditional probabilities of .792, .573, .805, and .984 respectively to these four possibilities. If, in addition to the interaction estimate, the estimator is also given the augmenting information $^{0,1}\hat{x}_n^{C0}$, then we often obtain far more decisive information. For example, if $^{0,1}\hat{x}_n^{C0}=1$, then from Table 6.19, the estimator assigns conditional probabilities of .0762, .208, .427, and .0264 which are only moderately different from the values with unaugmented interactions. This is the case because, from Table 6.19, we see that the maximum ambiguity occurs when $^{0,1}\hat{x}_n^{C0}=1$. On the other hand, if $^{0,1}\hat{x}_n^{C0}=9$, then we know with virtual certainty that the actual interaction triplet was (1,1,1).

Statistics on the global wave-tracking estimate from the estimators of Figures 5.6 and 6.24 are shown in Tables 6.21 and 6.22. Statistics for the global MAP estimate are also shown in these figures. Table 6.21 shows two misclassification tables for each estimator. The difference between the contents of the two tables is in the annotation matching, as in one case phase-shifted matches are permitted while in the other they are not. Also shown are six sets of false positive, false negative, and total error probabilities for both P and R waves.

The P wave estimates for the two suboptimal estimators are identical because the P wave estimates are constructed from the results of

LE 0 pass 1 which are the same in both estimators. Based on teP in the case when phase-shifted matches are allowed, the global MAP estimator performs (.334-.204)/.334 = 39 percent better than the suboptimal estimators. When considering the utility of this level of performance with respect to the ECG problem, note that the data is unrealistically noisy. The noise level was chosen so that a statistically significant number of errors would occur within a simulation of practical duration. In Section 6.21, we consider a more detailed example with more realistic noise levels.

With respect to the R wave performance of the suboptimal estimators, there are two effects. Both favor the estimator using augmented interaction estimates (i.e. the estimator of Figure 5.6). The two effects are:

(1) The total error probability (phase-shifts allowed) declines by (.0573-.0488)/.0573 = 15 percent.

(2) The imbalance between fpR and fnR declines. In the phase-shift-allowed case, the decline is from 9:65 to 27:36. That is, there is a greatly reduced tendency to reject actual R waves compared to the tendency to produce false detections.

We consider the decreased imbalance of (2) to be desirable for two reasons:

(1) A false positive detection could be found and deleted by further processing which can be limited to the times when waves were detected. On the other hand, retrieving an event that was not detected (i.e. a false negative detection) requires considering

all times at which no R wave was detected, a much larger set of times.

(2) Some false positives and false negatives are pairs. That is, they result from a wave detection too far from the true location to be considered a match, yet still nearby. On the basis of examining isolated plots of data, truth annotations, and estimated annotations, we believe that this is an important error mechanism. We have not, however, defined and collected statistics that would support this assertion. If a majority of the errors are paired, then it is natural to prefer a fpR:fnR ratio of 27:36 rather than 9:65 because the former would correspond to a significantly lower actual error rate (where an error pair is counted as a single error).

Now consider the better suboptimal estimator (i.e. the estimator of Figure 5.6) and the global MAP estimator. Based on teR in the case when phase-shifted matches are allowed, the global MAP estimator performs (.0488-.0348)/.0488 = 29 percent better than the suboptimal estimator.

Table 6.22 shows the joint P and R wave statistics described in Section 4.3.4. The tables for conducted beats are made from the subclass of matched truth and estimated annotations where, in the truth annotations, P is followed by R. The tables for dropped beats are made from the subclass where, in the truth annotations, P is not followed by R. (In these simple design models, that means it was followed by another P). As before, tabulations based on matches in which phase-shifts are and are not allowed are both shown.

The P wave estimates are the same for both suboptimal estimators.

With respect to R waves, the chief difference between the two cases is that (in the phase-shift-allowed tables) 24 (P,fn) pairs are transformed into (P,R) pairs and 5 (fn,fn) pairs are transformed into (fn,R) pairs. This reflects the improved R wave performance seen in the misclassification tables.

Now compare the better suboptimal estimator (i.e. the estimator of Figure 5.6) and the global MAP estimator. The dramatic difference is superior P wave performance in the global MAP estimator. Consider the tables based on matches in which phase shifts are allowed. For conducted beats, 57 of the 166 (fn,R) pairs of the suboptimal estimator are transformed into (P,R) pairs for the global MAP estimator. For dropped beats, 33 of the 54 fn detections in the suboptimal estimator become P detections in the global MAP estimator. While the global MAP estimator provides superior performance, the actual performance of the suboptimal estimator with respect to P waves is not bad. For example, in conducted beats where the R wave was correctly detected, only 166/1254 = 13 percent of the true P waves are not matched to estimated P waves (all of these errors are false negatives). The corresponding numbers for the global MAP estimator are 109/1251 = 8.7 percent.

On the basis of these results, the communication of augmented interaction estimates between LE 0 pass 1 and LE 1 pass 0 was chosen as the standard procedure.

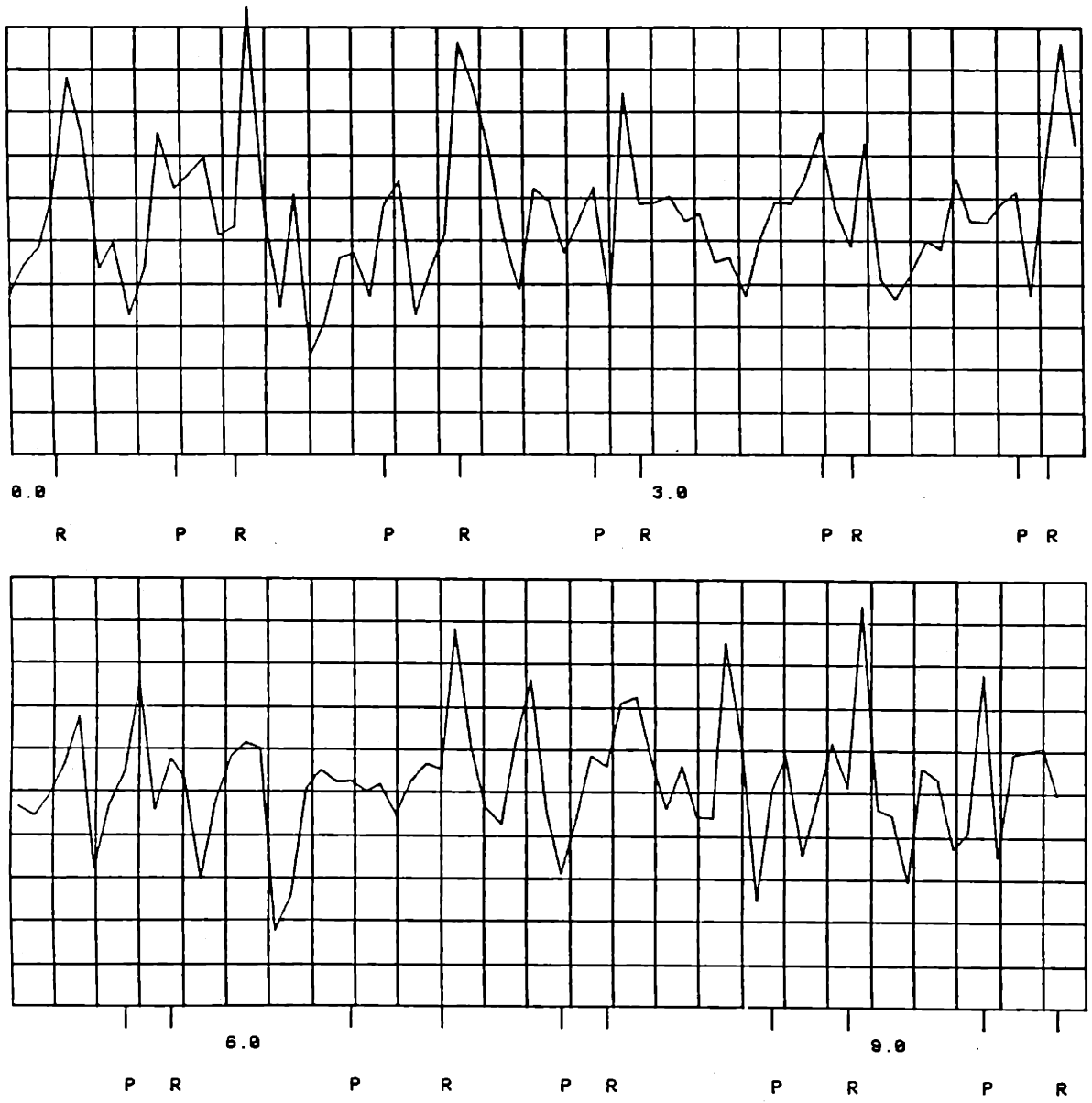


Figure 6.23 A Realization of the Stochastic Process Defined in Figure 6.22 and Table 6.18 Column 1.

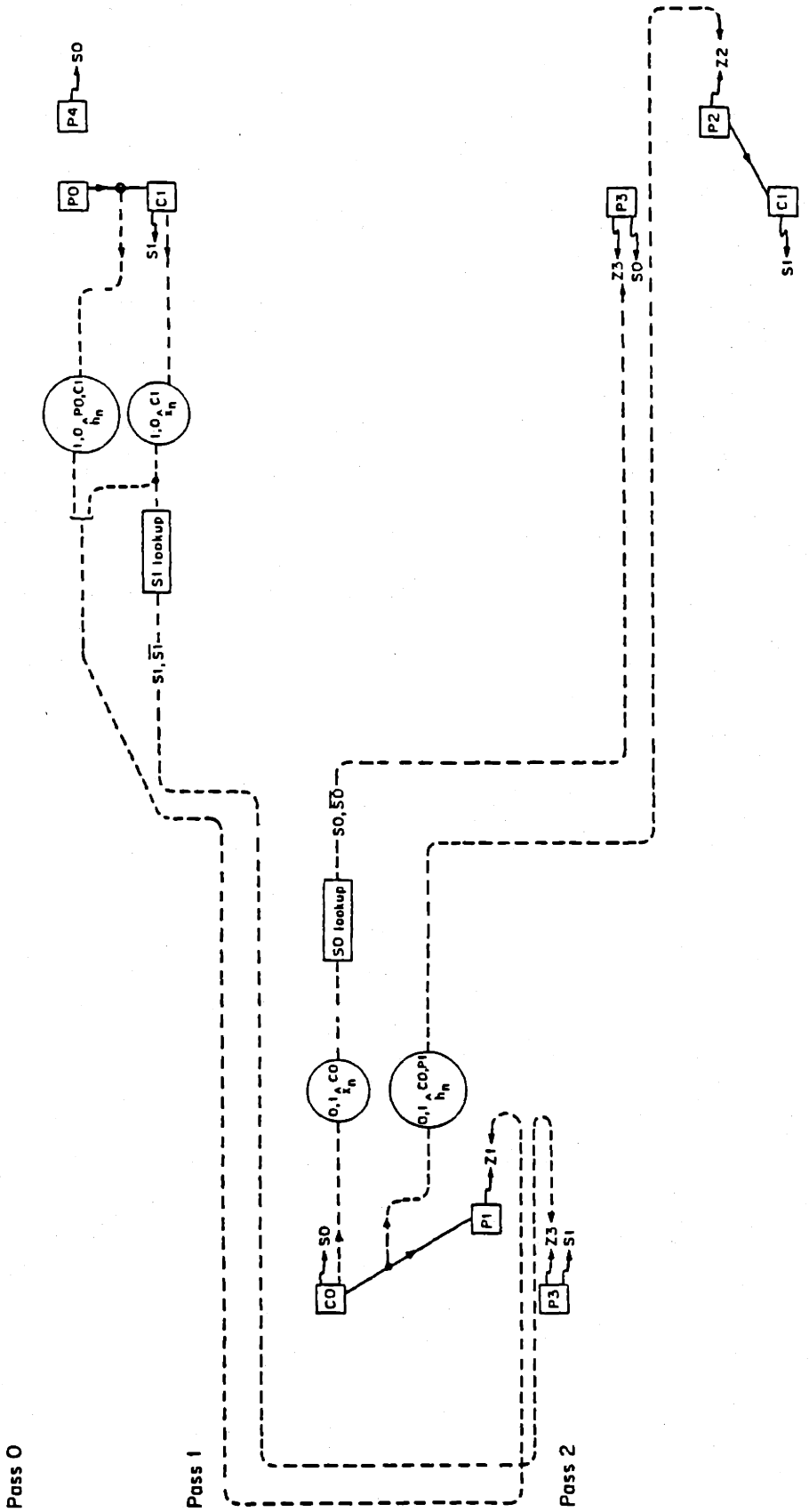


Figure 6.24 Comparison of Augmented and Unaugmented Interactions for Communication Between LE 0 Pass 1 and LE 1 Pass 2. Estimator for the Unaugmented Case.

Section 6.14.

			truth interaction triplet $(h^{0,1}_{n-1}, h^{0,1}_{n}, h^{0,1}_{n+1})$			
			(1,1,0)	(1,0,1)	(0,1,1)	(1,1,1)
estimated interaction $0,1\hat{h}^{C0,P1}_{n}$	0	estimated state $0,1\hat{x}^{C0}_{n}$ 0	297	611	279	244
		1	0	0	0	0
		2	0	0	0	0
		.	.	.	.	.
		.	.	.	.	.
		.	.	.	.	.
		16	0	0	0	0
	1	estimated state $0,1\hat{x}^{C0}_{n}$ 0	0	0	0	0
		1	109	297	611	414
		2	16	109	297	1009
		3	1	16	109	1305
		4	1	1	16	1341
		5	1	1	1	1330
		6	0	0	1	683
		7	0	0	0	63
		8	1	0	0	57
		9	0	2	0	1428
		10	0	0	2	1428
		11	5	0	0	1425
		12	16	5	0	1409
		13	43	16	5	1366
		14	159	43	16	1212
		15	475	159	43	753
		16	307	171	52	189

Table 6.19 Statistics on Augmented Interaction Estimates From LE 0 Pass 1.

		truth interaction triplet $(h^{0,1}_{n-1}, h^{0,1}_{n}, h^{0,1}_{n+1})$			
		(1,1,0)	(1,0,1)	(0,1,1)	(1,1,1)
estimated interaction $0,1\hat{h}^{C0,P1}_{n}$	0	297	611	279	244
	1	1134	820	1153	15412

Table 6.20 Statistics on Interaction Estimates From LE 0 Pass 1 (based on Table 6.19).

Estimator of Figure 5.6

With phase shifts

Truth \ Estimated	P	R	fn	rs
P	1190	9	230	1429
R	2	1255	34	1291
fp	237	18		255
cs	1429	1282	264	

fpP = 239/1429 = .167(.0590)
fnP = 239/1429 = .167(.0590)
teP = (239+239)/1429 = .334
fpR = 27/1282 = .0211(.190)
fnR = 36/1291 = .0279(.164)
teR = (27+36)/1291 = .0488

Without phase shifts

Truth \ Estimated	P	R	fn	rs
P	639	5	785	1429
R	1	1230	60	1291
fp	789	47		836
cs	1429	1282	845	

fpP = 790/1429 = .553(.0238)
fnP = 790/1429 = .553(.0238)
teP = (790+790)/1429 = 1.11
fpR = 52/1291 = .0403(.136)
fnR = 61/1282 = .0476(.125)
teR = (52+61)/1282 = .0881

Estimator of Figure 6.24

With phase shifts

Truth \ Estimated	P	R	fn	rs
P	1190	7	232	1429
R	2	1226	63	1291
fp	237	2		239
cs	1429	1235	295	

fpP = 239/1429 = .167(.0590)
fnP = 239/1429 = .167(.0590)
teP = (239+239)/1429 = .334
fpR = 9/1235 = .00729(.332)
fnR = 65/1291 = .0503(.121)
teR = (9+65)/1291 = .0573

Without phase shifts

Truth \ Estimated	P	R	fn	rs
P	639	3	787	1429
R	1	1208	82	1291
fp	789	24		813
cs	1429	1235	869	

fpP = 790/1429 = .553(.0238)
fnP = 790/1429 = .553(.0238)
teP = (790+790)/1429 = 1.11
fpR = 27/1235 = .0219(.190)
fnR = 83/1291 = .0643(.106)
teR = (27+83)/1291 = .0852

Table 6.21 Estimator Performance.
Misclassification Statistics.

Global MAP Estimator

With phase shifts

Truth \ Estimated	P	R	fn	rs
P	1285	3	143	1431
R	0	1252	40	1292
fp	146	2		148
cs	1431	1257	183	

fpP = 146/1431 = .102(.0784)
fnP = 146/1431 = .102(.0784)
teP = (146+146)/1431 = .204
fpR = 5/1257 = .00398(.446)
fnR = 40/1292 = .0310(.156)
teR = (5+40)/1292 = .0348

Without phase shifts

Truth \ Estimated	P	R	fn	rs
P	764	2	664	1430
R	0	1231	61	1292
fp	667	24		691
cs	1431	1257	725	

fpP = 667/1431 = .466(.0283)
fnP = 666/1430 = .466(.0283)
teP = (667+666)/1430 = .932
fpR = 26/1257 = .0207(.194)
fnR = 61/1292 = .0472(.125)
teR = (26+61)/1292 = .0673

Table 6.21 Continued.

Estimator of Figure 5.6

Conducted Beats

With phase shifts

		Est. Anno. Matched With Truth R Anno.			
		P	R	fn	rs
Est. Anno.	P	0	1088	22	1110
	R	0	0	5	5
Matched With Truth P Anno.	fn	2	166	7	175
	cs	2	1254	34	

Without phase shifts

		Est. Anno. Matched With Truth R Anno.			
		P	R	fn	rs
Est. Anno.	P	0	583	18	601
	R	0	0	4	4
Matched With Truth P Anno.	fn	1	646	38	685
	cs	1	1229	60	

Dropped Beats

With phase shifts

Est. Anno. Matched With Truth P Anno.			
P	R	fn	rs
80	4	54	138

Without phase shifts

Est. Anno. Matched With Truth P Anno.			
P	R	fn	rs
38	1	99	138

Table 6.22 Estimator Performance.
Joint P and R Wave Statistics.

Estimator of Figure 6.24

Conducted Beats

With phase shifts

		Est. Anno. Matched With Truth P	R	fn	rs
Est. Anno.	P	0	1064	46	1110
	R	0	0	4	4
Matched With Truth P Anno.	fn	2	161	13	176
	cs	2	1225	63	

Without phase shifts

		Est. Anno. Matched With Truth P	R	fn	rs
Est. Anno.	P	0	569	32	601
	R	0	0	3	3
Matched With Truth P Anno.	fn	1	638	47	686
	cs	1	1207	82	

Dropped Beats

With phase shifts

Est. Anno. Matched With Truth P	R	fn	rs
80	3	55	138

Without phase shifts

Est. Anno. Matched With Truth P	R	fn	rs
38	0	100	138

Table 6.22 Continued.

MAP Estimator

Conducted Beats

With phase shifts

		Est. Anno. Matched With Truth R Anno.			
		P	R	fn	rs
Est. Anno. Matched With Truth P Anno.	P	0	1142	25	1167
	R	0	0	1	1
	fn	0	109	13	122
	cs	0	1251	39	

Without phase shifts

		Est. Anno. Matched With Truth R Anno.			
		P	R	fn	rs
Est. Anno. Matched With Truth P Anno.	P	0	671	23	694
	R	0	0	1	1
	fn	0	559	37	596
	cs	0	1230	61	

Dropped Beats

With phase shifts

Est. Anno. Matched With Truth P Anno.			
P	R	fn	rs
117	2	21	140

Without phase shifts

Est. Anno. Matched With Truth P Anno.			
P	R	fn	rs
70	1	68	139

Table 6.22 Continued.

6.15. The P3 and P4 Chains

In this section we present results indicating that the subtractor chains (i.e. P3 and P4 chains) are not necessary for the ECG-motivated design models that we have considered. While this may seem clear for the low-energy P waves, it is somewhat surprising for the high-amplitude R waves.

In Chapter 5 we proposed a estimator, shown in Figure 5.6, for the DM2 design-model class. Here we call that estimator the nominal estimator and consider five variations, where each variation corresponds to

deleting one or more of the P3 or P4 chains. Since these chains are autonomous, no further change in the estimator is required. The five variants are listed in Table 6.23.

Consider the design model of Figure 6.22 and Table 6.18 column 4. A sample of the stochastic process defined by this design model is shown in Figure 6.25. The global wave-tracking estimate for the nominal estimator of Figure 5.6 and for each of the five variants described in Table 6.23 were computed. The global wave-tracking estimate of the global MAP estimator was also computed. The misclassification statistics for these seven estimators, based on annotation matches in which phase shifts are allowed, are shown in Table 6.24. These statistics indicate that all six suboptimal estimators provide equal performance. The performance of the global MAP estimator is provided for comparison. The misclassification statistics based on annotation matches in which phase shifts are not allowed and all of the joint P and R wave statistics indicate the same conclusion. Based on these results, we have done most calculations using estimators without subtractor chains.

The lack of subtractor submodels means that each pass of each LE processes data in which there are unmodeled signals. In the case of LE 0 pass 1, the unmodeled signal is the R wave which contains much more energy than the modeled P wave signal. The equal performance achieved by estimators with and without this subtractor therefore requires explanation.

We believe that it is the presence of pseudo-observations in LE 0 pass 1 that makes the subtractor chain superfluous. Intuitively, the pseudo-observations create a window, for the LE, during which the

estimator essentially cannot estimate a P wave, and (when LE 1 pass 0 is working well) this window usually coincides with the R wave. To see this, assume that the augmented interaction estimates are exact and that we do not round-up zeros in the pseudo-observation observation pmfs. Therefore, at each truth interaction, $^{0,1}\hat{x}_n^{C0}$ must be in state 0. Therefore, from the C0 tpm, the P wave must have occurred one or two samples earlier and the R wave must occur one to four samples later, or not at all (dropped beat). Therefore, it is not possible to put P waves on R waves. Rather, the R waves must be accepted as extraordinary noise bursts.

In the actual estimator, we round-up the zero elements in the pseudo-observation observation pmf and the augmented interaction estimates are sometimes in error. First consider the effect of round-up. Round-up does not change the situation qualitatively because:

(1) We consider noisy problems where the variance of the Gaussian observation noise is large compared to the R wave mean.

(2) We do not round-up very much.

Therefore, it is typically favorable for LE 0 pass 1 to take the R wave as noise.

Now consider errors in the augmented interaction estimates. The key is that the augmented interaction estimates are infrequently so badly in error that the window misses the R wave. Furthermore, if there is just an isolated large error, the pmf for the P-P interarrival time (which is contained implicitly in the tpm for chain C0) makes placing a P wave on the R wave that fell outside of the window a low probability

event. That is, matching a true P, R, and P wave sequence with three distinct, sequential, estimated P waves gives a very irregular, and therefore low probability, sequence of P-P intervals. Therefore, problems with the augmented interaction estimates give only a small contribution to the total error rate.

This explanation is supported by computations, to be described in Section 6.21, on a more realistic model. In that example, the SNR is much improved so that the probability of a noise burst that looks like an R wave is extraordinarily small. In this case, with the zero elements of the pseudo-observation observation pmfs rounded up, accepting low probability pseudo-observations was more favorable than interpreting R waves as noise bursts. The result is that the P wave estimates essentially exactly track the R waves (and also T waves--see the later discussion). However, by decreasing the round-up of the zero elements in the pseudo-observation observation pmfs (we actually left them at zero) we are able to force the P wave estimates onto the P waves leaving the R waves (and T waves) as extraordinary noise bursts. Note that one conclusion of this example is that pseudo-observation observation pmfs must be set consistently with respect to ECG SNR.

		LE 1 Pass 0 has P4?	LE 0 Pass 1 has P3?	LE 1 Pass 2 has P3?
standard		yes	yes	yes
variant	1	no	no	no
	2	no	yes	no
	3	no	yes	yes
	4	yes	no	yes
	5	yes	yes	yes

Table 6.23 Subtractor Chains in the Standard and in Each Variant Estimator.

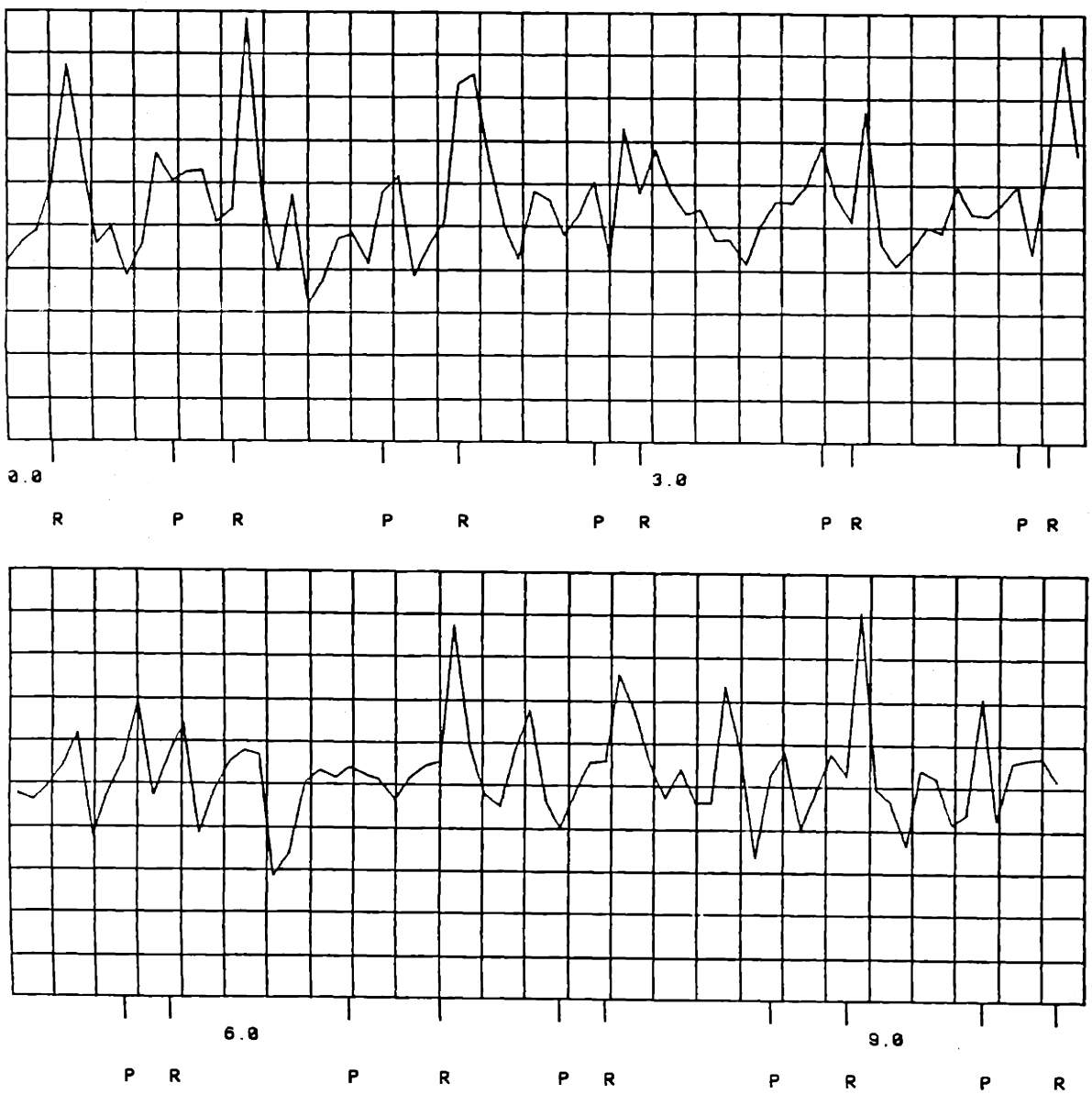


Figure 6.25 A Realization of the Stochastic Process Defined in Figure 6.22 and Table 6.18 Column 4.

Standard

Truth	Estimated P	R	fn	rs
P	1226	15	188	1429
R	4	1252	35	1291
fp	198	21		219
cs	1428	1288	223	

fpP = 202/1428 = .141(.0652)
fnP = 203/1429 = .142(.0650)
teP = (202+203)/1429 = .283
fpR = 36/1288 = .0280(.164)
fnR = 39/1291 = .0302(.158)
teR = (36+39)/1291 = .0581

Variant 1

Truth	Estimated P	R	fn	rs
P	1228	13	189	1430
R	4	1255	33	1292
fp	196	23		219
cs	1428	1291	222	

fpP = 200/1428 = .140(.0656)
fnP = 202/1430 = .141(.0652)
teP = (200+202)/1430 = .281
fpR = 36/1291 = .0279(.164)
fnR = 37/1292 = .0286(.162)
teR = (36+37)/1292 = .0565

Variant 2

Truth	Estimated P	R	fn	rs
P	1227	14	189	1430
R	4	1255	33	1292
fp	197	23		220
cs	1428	1292	222	

fpP = 201/1428 = .141(.0654)
fnP = 203/1430 = .142(.0650)
teP = (201+203)/1430 = .283
fpR = 37/1292 = .0286(.162)
fnR = 37/1292 = .0286(.162)
teR = (37+37)/1292 = .0573

Variant 3

Truth	Estimated P	R	fn	rs
P	1227	14	188	1429
R	4	1254	33	1291
fp	197	23		220
cs	1428	1291	221	

fpP = 201/1428 = .141(.0654)
fnP = 202/1429 = .141(.0654)
teP = (201+202)/1429 = .282
fpR = 37/1291 = .0287(.162)
fnR = 37/1291 = .0287(.162)
teR = (37+37)/1291 = .0573

Variant 4

Truth	Estimated P	R	fn	rs
P	1227	13	189	1429
R	5	1252	34	1291
fp	197	21		218
cs	1429	1286	223	

fpP = 202/1429 = .141(.0652)
fnP = 202/1429 = .141(.0652)
teP = (202+202)/1429 = .283
fpR = 34/1286 = .0264(.169)
fnR = 39/1291 = .0302(.158)
teR = (34+39)/1291 = .0565

Variant 5

Truth	Estimated P	R	fn	rs
P	1226	15	189	1430
R	4	1253	35	1292
fp	198	21		219
cs	1428	1289	224	

fpP = 202/1428 = .141(.0652)
fnP = 204/1430 = .143(.0648)
teP = (202+204)/1430 = .284
fpR = 36/1289 = .0279(.164)
fnR = 39/1292 = .0302(.158)
teR = (36+39)/1292 = .0580

Table 6.24 Estimator Performance.
Misclassification Statistics.
All six suboptimal estimators have α=6 in LE 1 pass 0.

		MAP Estimated			
		P	R	fn	rs
Truth	P	1332	2	97	1431
	R	1	1259	33	1293
	fp	98	1		99
	cs	1431	1262	130	

fpP = 99/1431 = .0692(.0970)
fnP = 99/1431 = .0692(.0970)
teP = (99+99)/1431 = .138
fpR = 3/1259 = .00238(.577)
fnR = 34/1293 = .0263(.169)
teR = (3+34)/1293 = .0286

Table 6.24 Continued.

6.16. Less Deterministic Design Models

In the models considered previously, there are three prominent kinds of qualitative behavior on which the estimators could depend. The three properties are:

(1) the P-P interval is, with probability one, tightly bounded above and below;

(2) if an excitatory interaction does not occur, then 0 is a trapping state in submodel 1 (that is, no R wave occurs without a causally-related excitatory interaction); and

(3) if $x_n^1=0$ and an excitatory interaction occurs, then, with probability one, an R wave follows.

Since real ECGs often are somewhat more irregular than our design models, it is important to make sure that the performance of the suboptimal estimators does not depend too heavily on the three more-or-less

deterministic properties described above. For this reason we have considered three additional design models. In each design model, one of these three properties is relaxed. The three design models are the design model of Figure 6.22 with parameters from Table 6.18 columns 5-7. The corresponding design model that exhibits all three properties is the design model of Figure 6.22 with parameters from Table 6.18 column 2.

Column 5 is relevant to property (1). In this model, the support of the P-P interval pmf extends out to $+\infty$. This model corresponds to a situation in which the SA node can try to excite the atria and fail. Note that a P wave still precedes all attempts to excite submodel 1.

Column 6 relaxes property (2). R waves can now occur without a causally-related excitation from submodel 0. This model corresponds to a situation in which the ventricles are occasionally excited by ectopic foci. Because the R wave morphology is not changed, the ectopic focus must be relatively proximal (e.g. the Bundle of His).

Column 7 effects property (3). Submodel 1 can now ignore excitation attempts even when submodel 1 is in the resting state (i.e. $x_n^1=0$). This model corresponds to a complete, but intermittent, AV block.

When considering properties (2) and (3) and the effect of the parameters in column (6) and (7), note the different rates with which property (2) is violated versus property (3). Under column (6), every Markov chain cycle during which $x_n^1=0$ and an excitation attempt does not occur has probability $1-q_2=.05$ of violating property (2). Under column (7), every Markov chain cycle during which $x_n^1=0$ and an excitation attempt does occur has probability $1-4q_1=.05$ of violating property (3).

However, the fraction of time during which $x_n^1=0$ and $h_n^{0,1}=1$ is much larger than the fraction of time during which $x_n^1=0$ and $h_n^{0,1}=0$. Therefore, the rates with which properties (2) and (3) are violated are not equal.

The suboptimal estimators have the structure shown in the block diagram of Figure 5.6 (with $\alpha=6$ in the P0 chain of the LEPM for LE 1 pass 0). Except in the case of the suboptimal estimator for column 5, a new estimator (suboptimal or global MAP) was designed for each new design model. Therefore this is not a test of robustness--the estimators and observations are matched. In the case of the suboptimal estimator for column 5, due to limitations in the computer code, the tpms for the P0 and P4 chains of the LEPM for LE 1 pass 0 are not matched. Rather, they are the same as the tpms for the corresponding chains in the suboptimal estimators for columns 1-4, 6, and 7 (with $\alpha=6$ for the P0 chain). In all other respects, the suboptimal estimator for column 5 is matched. This mismatch represents only a small perturbation early in the estimator.

Tables 6.25, 6.26, and 6.27 show statistics on the global wave-tracking performance of both the suboptimal and the global MAP estimators for the three design models. The figures show only misclassification results based on matches in which phase shifts are allowed. The results for other measures of performance are similar. Table 6.21 shows corresponding statistics for the design model of column 2. (Recall that this is the design model in which none of the three properties have been relaxed). The statistics from all of these tables are summarized in Tables 6.28 and 6.29. Table 6.28 shows the estimated total error

probability and an estimate of its fractional standard deviation for P waves and R waves of each design model. Table 6.29 shows the performance ratio between suboptimal and global MAP estimators. Because the pdfs for the ratios are very complex, not even estimates of the standard deviations of the ratios have been computed. However, it is clear that the standard deviations of the ratios, especially for the R waves, is quite large due to the large standard deviations of the numerator and denominator.

First note that relaxing each of the three properties reduces the performance of the global MAP estimator by a moderate amount. Therefore it is not surprising that the suboptimal estimators perform more poorly also. The important point is that the performance of the suboptimal estimators does not decrease dramatically when any of the three properties are relaxed. Therefore, the suboptimal estimators are not relying to an extraordinary degree on any of the three properties.

Nonoptimal

Truth	Estimated P	R	fn	rs
P	1134	5	225	1364
R	0	1209	31	1240
fp	289	32		321
cs	1423	1246	256	

MAP

Truth	Estimated P	R	fn	rs
P	1206	3	159	1368
R	0	1203	40	1243
fp	203	2		205
cs	1409	1208	199	

fpP = 289/1423 = .203(.0525)
fnP = 230/1364 = .169(.0601)
teP = (289+230)/1364 = .380
fpR = 37/1246 = .0297(.162)
fnR = 31/1240 = .0250(.177)
teR = (31+37)/1240 = .0548

fpP = 203/1409 = .144(.0649)
fnP = 162/1368 = .118(.0738)
teP = (203+162)/1368 = .267
fpR = 5/1208 = .00414(.446)
fnR = 40/1243 = .0322(.156)
teR = (5+40)/1243 = .0362

Table 6.25 Estimator Performance With the Parameters of Table 6.18 Column 5.

Nonoptimal

Truth	Estimated P	R	fn	rs
P	1064	14	351	1429
R	4	1298	29	1331
fp	365	8		373
cs	1433	1320	380	

MAP

Truth	Estimated P	R	fn	rs
P	1181	9	240	1430
R	7	1294	31	1332
fp	242	3		245
cs	1430	1306	271	

fpP = 369/1433 = .258(.0449)
fnP = 365/1429 = .255(.0452)
teP = (369+365)/1429 = .514
fpR = 22/1320 = .0167(.211)
fnR = 33/1331 = .0248(.172)
teR = (22+33)/1331 = .0413

fpP = 249/1430 = .174(.0576)
fnP = 249/1430 = .174(.0536)
teP = (249+249)/1430 = .348
fpR = 12/1306 = .00919(.287)
fnR = 38/1332 = .0285(.160)
teR = (12+38)/1332 = .0375

Table 6.26 Estimator Performance With the Parameters of Table 6.18 Column 6.

Nonoptimal

Truth \ Estimated	P	R	fn	rs
P	1162	12	254	1428
R	5	1184	35	1224
fp	263	19		282
cs	1430	1215	289	

MAP

Truth \ Estimated	P	R	fn	rs
P	1244	3	184	1431
R	3	1186	37	1226
fp	184	9		193
cs	1431	1198	221	

fpP = 268/1430 = .187(.0551)
fnP = 266/1428 = .186(.0553)
teP = (268+266)/1428 = .374
fpR = 31/1215 = .0255(.177)
fnR = 40/1224 = .0327(.156)
teR = (31+40)/1224 = .0580

fpP = 187/1431 = .131(.0682)
fnP = 187/1431 = .131(.0682)
teP = (187+187)/1431 = .261
fpR = 12/1198 = .0100(.287)
fnR = 40/1226 = .0326(.156)
teR = (12+40)/1226 = .0424

Table 6.27 Estimator Performance With the Parameters of Table 6.18 Column 7.

		Nonoptimal Estimator		MAP estimator	
		teP	teR	teP	teR
model	2	.334(.0373)	.0488(.123)	.204(.0522)	.0348(.146)
	5	.380(.0345)	.0548(.118)	.267(.0448)	.0362(.146)
	6	.514(.0357)	.0413(.132)	.348(.0362)	.0375(.139)
	7	.374(.0342)	.0580(.155)	.261(.0444)	.0424(.136)

Table 6.28 Summary of Performance.
Total Error Probabilities.

	Ratio: Nonoptimal/MAP	
	teP	teR
model 2	1.64	1.40
model 5	1.43	1.51
model 6	1.47	1.10
model 7	1.43	1.37

Table 6.29 Summary of Performance.
Ratios of Total Error Probabilities: Suboptimal Estimator Performance Divided By Global MAP Estimator Performance.

6.17. Consistency of the Global Estimate

In this section we consider the consistency of the global state-trajectory estimate on which the global wave-tracking estimate is based. The subject of consistency was discussed in the abstract previously in Section 5.3.

We have measured consistency in two senses. The first sense is the probability that the global state-trajectory occupies a state in the transient subset of the global state space. The initial conditions are chosen to be in the persistent subset of the global state space. Therefore, in order to occupy a state in the transient subspace, the state must have made a zero-probability transition. The probability we compute is a measure of the amount of time spent in the transient subspace, so it is a function of the number of transitions into the transient subspace and the duration of each stay.

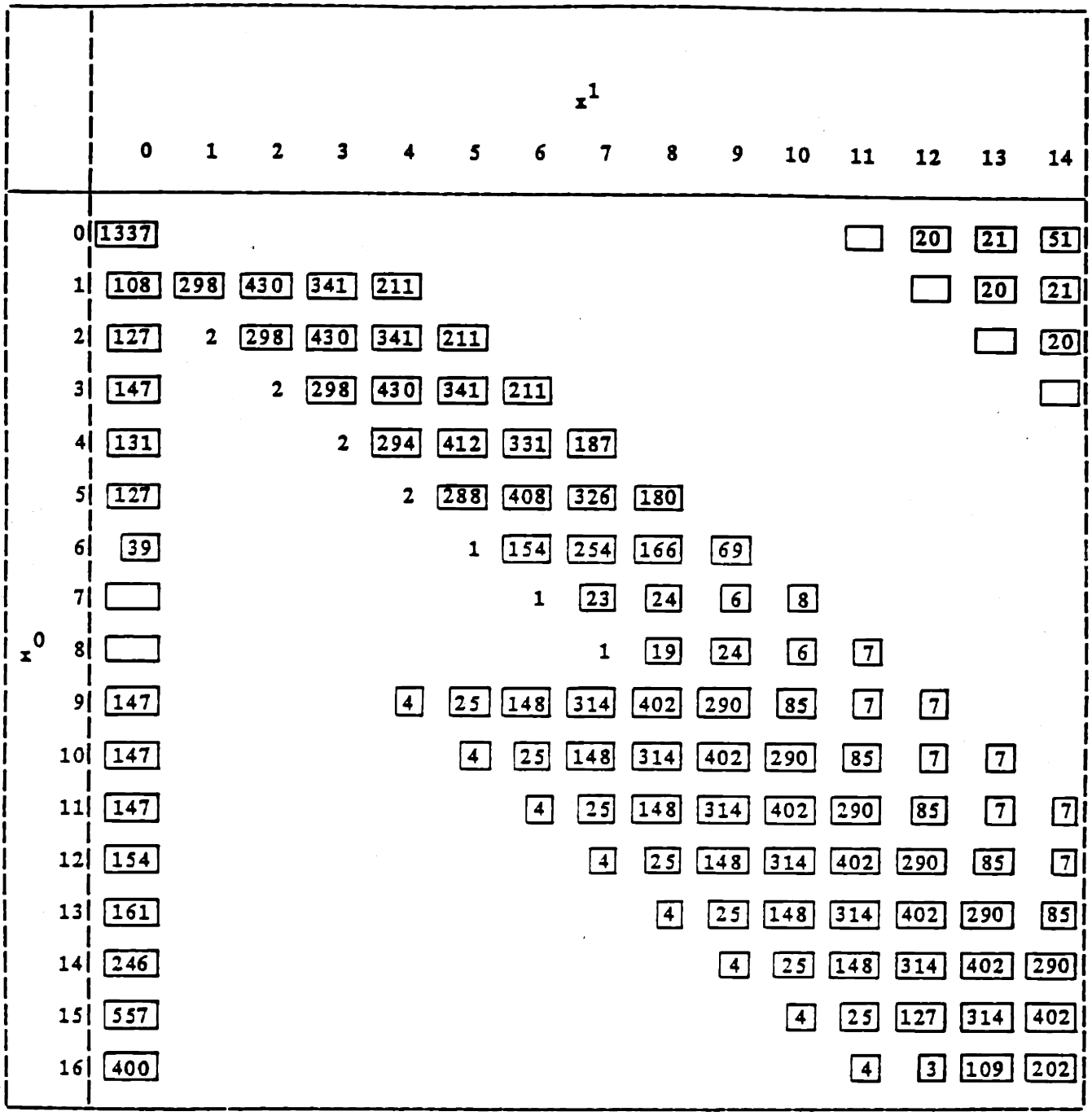
The second sense in which we have measured consistency is the probability that a global state transition in the global state-trajectory estimate has zero probability. Note that the zero-probability transition could be between two states both in the persistent subspace, a situation not reflected in the first measure of consistency.

Consider the design model of Figure 6.22 and Table 6.18 column 2. Apply the estimator with the block diagram shown in Figure 5.6 (α=6 in the P0 chain of LE 1 pass 0) to compute the global state-trajectory estimate.

To evaluate consistency in the first sense, consider Figure 6.26. This is a two dimensional representation of the global state space of

the design model. The persistent states are shown as boxes. The number of times that the global state-trajectory estimate enters each global state is indicated if it is positive. Therefore, the sum of the numbers outside of the boxes is the number of times the global state-trajectory estimate visited a transient state. There are 11 such Markov chain cycles out of a total of 19927. Therefore, the probability that the global state trajectory estimate occupies a transient state is 11/19927 = .000552. In this sense, the estimate is very consistent. (Independently of interest in Figure 6.26 is the sparseness of the persistent global states. Only 116 out of 15x17 = 255 global states are persistent).

To evaluate consistency in the second sense, the number of zero-probability global state transitions in the global state-trajectory estimate were counted. This number was found to be 59 out of 19926 total global transitions. Therefore, the probability that the global state-trajectory estimate makes a zero-probability transition is 59/19926 = .00296. [Of the 59 illegal transitions, 57 were (0,0) to (1,0) and 2 were (1,0) to (2,1)]. In this second sense the estimate is also very consistent. In light of these results, we do not discuss consistency further.



x^0 \ x^1	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
0	1337												20	21	51
1	108	298	430	341	211									20	21
2	127	2	298	430	341	211									20
3	147		2	298	430	341	211								
4	131			2	294	412	331	187							
5	127				2	288	408	326	180						
6	39					1	154	254	166	69					
7							1	23	24	6	8				
8								1	19	24	6	7			
9	147				4	25	148	314	402	290	85	7	7		
10	147					4	25	148	314	402	290	85	7	7	
11	147						4	25	148	314	402	290	85	7	7
12	154							4	25	148	314	402	290	85	7
13	161								4	25	148	314	402	290	85
14	246									4	25	148	314	402	290
15	557										4	25	127	314	402
16	400											4	3	109	202

Figure 6.26 Consistency, Definition 1.

6.18. Numerical Results for Design Model Class DM3--Introduction

The following sections use numerical results for design models belonging to class DM3 to investigate three issues:

(1) the fact that, in the ECG-motivated design models of this thesis, the complete six stage estimator of Figure 5.5 is not necessary;

(2) the robustness of the estimator, when it models phenomena that do not occur in the data; and

(3) the presentation of a culminating estimation example in which realizations of the stochastic process defined by the design model more closely resemble an ECG.

The design model used for the presentation of the first two issues is shown in Figure 6.27. Note the two possible values for the parameter $p^1_{14,15}$. Example realizations for this design model, for both values of $p^1_{14,15}$, are shown in Figures 6.28 and 6.29. The design model for the third issue is presented in Section 6.21.

In light of the results of Section 6.15, none of the estimators presented in the following sections have any subtractor chains (i.e. submodels of type P3 or P4).

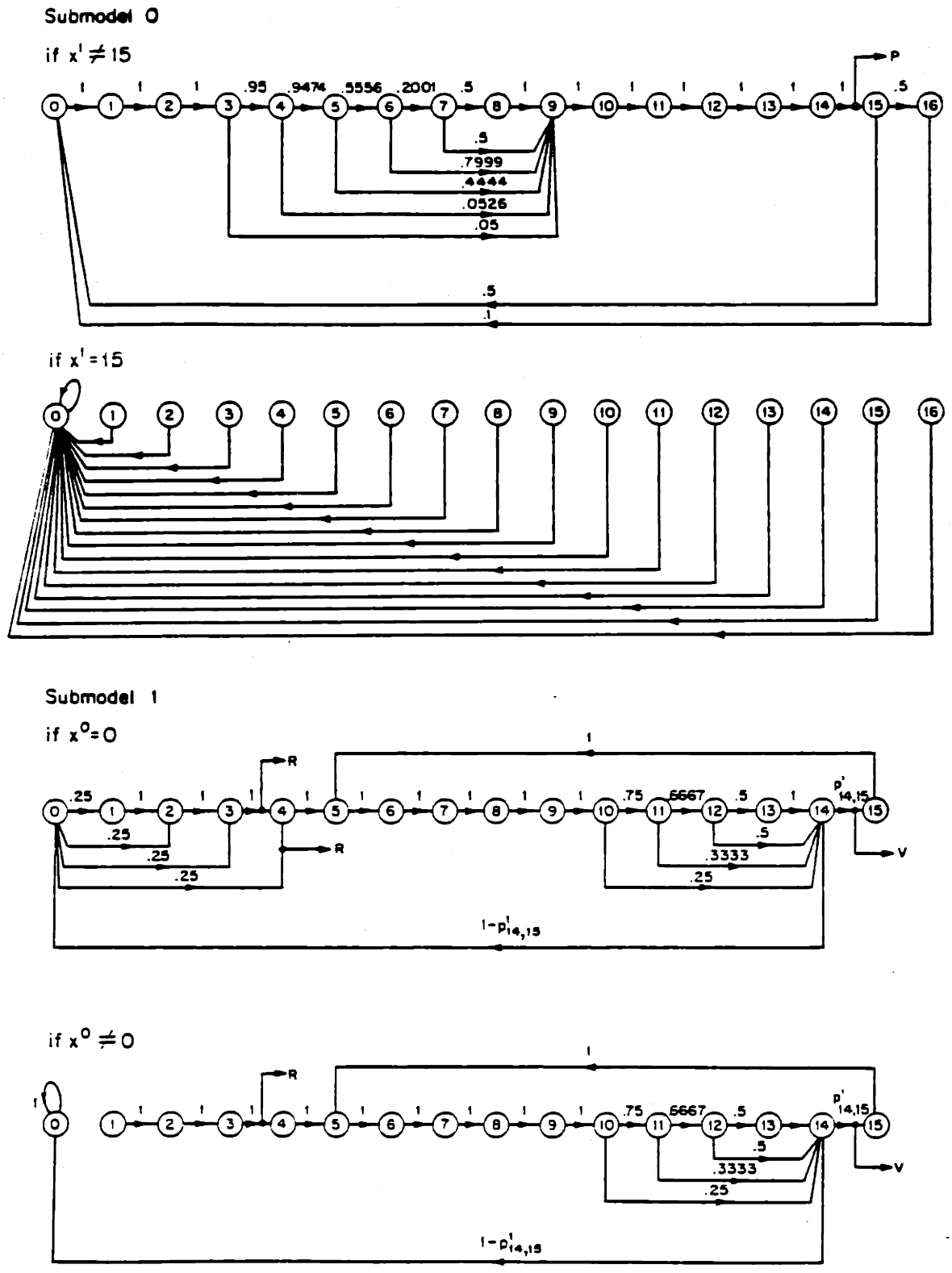


Figure 6.27 A Normal Rhythm Design Model With Reentrant-Mechanism PVCs.

Markov chain cycle interval = signature sampling interval = 1 (normalized time).

P: 2 samples long, each sample having mean .9 and standard deviation $\sqrt{.1}$.

R: 3 samples long, with means .75, 4.5, and .75 and common standard deviation $\sqrt{.5}$.

V: 6 samples long, with means .75, 4.5, .75, -.75, -4.5, and -.75 and common standard deviation 1.

Observation noise: mean 0 and standard deviation 1.

$p^1_{14,15} = 0$ or .05.

Figure 6.27 Continued.

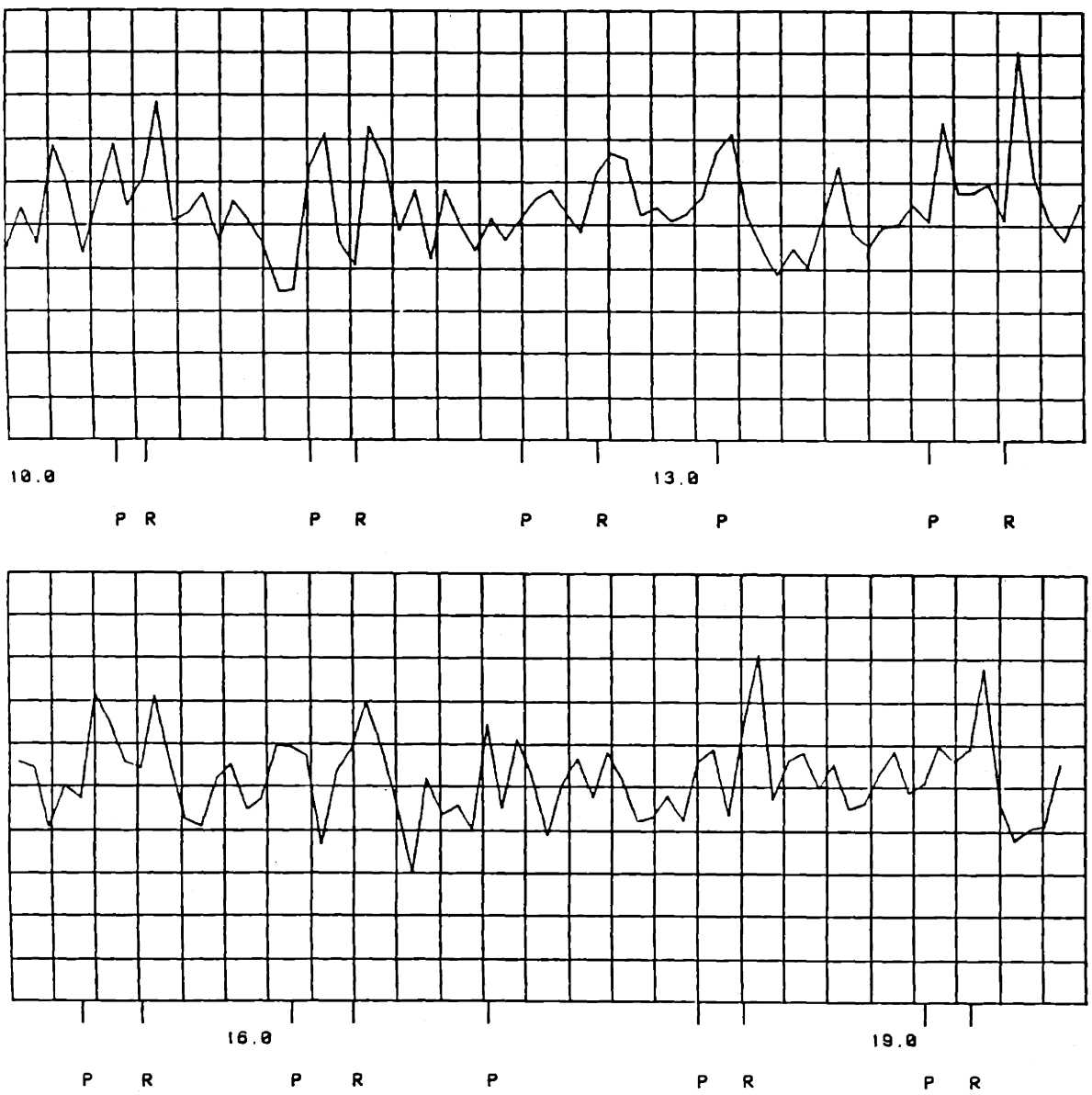


Figure 6.28 A Realization of the Stochastic Process Defined in Figure 6.27 in the Case of $p^{1}_{14,15}=0$.

Section 6.18.

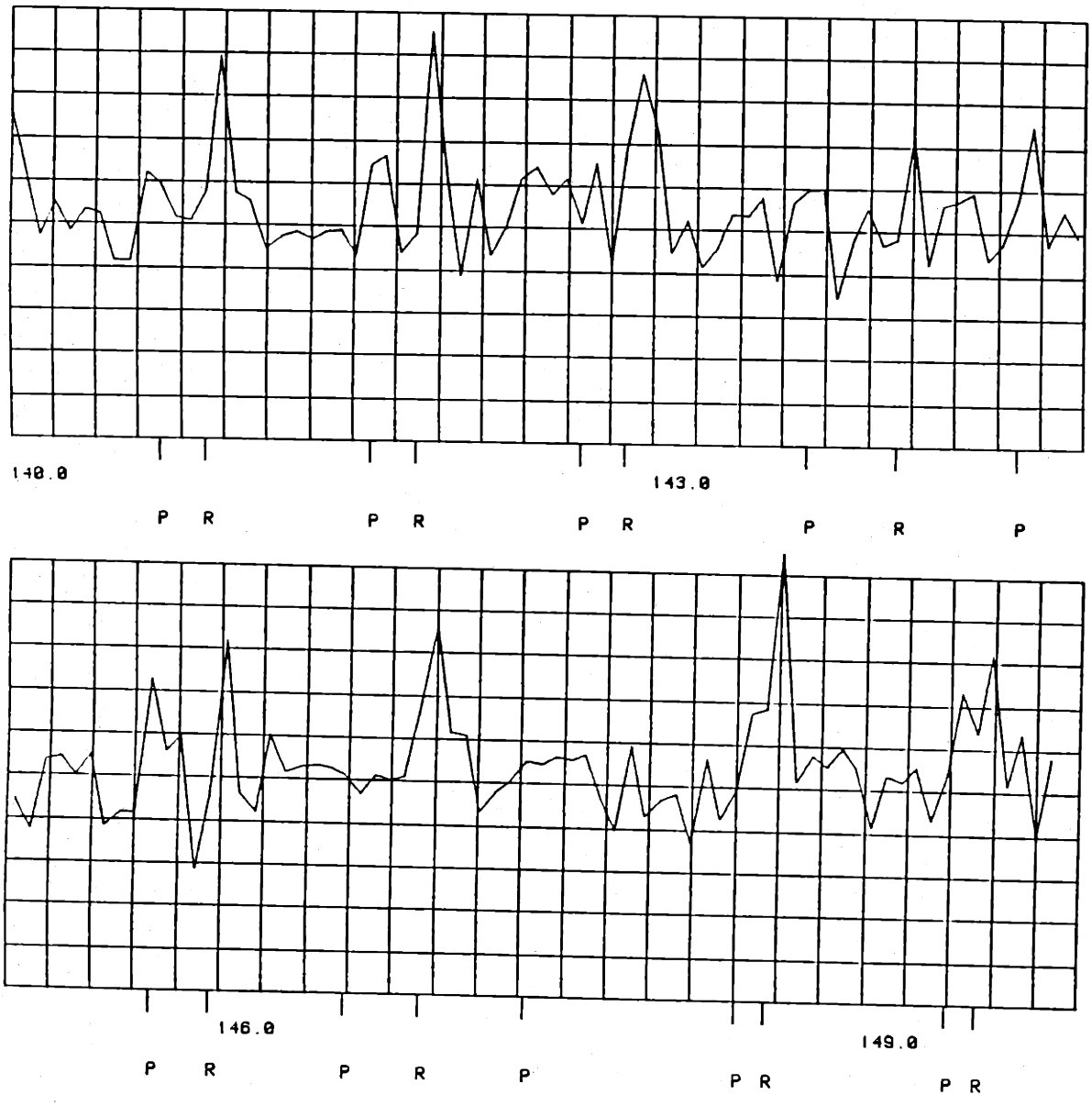


Figure 6.29 A Realization of the Stochastic Process Defined in Figure 6.27 in the Case of $p^1_{14,15}=.05$.
Note the dropped R waves near 145 and 147.8 seconds.

Section 6.18.

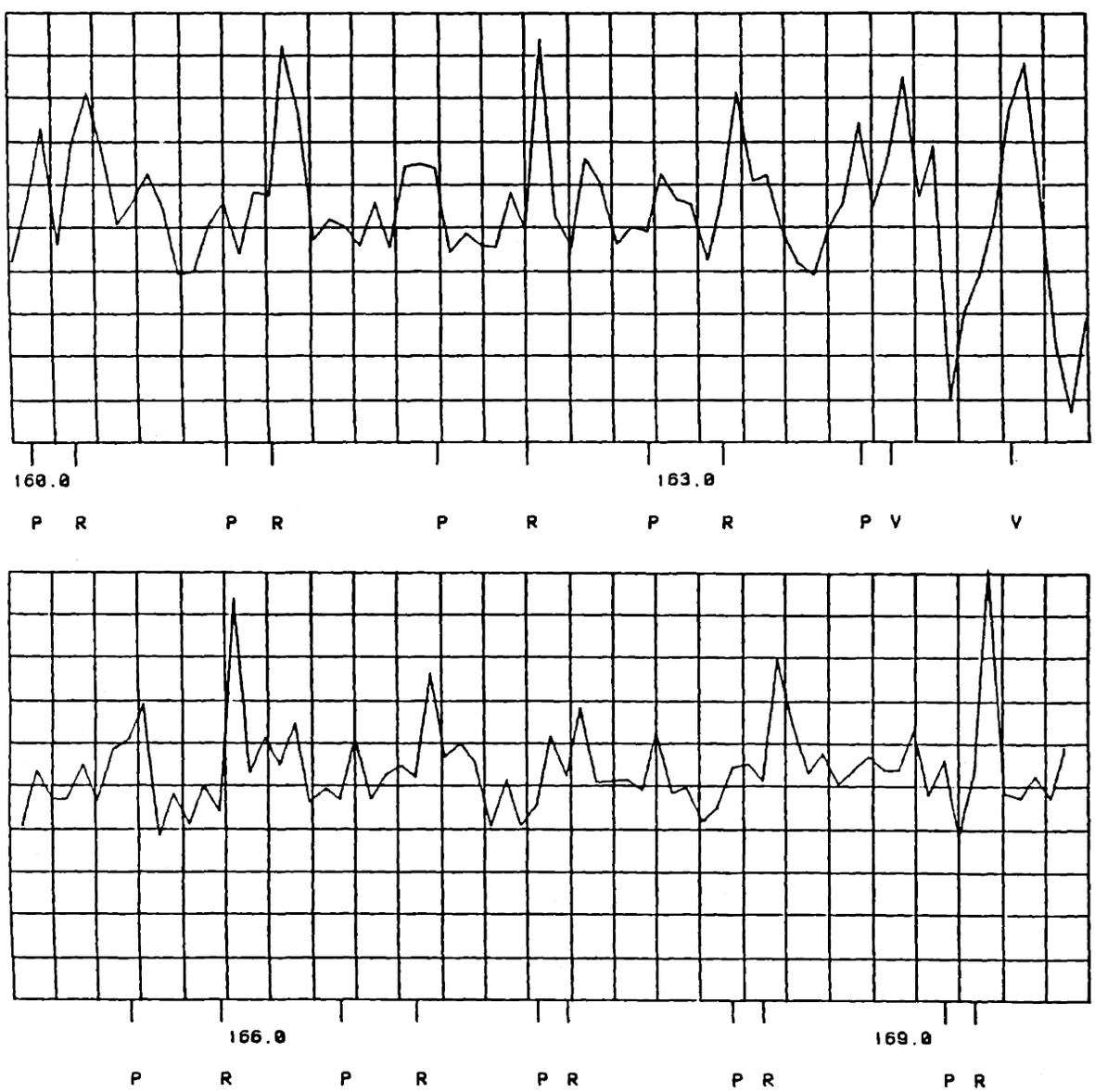


Figure 6.29 Continued.
Note the pair of V waves.

6.19. Simpler Suboptimal Estimators for DM3-Class Design Models

For the ECG-motivated design models with which we are concerned, it is clear that the six-part estimator of Figure 6.30 performs calculations that do not have a significant effect on the final estimate. (Figure 6.30 is Figure 5.5 with all subtractor chains deleted and with the obvious generalization to multiple signatures initiated by submodel 1). Specifically, LE 0 pass 0 is of little value because the SNR of the waves initiated by submodel 0 (i.e. P waves) is so much lower than the SNR of the waves initiated by submodel 1 (i.e. R and V waves). Thus the estimation of P waves makes sense only as a subsequent calculation (i.e. pass 1) to the initial detection of R and V waves (pass 0).

Consider the design model of Figure 6.27, with $p^1_{14,15}=.05$. Table 6.30 shows the misclassification table for LE 0 pass 0 of the estimator shown in Figure 6.30. The table is based on matches in which phase shifts are allowed. Note that the R and V waves are unmodeled in the estimator and therefore there are no R or V wave columns in the table. As expected, performance is very poor.

Unlike the situation in Section 6.15, here there are no pseudo-observations to generate windows around the unmodeled waveforms (i.e. R and V waves). Therefore, the MAP solution is to ignore the true P waves, accepting them as unusual noise bursts, and attempt to explain the true R and V waves by calling them P waves. Note that interpreting an R or V wave as a P wave (versus an unusual noise burst) is more favorable than it initially appears because the variance of the P wave added to the variance of the observation noise greatly increases the a posteriori probability of the data. Use of P4 subtractor chains (two

chains, one for R and one for V) in LE 0 pass 0 would probably improve the results. However, removing LE 0 pass 0 and taking advantage of the temporal window created in LE 0 pass 1 by the pseudo-observations originating from LE 1 pass 0 is a much less expensive solution computationally.

Table 6.31 shows the statistics for $^{0,0}\hat{h}_n^{C0,P0}$ based on the same simulation. The fact that performance is very poor is reflected, for instance, in the 1385 occurrences of $^{0,0}\hat{h}_n^{C0,P0}=0$, $^{0,0}\hat{x}_n^{C0}=0$, $(h_{n-1}^{0,1},h_n^{0,1},h_{n+1}^{0,1})=(1,1,1)$. It is not clear that later passes can extract any useful information from these augmented interactions. In order to consider this question, we investigated a second estimator.

A block diagram of the second estimator is shown in Figure 6.31. The only difference between the first and second estimators is that in the second estimator there is no LE 0 pass 0 and the ICS has been adjusted (in the usual manner) to accommodate the deletion of LE 0 pass 0. Statistics for the global wave-tracking estimate are shown in Tables 6.32 and 6.33 for the first and second estimators respectively. These two tables show misclassification performance based on matches in which phase shifts are allowed. The difference in performance between these two estimators is small and actually favors the estimator lacking LE 0 pass 0.

Unlike prior estimators, the estimator of Figure 6.31 reestimates the submodel 0 state trajectory. That is, LE 0 pass 2 exists. It is questionable whether this pass can improve on the estimate from LE 0 pass 1. To investigate this issue, consider the estimator of

Figure 6.32. This estimator lacks both LE 0 pass 2 and LE 1 pass 1 and the ICS has been adjusted (in the usual manner) to accommodate the deletions. The result is an estimator in very close analogy with the DM1 and DM2 estimators.

Using the same stochastic process realization considered previously in this section, the statistics for the global wave-tracking estimate of this third estimator are shown in Table 6.34. As usual, phase shifts are allowed. Comparison of Table 6.33 and 6.34, which exhibit results for the second and third estimators respectively, shows essentially identical performance. Therefore, for the ECG-motivated design models with which we are concerned, the reestimation of the submodel 0 state trajectory provided by LE 0 pass 2 does not appear able to improve on the initial estimate provided by LE 0 pass 1. Thus, in the following sections, only estimators with the structure shown in Figure 6.32 are considered.

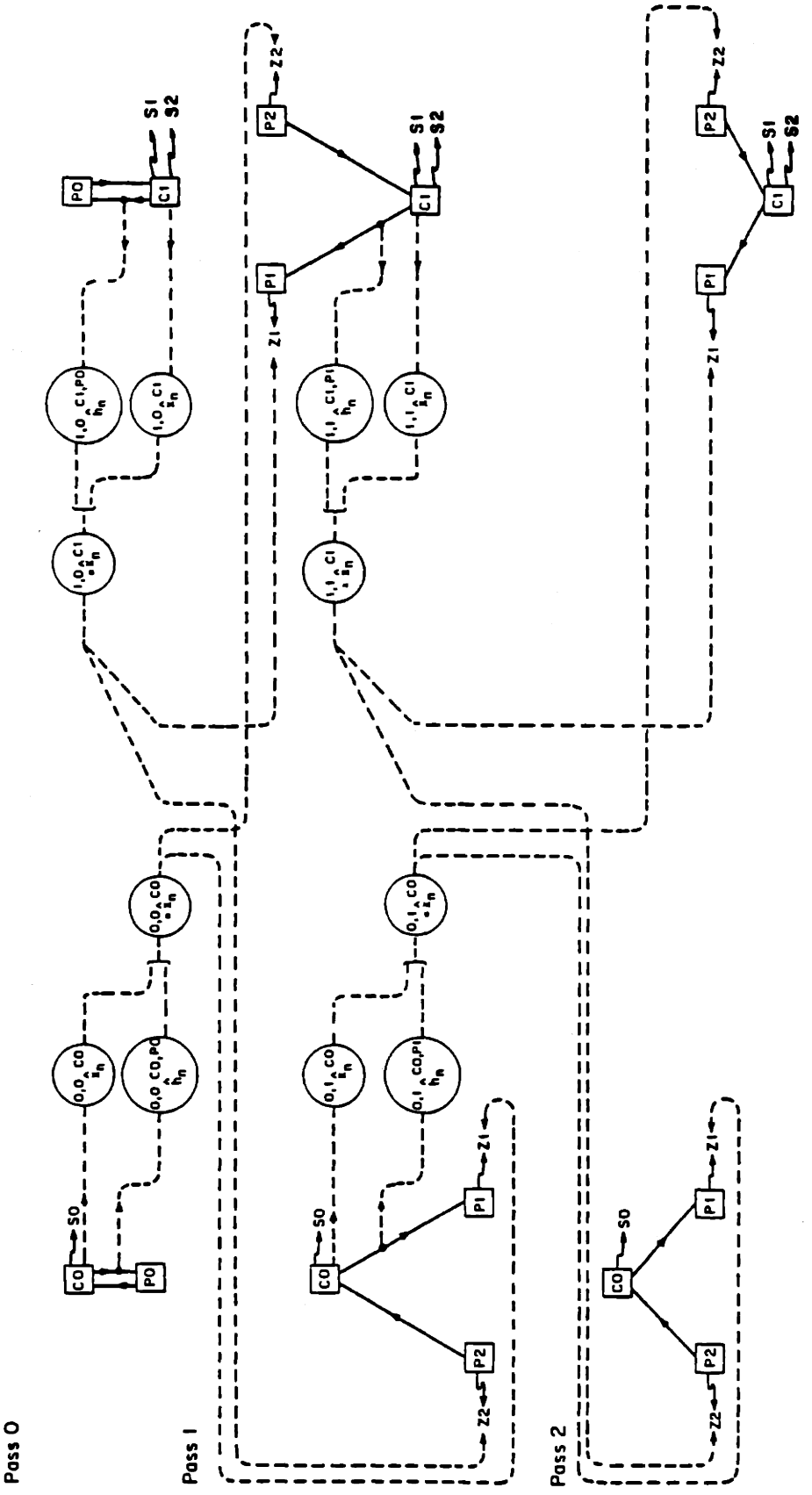


Figure 6.30 A Modified Version of the DW3 Estimator of Figure 5.5.
The changes are that the subtractor chains are deleted and there are multiple signatures from chain C1.

Section 6.19.

		Estimated		
		P	fn	rs
Truth	P	40	1372	1412
	R	1182	72	1254
	V	36	15	51
	fp	175		175
	cs	1433	1459	

fpP = 1393/1433 = .972(.00448)
fnP = 1372/1412 = .972(.00454)
teP = (1393+1372)/1412 = 1.96

Table 6.30 Misclassification Performance of LE 0 Pass 0 in the Estimator of Figure 6.30.

			truth interaction triplet $(h_{n-1}^{0,1}, h_n^{0,1}, h_{n+1}^{0,1})$						
			(1,0,0)	(0,1,0)	(1,1,0)	(0,0,1)	(1,0,1)	(0,1,1)	(1,1,1)
estimated interaction $0,0\hat{h}_n^{C0,P0}$	0	estimated state $0,0\hat{x}_n^{C0}$ 0	0	0	8	0	14	26	1385
		1	0	0	0	0	0	0	0
		2	0	0	0	0	0	0	0
		.	.	.	.	.	.	.	.
		.	.	.	.	.	.	.	.
		.	.	.	.	.	.	.	.
		16	0	0	0	0	0	0	0
	1	estimated state $0,0\hat{x}_n^{C0}$ 0	0	0	0	0	0	0	0
		1	0	0	1	0	8	15	1410
		2	0	0	0	0	1	8	1424
		3	0	0	2	0	0	1	1430
		4	0	0	2	0	1	0	1288
		5	0	0	13	0	2	1	1275
		6	0	0	18	0	10	1	801
		7	1	0	2	0	0	2	122
		8	0	0	1	1	1	0	98
		9	1	0	129	0	22	11	1270
		10	1	0	390	1	128	22	891
		11	0	0	359	1	390	129	554
		12	0	0	323	0	359	391	360
		13	1	0	137	0	322	359	614
		14	2	2	22	1	135	320	951
		15	0	1	29	2	24	135	1242
		16	0	0	3	0	19	19	515

Table 6.31 Statistics on Augmented Interaction Estimates From LE 0 Pass 0.

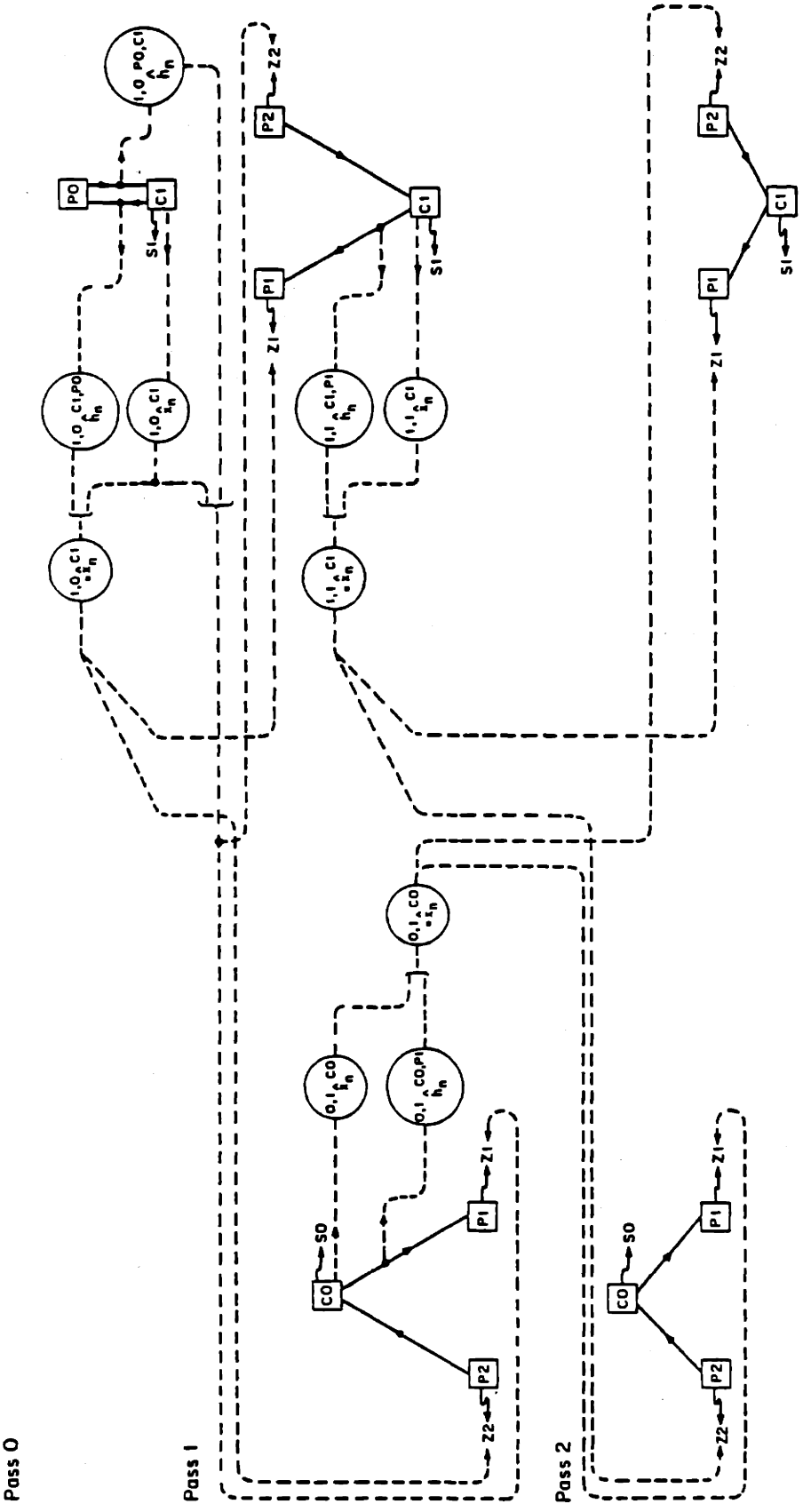


Figure 6.31 A Modified Version of the DM3 Estimator of Figure 6.30.
The change is that LE 0 pass 0 is deleted.

Section 6.19.

		Estimated				
		P	R	V	fn	rs
Truth	P	1184	14	1	207	1406
	R	5	1200	1	42	1248
	V	0	0	51	0	51
	fp	237	18	5		260
	cs	1426	1232	58	249	

fpP = 242/1426 = .170(.0586)
fnP = 232/1406 = .165(.0600)
teP = (242+232)/1406 = .337
fpR = 32/1232 = .0260(.174)
fnR = 48/1248 = .0385(.142)
teR = (32+48)/1248 = .0641
fpV = 7/58 = .121(.354)
fnV = 0/51 = 0
teV = (7+0)/51 = .137

Table 6.32 Misclassification Performance of the Estimator of Figure 6.30.

		Estimated				
		P	R	V	fn	rs
Truth	P	1188	15	1	204	1408
	R	3	1216	1	30	1250
	V	0	0	51	0	51
	fp	236	22	3		261
	cs	1427	1253	56	234	

fpP = 239/1427 = .167(.0590)
fnP = 220/1408 = .156(.0619)
teP = (239+220)/1408 = .326
fpR = 37/1253 = .0295(.162)
fnR = 34/1250 = .0272(.169)
teR = (37+34)/1250 = .0568
fpV = 5/56 = .0893(.427)
fnV = 0/51 = 0
teV = (5+0)/51 = .0980

Table 6.33 Misclassification Performance of the Estimator of Figure 6.31.

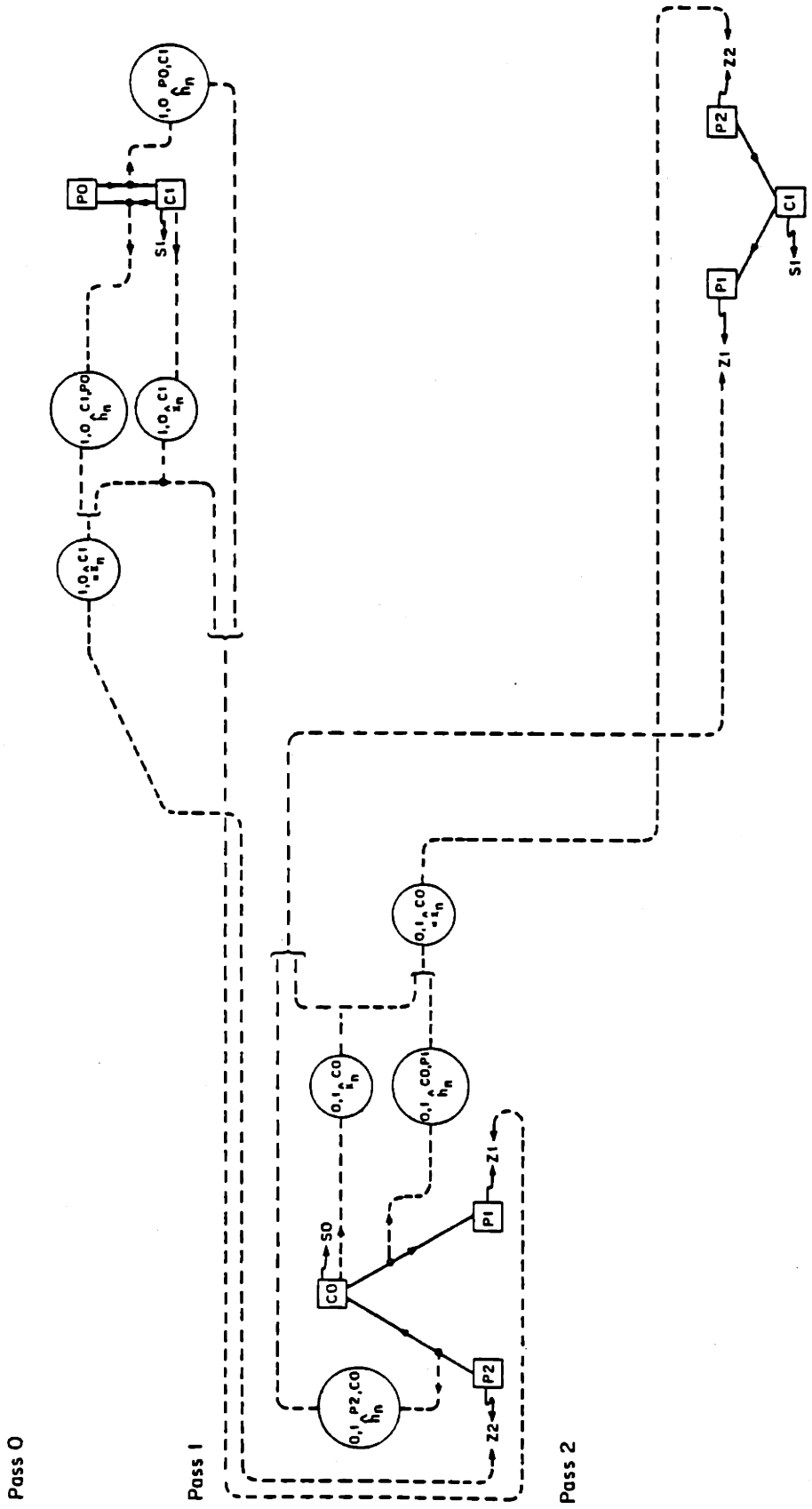


Figure 6.32 A Modified Version of the DM3 Estimator of Figure 6.31.
The changes are that LE 1 pass 1 and LE 0 pass 2 are deleted.

Section 6.19.

		Estimated				
		P	R	V	fn	rs
Truth	P	1188	15	0	207	1410
	R	3	1216	2	30	1251
	V	0	0	51	0	51
	fp	238	22	3		263
	cs	1429	1253	56	237	

fpP = 241/1429 = .169(.0587)
fnP = 222/1410 = .157(.0616)
teP = (241+222)/1410 = .328
fpR = 37/1253 = .0295(.162)
fnR = 35/1251 = .0280(.167)
teR = (37+35)/1251 = .0576
fpV = 5/55 = .0909(.426)
fnV = 0/51 = 0
teV = (5+0)/51 = .0980

Table 6.34 Misclassification Performance of the Estimator of Figure 6.32.

6.20. Robustness Performance of Estimators Modeling Behavior Not Seen in the Observations

Previously we discussed a robustness example in which the design model defining the simulated data was less constrained than the design model on which the estimator was based. The reverse situation arises when a normal ECG is processed by an estimator which models some abnormality. The situation is closely related to false positive performance. That is, a high performance estimator must be sensitive in order to detect the abnormality in the first place. The question investigated here is whether the sensitivity leads to false positive detections of the abnormality. The investigation is simplified by using data in which the abnormality never occurs, since in that case all detections are false positive detections.

Consider the stochastic process defined by the design model of Figure 6.27 with $p^1_{14,15}=0$. In this design model, there are no V waves and

submodel 0 is never reset. The estimator, however, is based on the design model of Figure 6.27 with $p^1_{14,15}$=.05. Therefore, the estimator expects V waves and the associated resetting of submodel 0. The estimator has the structure shown in the block diagram of Figure 6.32. This structure is the third of the three structures considered in the prior section.

Misclassification statistics for the global estimate are shown in Table 6.35. Performance is essentially indistinguishable from the matched-estimator situation of the prior section. In order to understand the cause of the two false positive V wave detections, in Figure 6.33 we plot the observations, the truth annotations, and the global estimate annotations. The false positive V wave detections are at approximately 563.1 and 846.6 seconds. Figure 6.33 also shows the same observations but with all the Gaussian noise removed. That is, $\sigma_P = \sigma_R = \sigma_{obs} = 0$. (The value of σ_V does not matter since $p^1_{14,15}$=0). Figure 6.34 shows the means of the P, R and V waves. Note when considering these plots that the temporal sampling rate is extremely coarse. Specifically, there are .4 gridlines/sample. It is clear that the two false positive V wave detections are simply due to extraordinary noise bursts which the estimator found to be less costly to account for if they were taken to be V waves.

		Estimated				
		P	R	V	fn	rs
Truth	P	1201	17	1	212	1431
	R	2	1252	1	32	1287
	fp	224	20	0		244
	cs	1427	1289	2	244	

fpP = 226/1427 = .158(.0610)
fnP = 230/1431 = .161(.0604)
teP = (226+230)/1431 = .319
fpR = 37/1289 = .0287(.162)
fnR = 35/1287 = .0272(.167)
teR = (37+35)/1287 = .0559

Table 6.35 Misclassification Statistics.

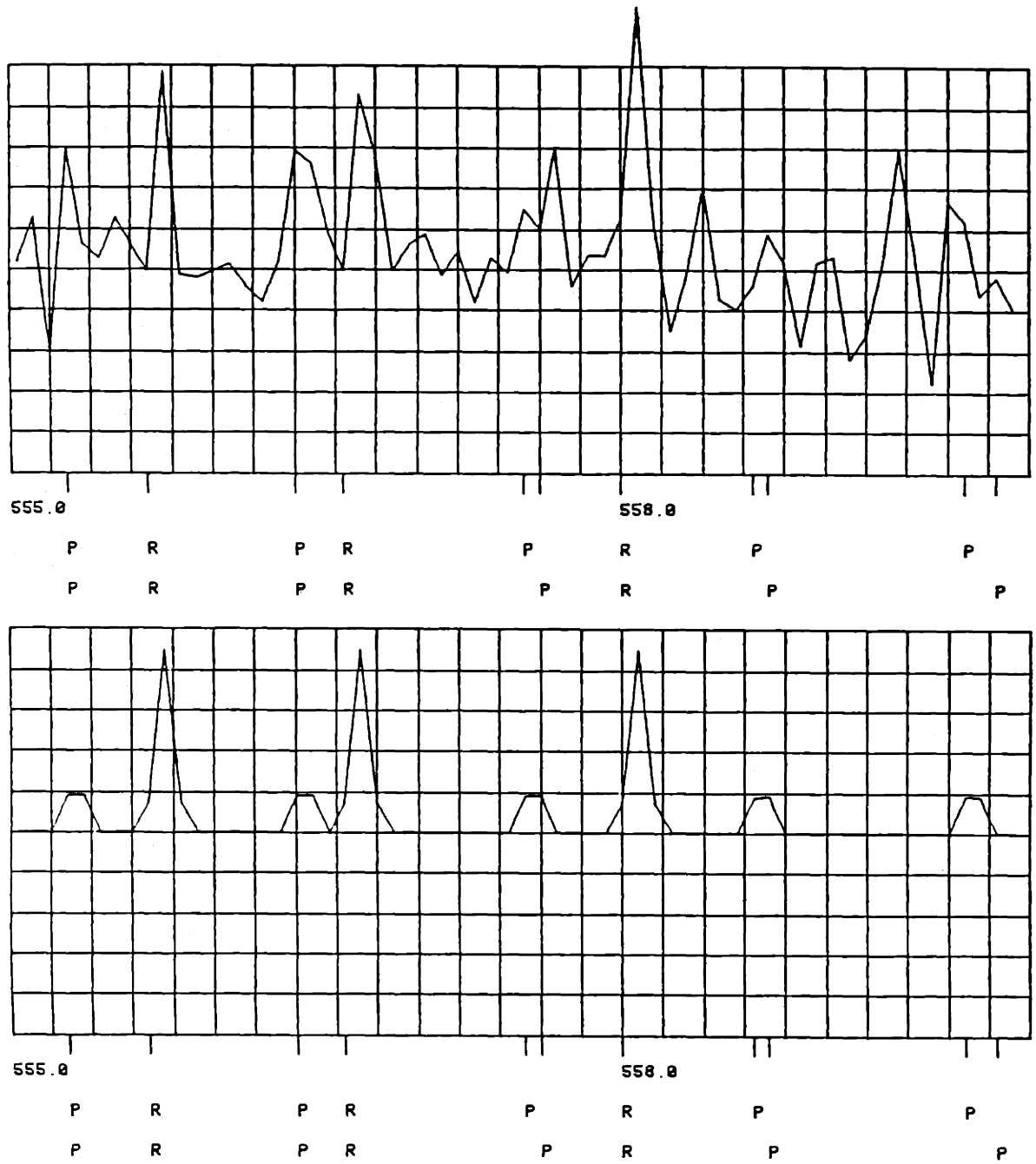


Figure 6.33 ECG Observations with Estimation Results.
The upper trace shows the actual signal while the lower trace shows the signal with the Gaussian noise removed. Beneath each trace there are two sets of annotations, of which the upper is true and the lower is estimated.

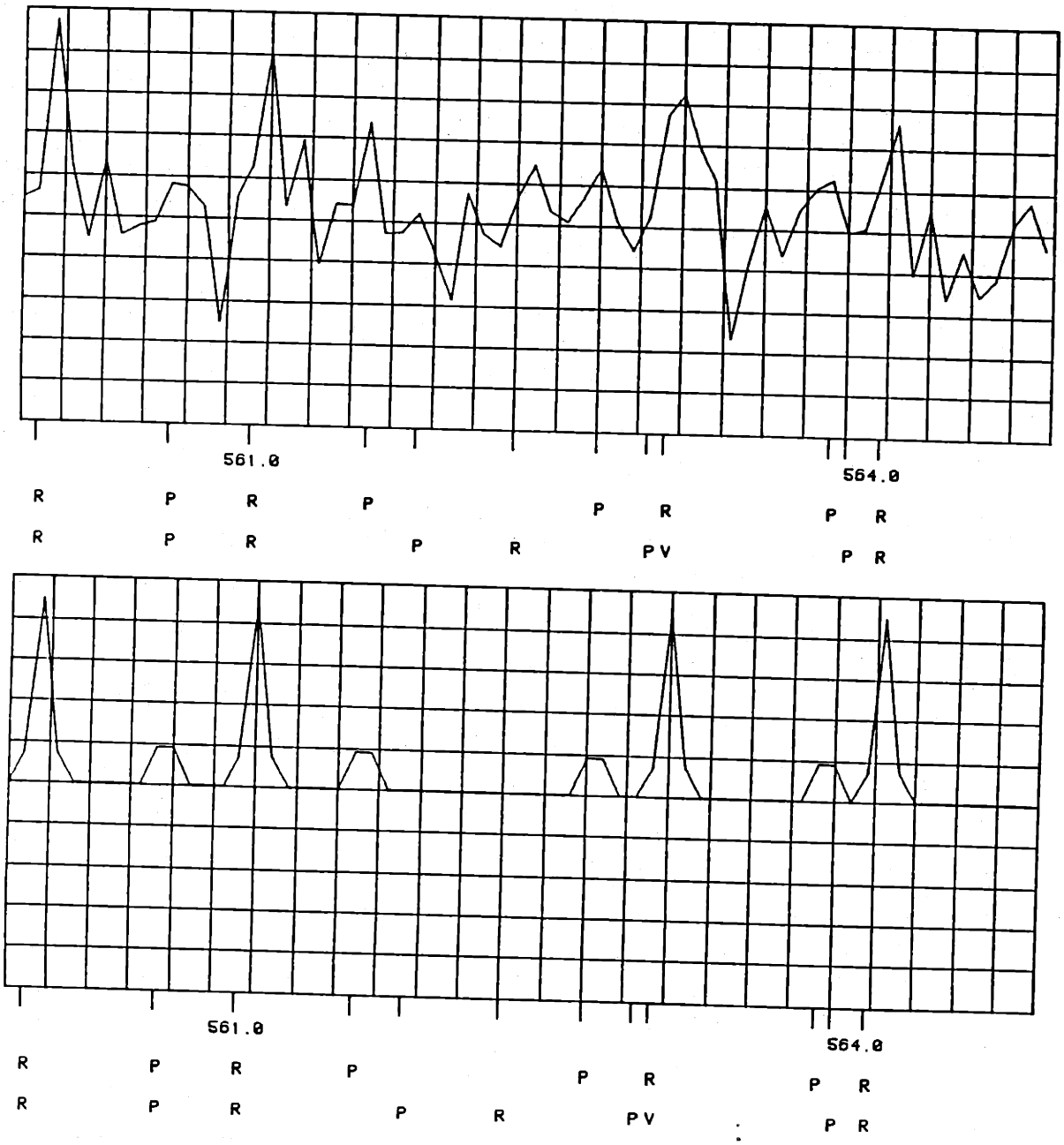


Figure 6.33 Continued.

Section 6.20.

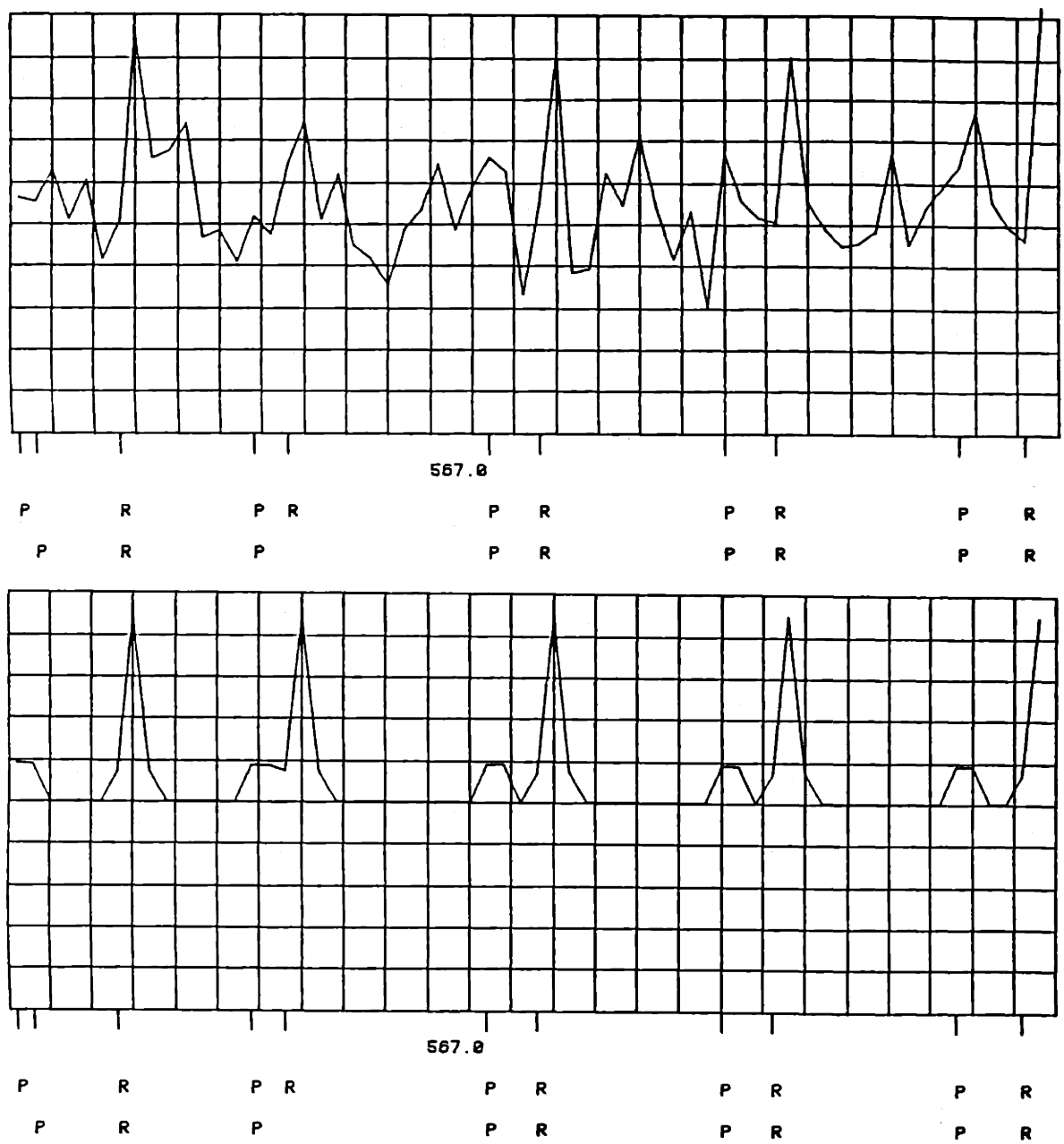


Figure 6.33 Continued.

Section 6.20.

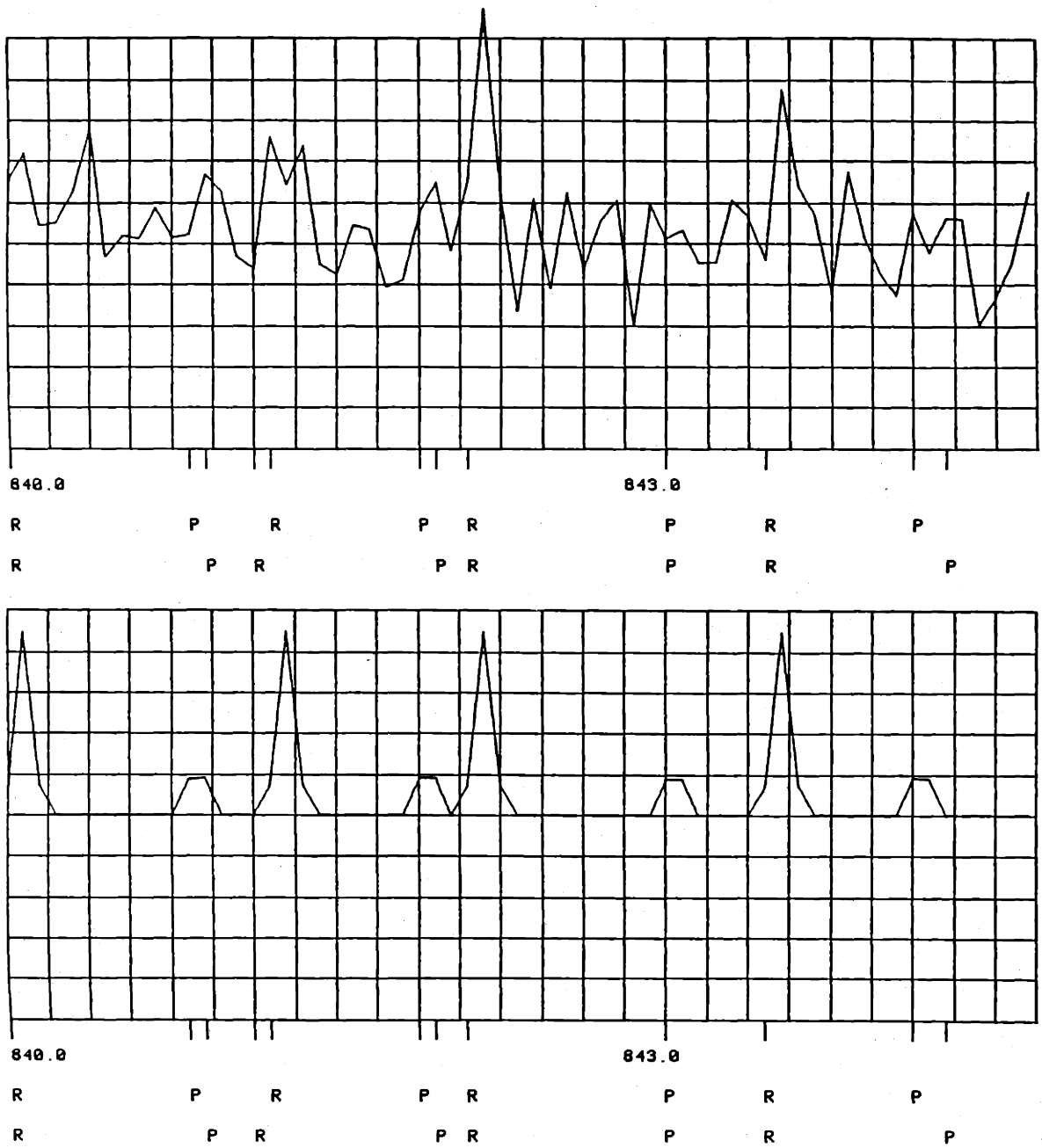


Figure 6.33 Continued.

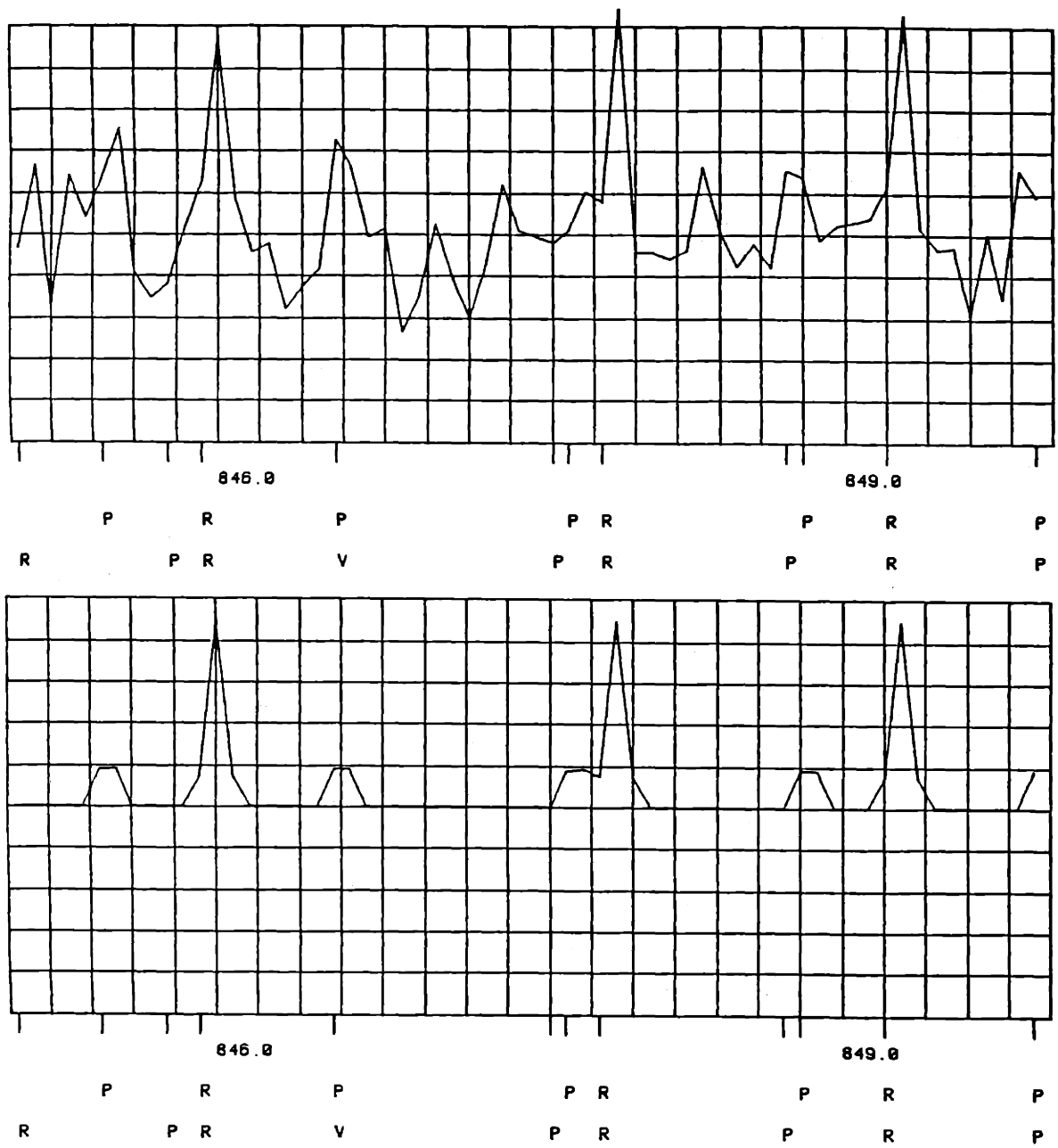


Figure 6.33 Continued.

Section 6.20.

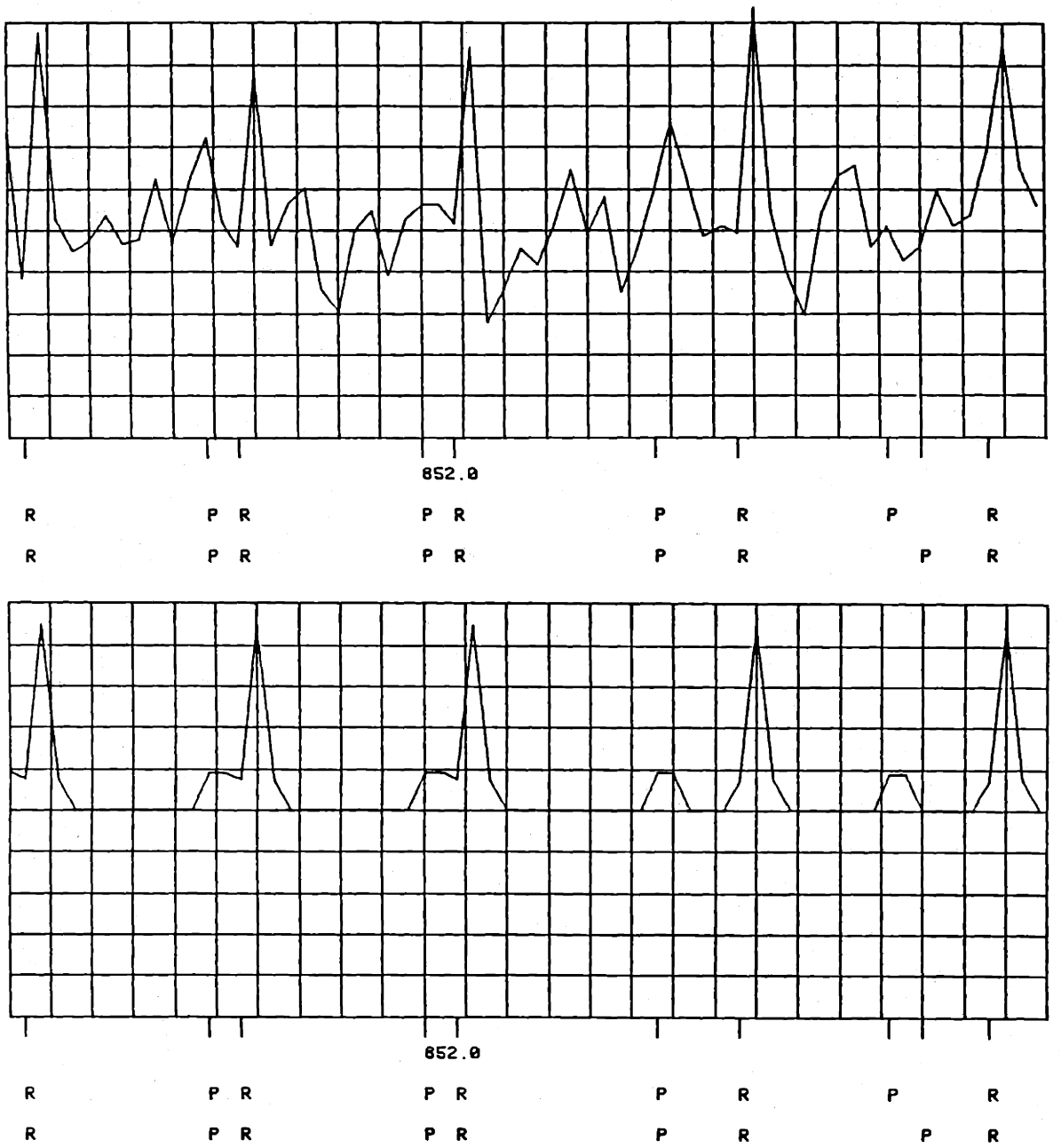


Figure 6.33 Continued.

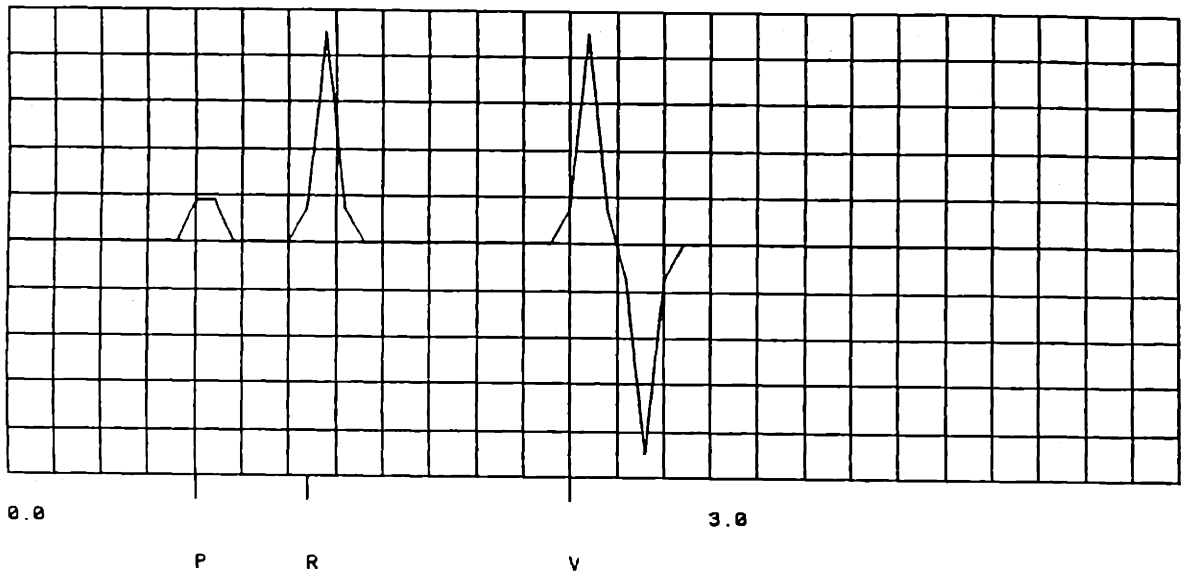


Figure 6.34 The Means of the P, R, and V Waves.

6.21. A Realistic Estimation Example

In this section we present a culminating estimation example. In comparison with previous examples, the realizations of the stochastic process defined by this design model more closely resemble an ECG, both in morphology and SNR. (Recall that the motivation for our previous use of unreasonably low SNRs was the need to compute statistically significant results in a simulation of practical duration).

The design model is shown in Figure 6.35. Similar to the design model of Figure 6.27 with $p^1_{14,15}=.05$, the present design model is also a

design model for normal rhythm with reentrant-mechanism PVCs. In comparison with the design model of Figure 6.27, there are three major differences:

(1) Given that a PVC occurred, the probability that the next ventricular depolarization is a PVC is increased. This explains states 15-21 in submodel 1, and especially the transition probabilities out of state 21 in submodel 1.

(2) If a PVC occurs, submodel 1 resets submodel 0 with probability 1/2 rather than with probability 1.

(3) The signature sampling interval is one fourth the Markov chain cycling interval. In order to make it possible for signatures to start at any observation sample, each signature appears four times with 0, 1, 2, or 3 leading zeros in the mean and covariance sequences. (The subscripts on the annotation names indicate the number of leading zeros). This explains states 22-24 in submodel 0 and states 22-30 in submodel 1.

The Markov chain cycling interval is .04 second, leading to a mean heart rate of 75 beats per minute, and the signature sampling interval is .01 second.

Figure 6.36 shows a plot of the simulated ECG and the truth and estimated annotation types and times. Estimated annotations are shown for each pass and for the global reconstruction. This simulated data resembles an ECG, but it is not entirely realistic. The chief difference is the character of the noise. First, no bandlimited noise, "shot noise" like events, or baseline instabilities are present. Second, the

noise that is present, though it resembles electromyogram artifact when considered locally, appears to have a time-varying intensity (higher during the P, R, T, and V waves) which is not realistic for true electromyogram artifact. (Recall that the reason for the time-varying intensity was to give a crude model of morphology variability).

Using an estimator with the structure shown in Figure 6.32, excellent performance was obtained.[4] Specifically, all R, T, and V waves were detected with the exception of the initial R and T (which were missed only due to edge effects generated by the multiple pass structure of the estimator). The initial P wave was also dropped (for the same reason). Counting this initial P wave problem as an error, the P wave performance was fpP = fnP = 25/231 = .108 (.189). Of these errors, there were two isolated false negatives (the initial P wave and one other P wave), two isolated false positives, and all the remaining false positive and false negatives were paired. That is, the remaining false positive and false negatives occurred as pairs but with a separation great enough so that the truth and estimated waves did not overlap. In all of these paired errors, the estimated time of the wave was closer to the R wave (i.e. smaller P-R interval) than the true time. This may be due to the absence of subtractor chains. That is, the energy at the beginning of the R wave, which is not accounted for since there is no subtractor, may bias the P wave estimate towards the R wave.

In the estimator described here, the zeros in the estimated observation pmfs for the pseudo-observations were not rounded up. The reason

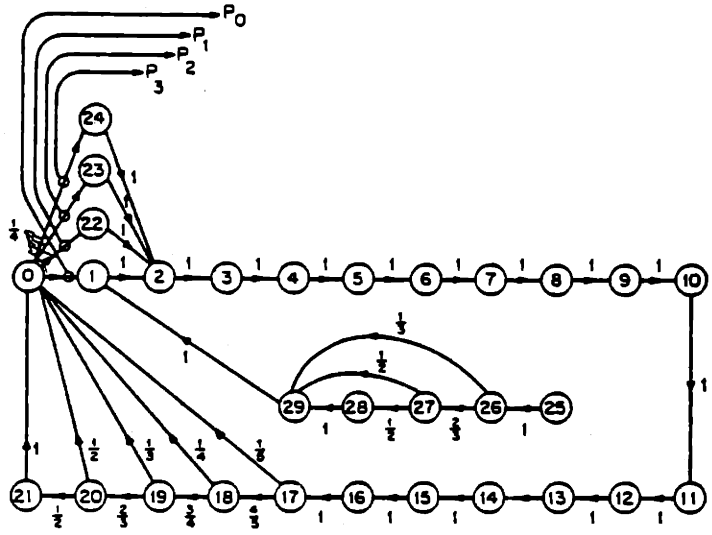
---

[4] In contrast, the global MAP estimator failed due to exceeding the available computer storage.

for this change was that without it, in a similar model in which the equivalent of state 29 was connected to the equivalent of state 0 rather than state 1, the P wave estimates tracked the R and T waves. (It was possible to track both because it was possible to estimate that submodel 1 was reset each beat). This is the calculation alluded to at the end of Section 6.15.

SA - Atria: x^0

if $x^1 \neq 13$:



if $x^1 = 13$: then the tpm is $\frac{1}{2}(P^0+Q^0)$ where $Q^0 = \{q^0_{i,j}\}$ and $q^0_{i,j} = \begin{cases} 1 & j = 25 \\ 0 & \text{otherwise} \end{cases}$

AV - Ventricles: x^1

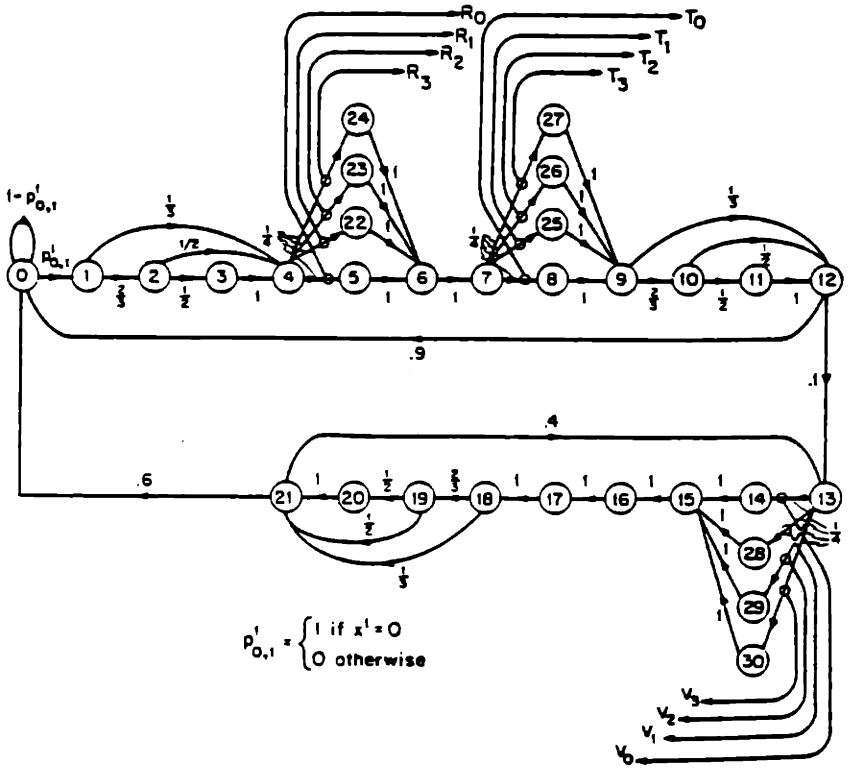


Figure 6.35 A Relatively Realistic Model for Normal Rhythm with Reentrant-Mechanism PVCs.

Markov chain cycling interval = .04 seconds; signature sampling interval = .01 seconds.

P_0: 6 samples long; mean is .069985, .090351, .098974, .098974, .090351, .069985; standard deviation is constant at .02.

R_0: 4 samples long; mean is .5, 1.0, .66, .33; standard deviation is constant at .2.

T_0: 14 samples long; mean is .05, .1, .2, .3, .35, .4, .4, .35, .3, .2, .1, .05; standard deviation is constant at .12;

V_0: 8 samples long; mean is .25, .75, 1.0, 1.0, .75, .25, -.5, -.25; standard deviation is constant at .4.

X_i $X \in \{P,R,T,V\}$, $i \in \{1,2,3\}$: same as X_0 except that there are i leading zeros in both the mean and the covariance sequences.

Observation noise: mean 0 and standard deviation .02.

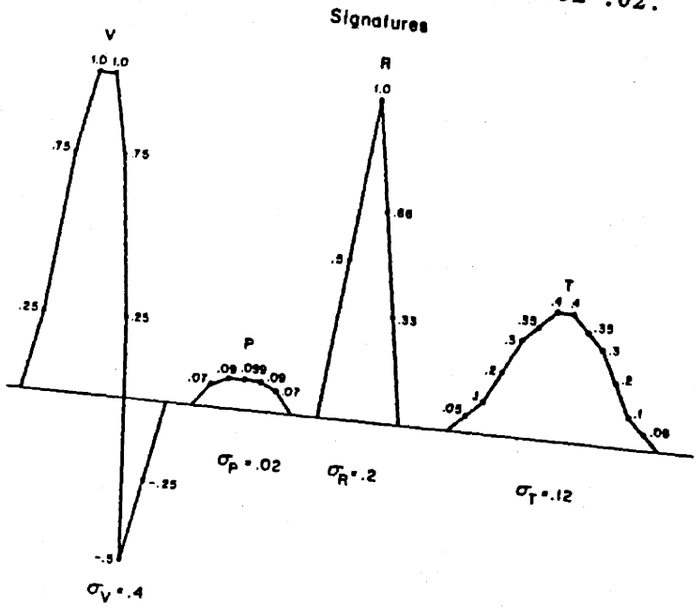


Figure 6.35 Continued.

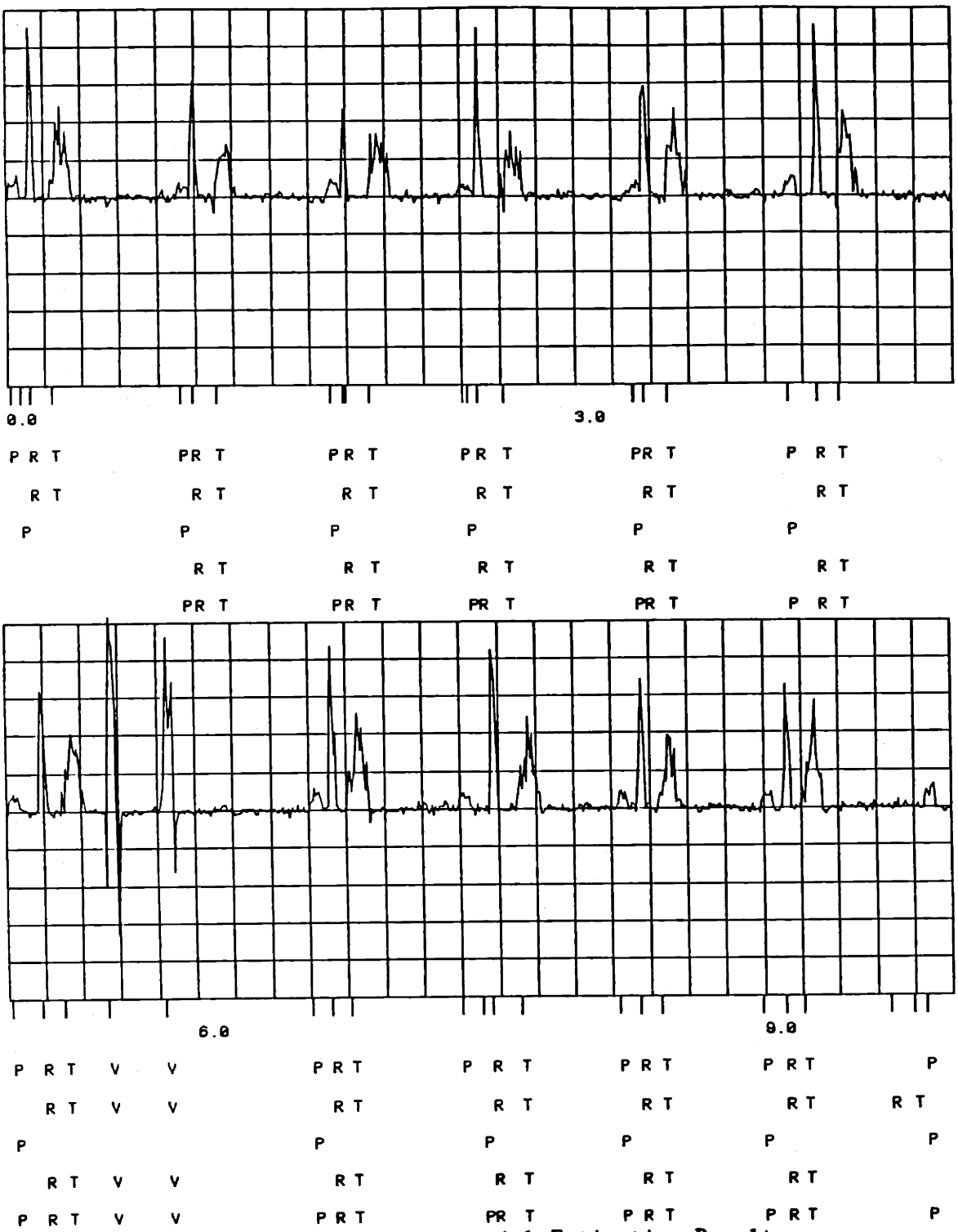


Figure 6.36 ECG Observations with Estimation Results.
Beneath each trace are five sets of annotations. From top to bottom these are: (1) true, (2) LE 1 pass 0, (3) LE 0 pass 1, (4) LE 1 pass 2, and (5) the result of the global reconstruction. The first four pages are continuous, following which are shown several pages of interesting, isolated segments.
Note the error (an extra R and T wave pair) in LE 1 pass 0 near 9.7 seconds which, however, did not adversely affect LE 0 pass 1 and was corrected by LE 1 pass 2. Note also the phase-shift error in LE 0 pass 1 near 7.4 seconds resulting in a P-R interval estimate that is too short.

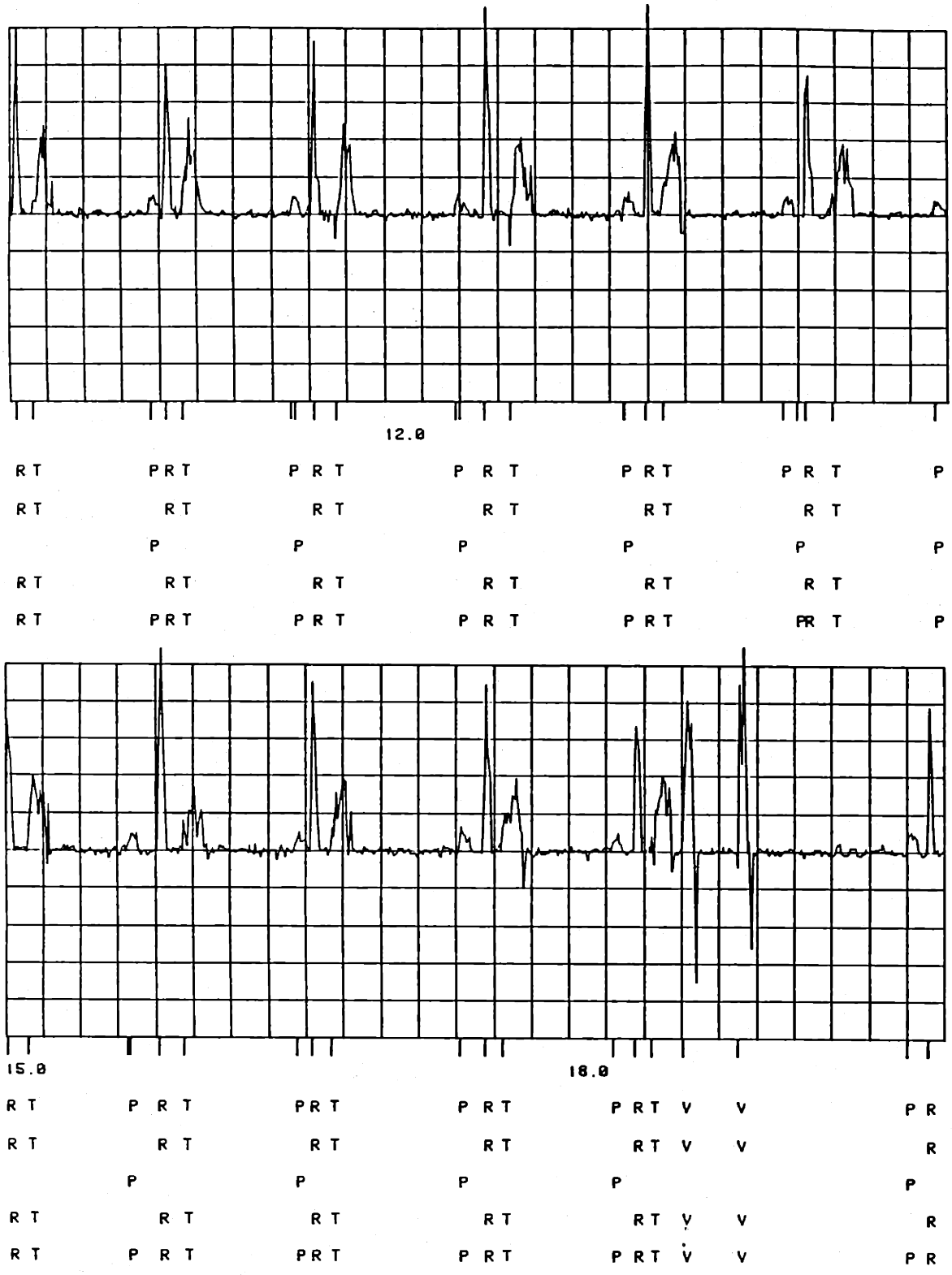


Figure 6.36 Continued.
Note the phase-shift error in LE 0 pass 1 near 14.1 seconds resulting in a P-R interval estimate that is too short.

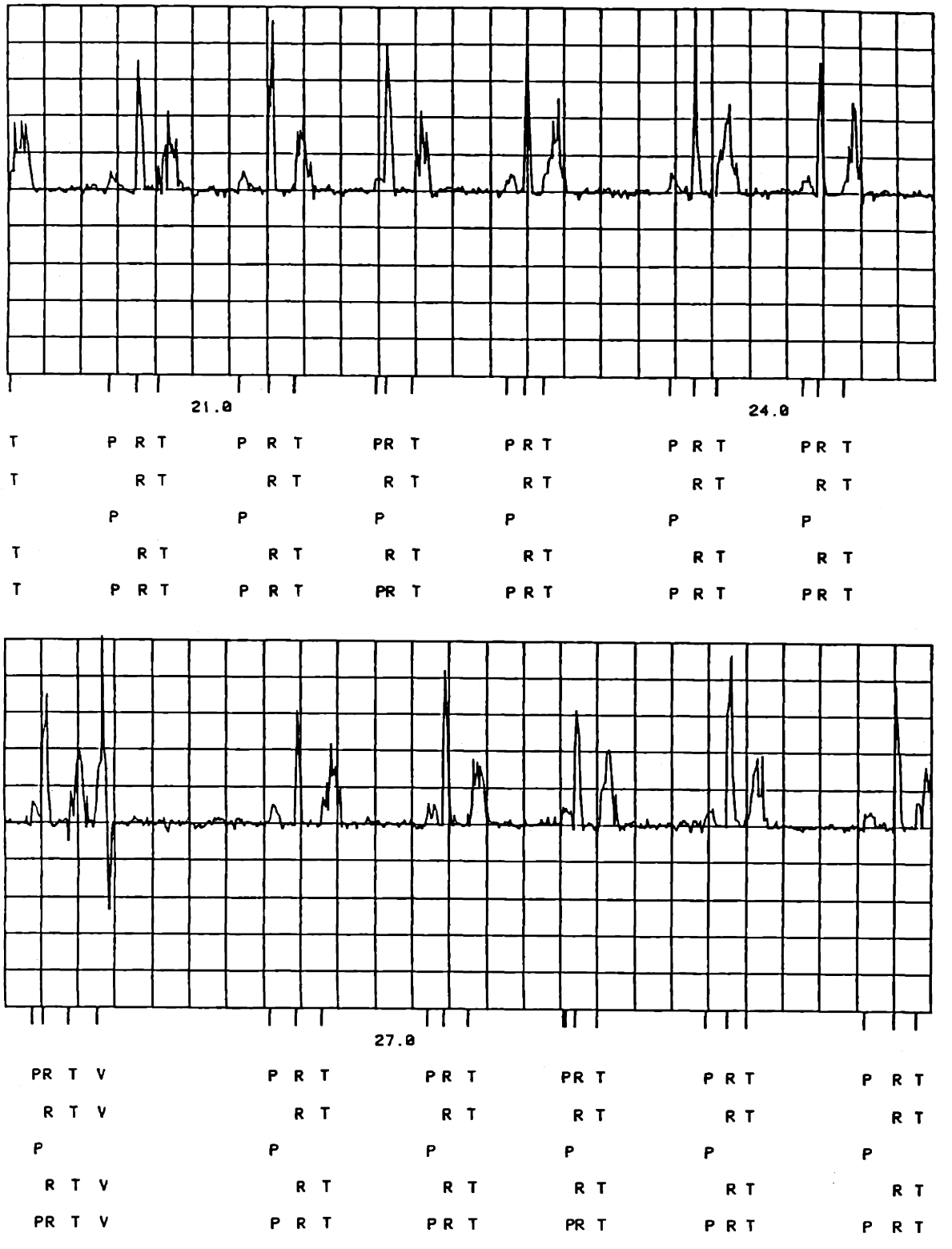


Figure 6.36 Continued.

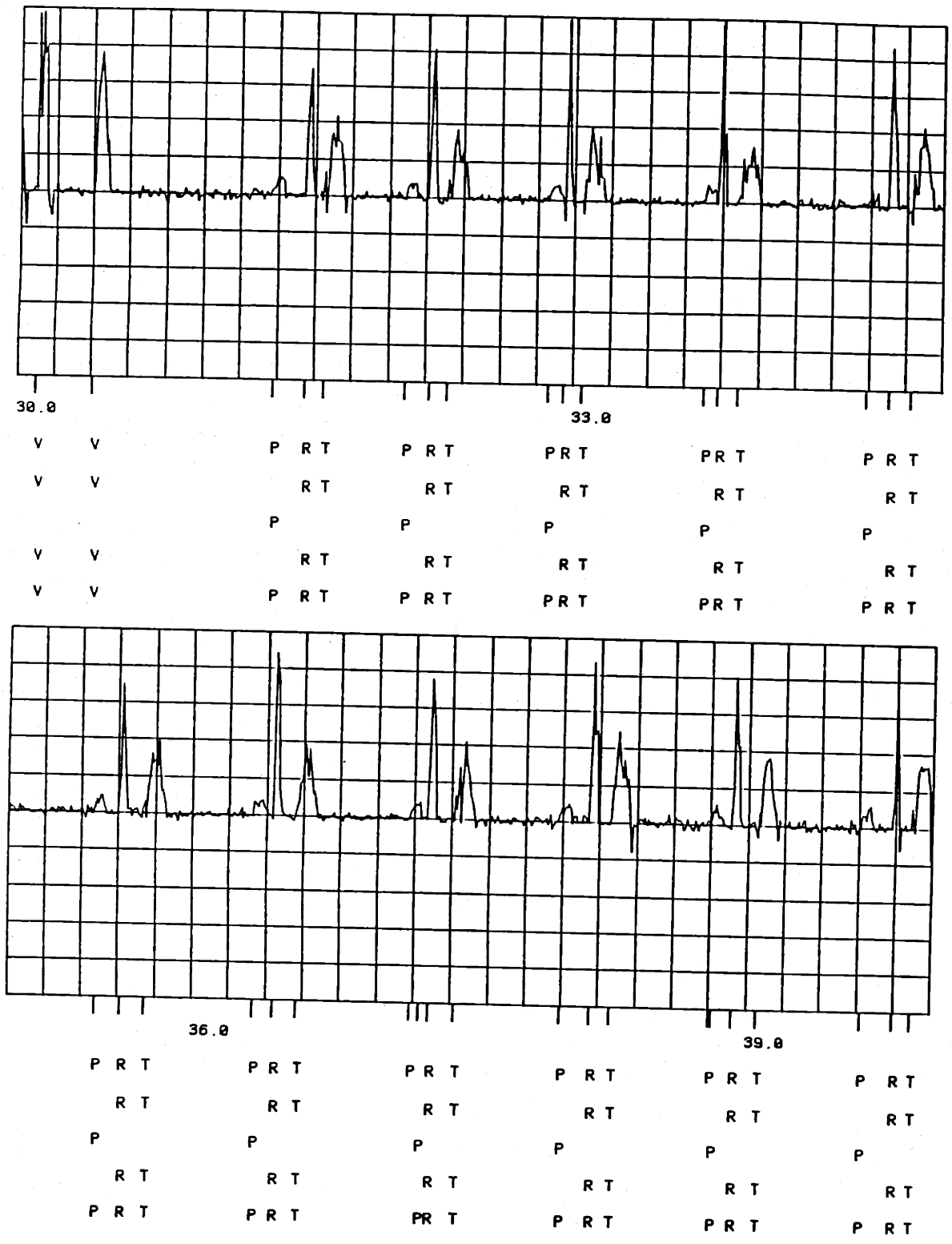


Figure 6.36 Continued.

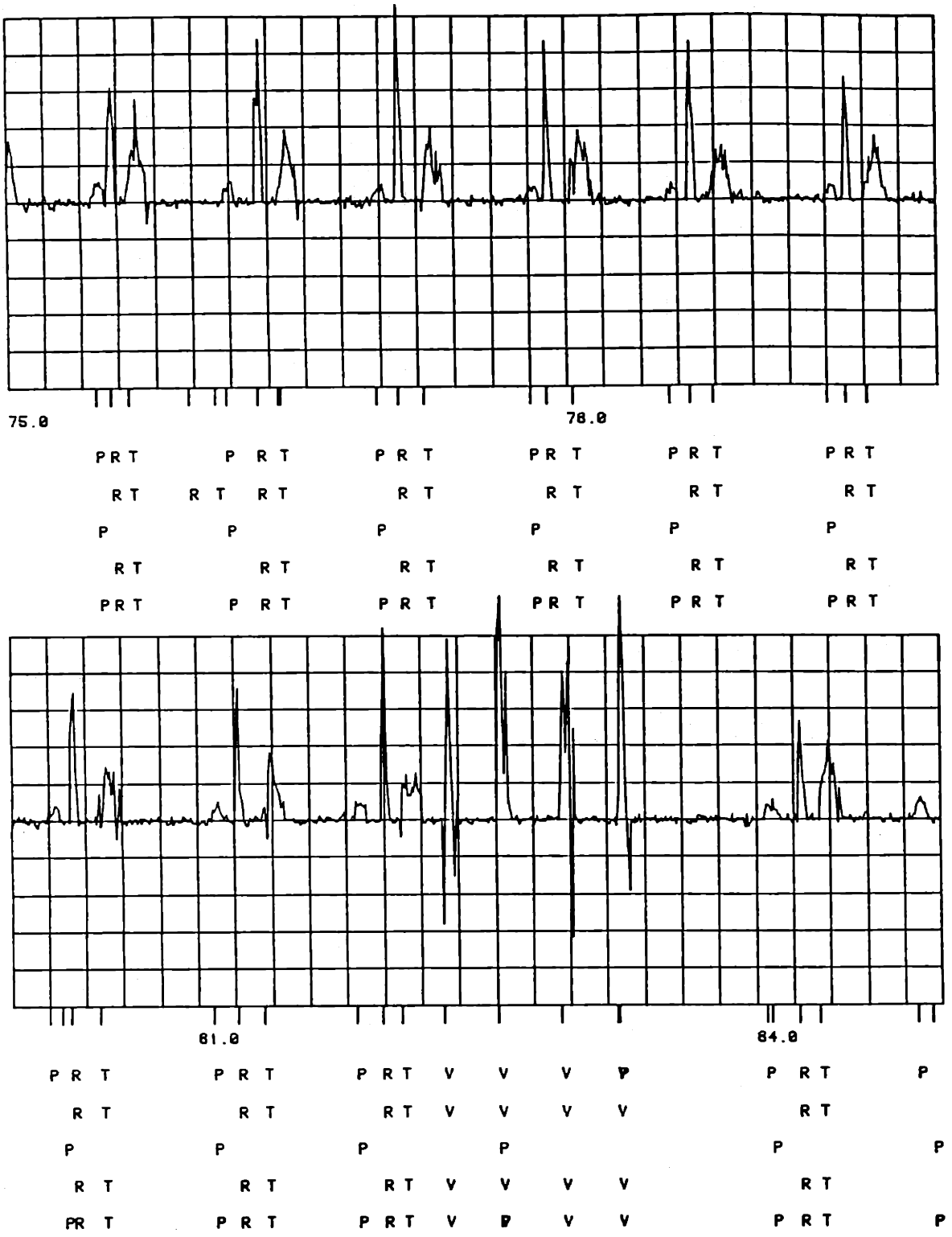


Figure 6.36 Continued.
Note the error (an extra R and T wave pair) in LE 1 pass 0 near 76 seconds (which is corrected by LE 1 pass 2), the phase-shift error in LE 0 pass 1 near 80.2 seconds (resulting in a P-R interval estimate that is too short), the false-positive error in LE 0 pass 1 near 82.6 seconds, and the false-negative error in LE 0 pass 1 near 83.3 seconds.

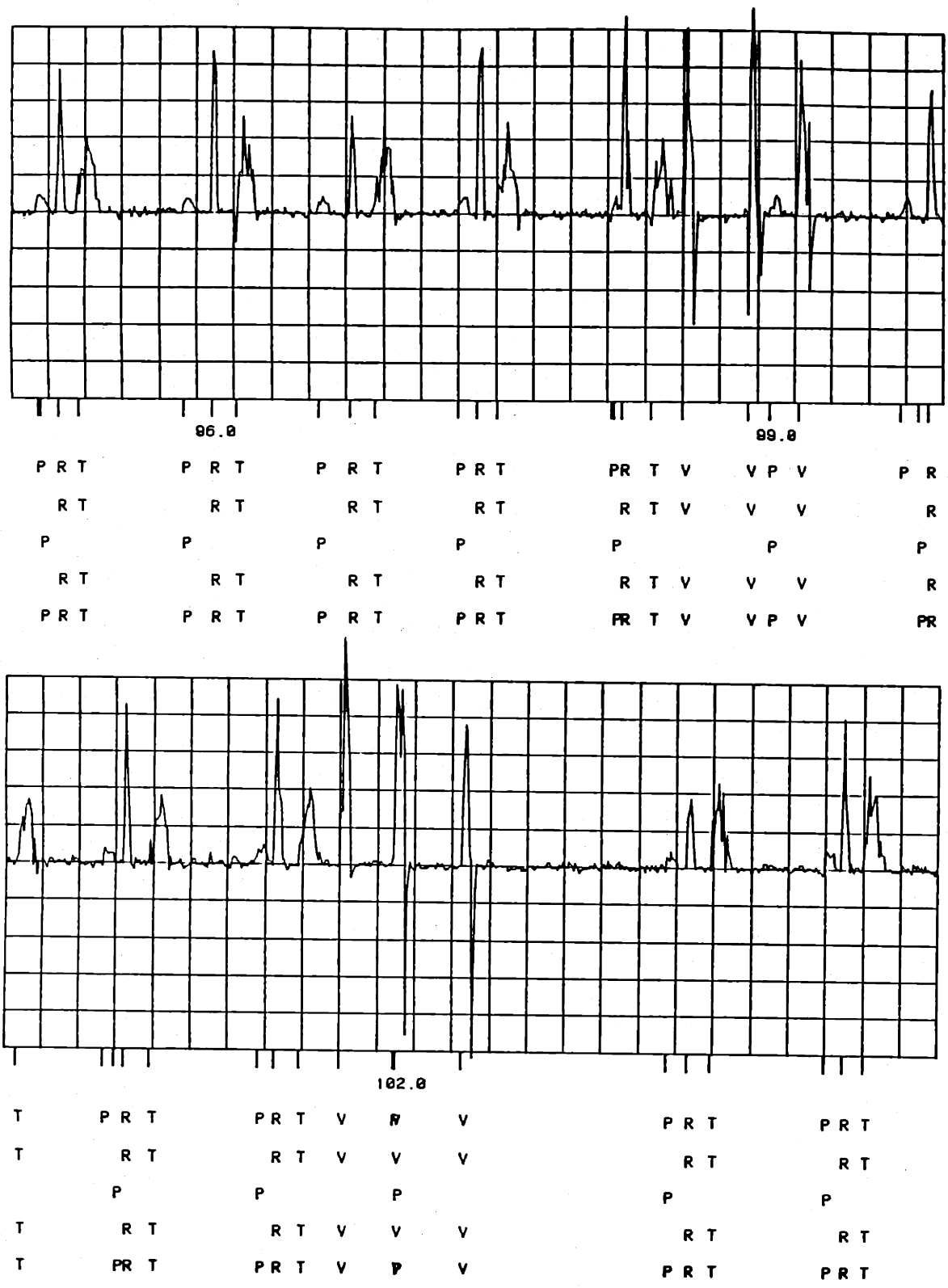


Figure 6.36 Continued.
Note the phase-shift errors in LE 0 pass 1 near 99.9 and 100.6 seconds (resulting in P-R interval estimates that are too short), but also the correct detection by LE 0 pass 1 of the P waves near 99 and 102 seconds, which both occur in difficult circumstances.

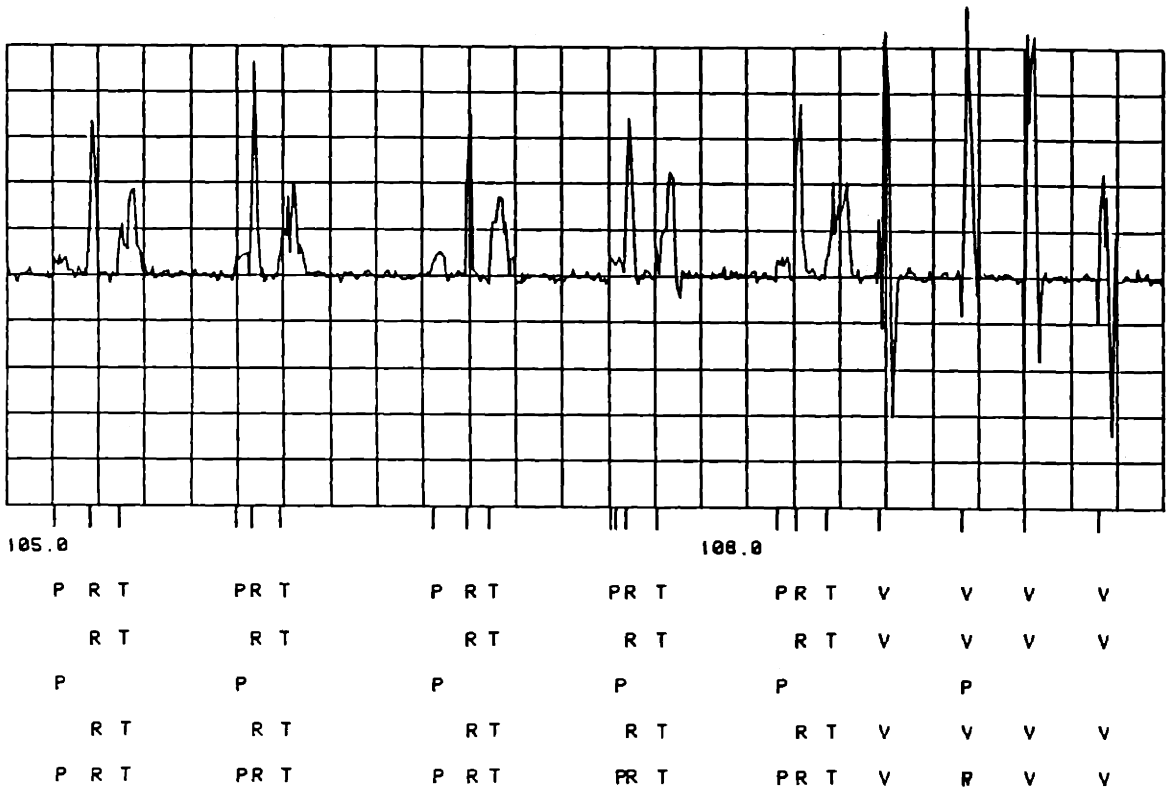


Figure 6.36 Continued.
Note the false-positive error in LE 0 pass 1 near 109.1 seconds.

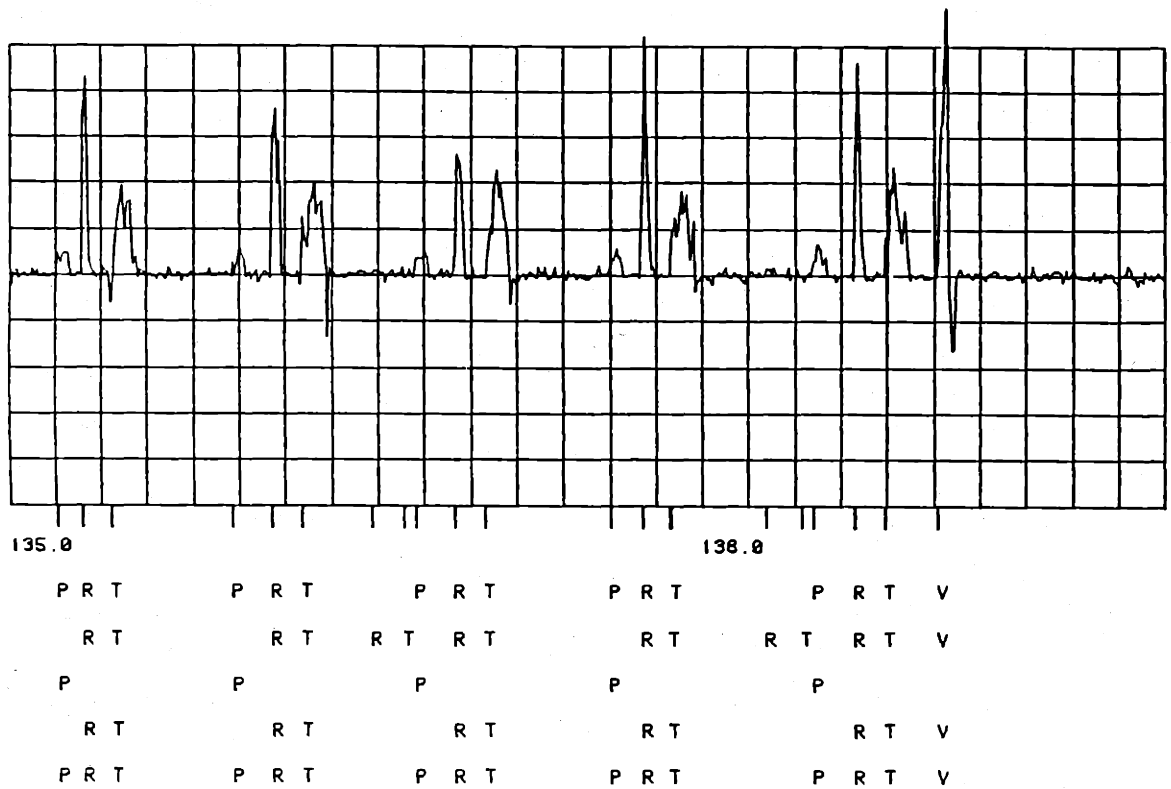


Figure 6.36 Continued.
Note the two errors (two extra R and T wave pairs) in LE 1 pass 0 near 136.6 and 138.3 seconds (both of which are corrected by LE 1 pass 2).

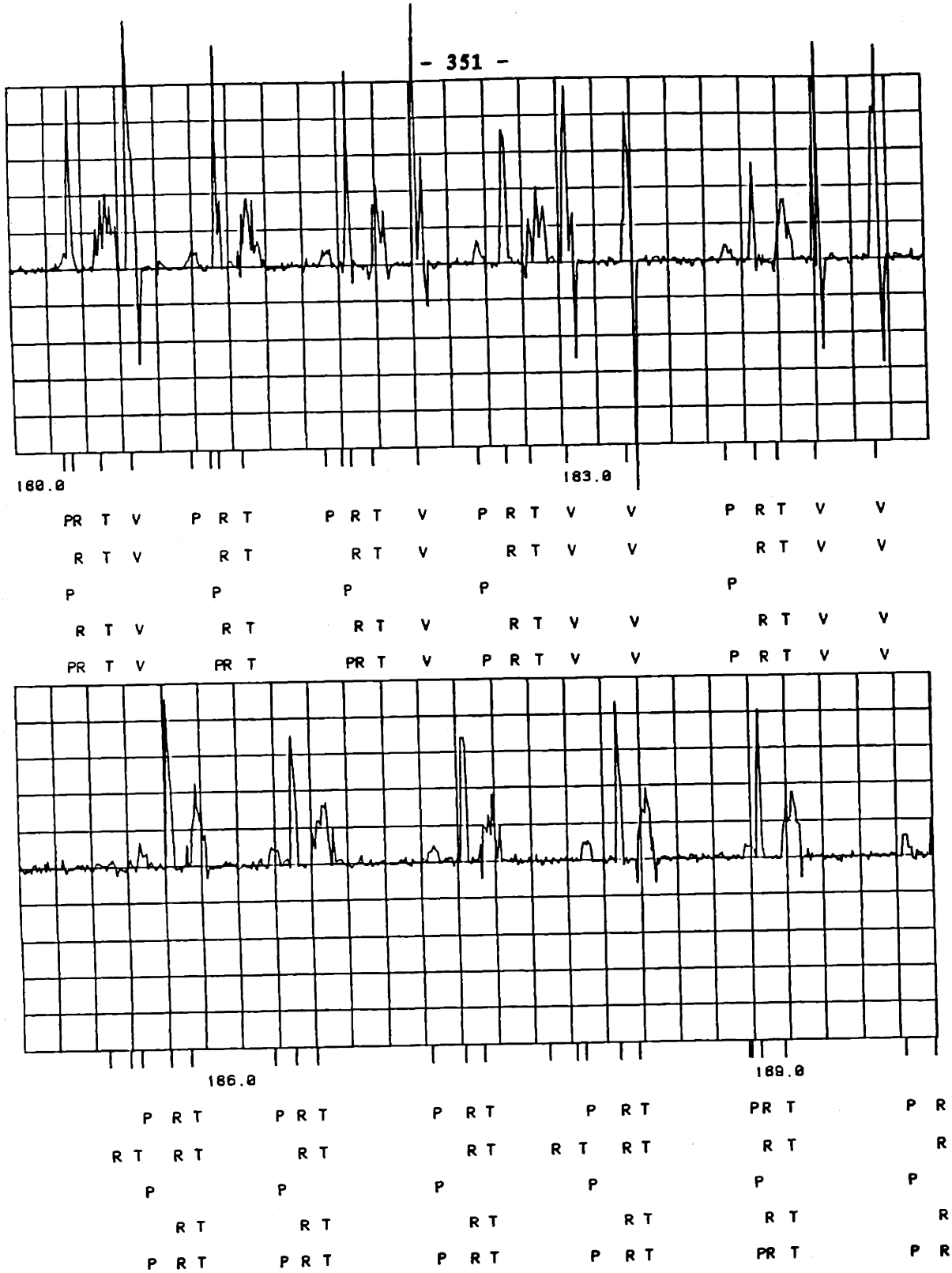


Figure 6.36 Continued.
Note the two errors (two extra R and T wave pairs) in LE 1 pass 0 near 185.5 and 187.9 seconds (both of which are corrected by LE 1 pass 2) and the phase-shift errors in LE 0 pass 1 near 181 and 181.7 seconds (resulting in P-R interval estimates that are too short). However, note also that LE 0 pass 1 did not make any false-positive or false-negative errors in spite of substantial PVC activity and errors in the LE 1 pass 0 estimate.

6.22. Delay-Line Topology

Every LE is a sequence of MAP estimators. As mentioned in Section 3.6, the choice of the MAP criterion has implications for the topology of the delay lines in the Markov chains.

The choice of MAP estimators has an important effect on the design of Markov chains. Specifically, there are many portions of the Markov chain which exist solely to arrange the proper time interval between two observable events--interactions or annotations. If there are two or more indistinguishable paths between these observable events then, as illustrated in the following example, the MAP estimate may have properties that the designer probably does not intend. (Here we assume that the designer's intent is to achieve some condition on the intervals between observable events since these are the measurable quantities).

Consider the example in Figure 6.37. Note that there are three ways to proceed between the two signatures, which are the observable events. One path (a->b->f) has length 2 and a priori probability .4. Two paths (a->c->d->f and a->c->e->f) have length 3 and each has a priori probability .3. In a simulation of the model, signatures S1 and S2 are separated by 2 cycles 40 percent of the time and 3 cycles 60 percent of the time. Presumably this is the designer's a priori expectation. However, because the 3 cycle path is divided into two indistinguishable subpaths, this is not the MAP estimator's view of the designer's intentions. Specifically, the MAP estimator must settle on a single trajectory. There are three choices, with a priori probabilities .4, .3, and .3. For the sake of illustration, assume that the likelihood, conditioned on the observations, of a length 2 interval

between signatures S1 and S2 equals the likelihood of a length 3 interval. Then the MAP estimator picks the a→b→f trajectory because it has the highest a priori probability of any of the three trajectories, and therefore the chosen interval has length 2. However, the designer, concerned with interval lengths, intended the choice of a length 3 interval since the a priori probability of a length 3 interval is .6 which is greater than .4. Because the length 3 interval was divided into two indistinguishable trajectories and the MAP estimator picks trajectories, not intervals, the estimator did not make the expected decision.

The prior example is admittedly artificial. However this type of behavior can occur in design models. Sometimes it is hidden, but in delay lines it is very obvious. Explicitly, all delay lines are sets of states in which paths are only distinguishable by their lengths. Therefore, if there are two paths of equal length then this type of behavior will occur. It is for this reason that design models use only delay lines with the tapped delay line topology since the alternative topology, the feedback topology, always introduces multiple indistinguishable paths.

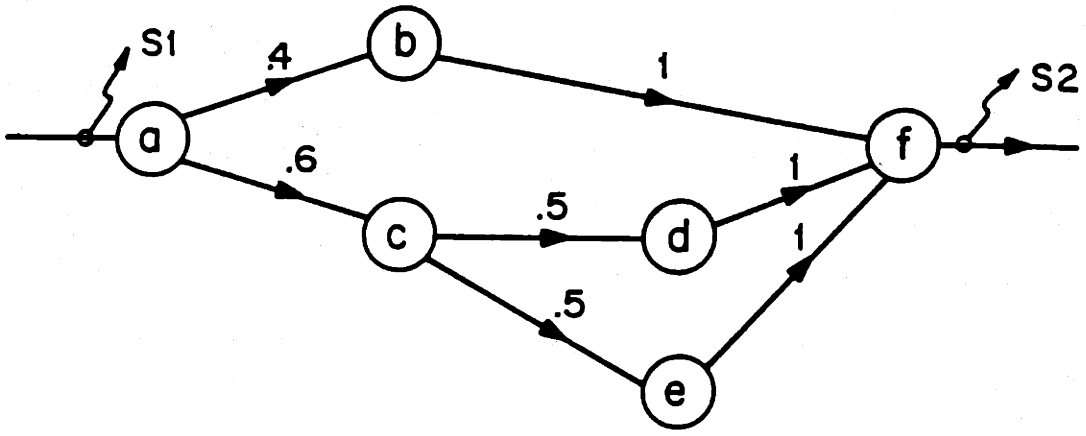


Figure 6.37 A Part of a Markov Chain Which Exhibits Indistinguishable Alternative Paths.

6.23. Utility of the P0-P4 Chains

We believe that the five generic chains P0-P4 are widely applicable. In this section, we attempt to demonstrate this by example. The example considered is the Wenckebach design model of Section 3.8. For simplicity, we ignore the T waves. Therefore, we are concerned with a design model with a block diagram as shown in Figure 6.38.

The basic idea of the design procedure is to symmetrize the design model, write down a symmetric estimator, and then simplify the estimator to reflect the asymmetries of the original design model. Symmetrizing gives a design model with a block diagram as shown in Figure 6.39.

A natural estimator for design models with the block diagram shown in Figure 6.39 is given in Figure 6.40. Initialization and refinement

passes are shown. The number of refinement passes required is not specified but each would have the structure shown in Figure 6.40. The global reconstruction is obvious. Note the symmetry of the estimator. In Figure 6.40, we have used simplified notation.

(1) Augmentation of interaction estimates according to the usual rules is used but not shown.

(2) The calculation of binary annotations is indicated by an arrow leaving the annotation label.

(3) For legibility, the ICS is shown in solid lines.

(4) For each P0-P2 chain, the corresponding design-model submodel is noted.

The first simplification step is to strike out those parts of the estimator dealing with interactions in the Figure 6.39 design model that do not occur in the Figure 6.38 design model. The resulting estimator, with the drawing slightly rearranged, is shown in Figure 6.41. In Figure 6.41, a type of chain labeled U, for unobservable, occurs twice in the initialization pass. This chain does not exist in the estimator and in the diagram is used only to easily indicate the nature of interactions. That is, in LE 0 pass 0, the occurrence of U makes it easy to indicate the C0-to-C1 interactions. The U chain is not necessary because these interactions are functions of the C0 subprocess state space alone.

The second simplification step is to strike out the S1 signature, as it is identically zero in the design model of Figure 6.38. Without

the S1 signature, LE 1 pass 0 is useless. Therefore, the estimate of the interactions that submodel 1 initiates with submodel 2 is now taken from LE 2 pass 0. The appropriate change in the augmenting information is made also. The resulting estimator is shown in Figure 6.42.

The last step is to take account of the fact that the SNR of S0 (i.e. the P wave) is much less than the SNR of S1 (i.e. the R wave). First consider the case when there is no S0. In this case, the information from the real observation (i.e. the ECG) enters the estimator through LE 2 exclusively. Therefore, the initialization process requires two passes to allow this information to propagate back through the submodels--first submodel 1, then submodel 0. We call this phenomenon ripple. The two initialization passes and the refinement pass of the estimator for the case when there is no S0 signature are shown in Figure 6.43.

In Figure 6.43, note that only the first initialization and last refinement passes need be performed for LE 2. That is, the results of the second initialization pass for LE 2, which is indicated in this figure, are never used so the pass can be deleted. Similarly for all but the last refinement pass of LE 2. This is related to a more fundamental problem, that is, it is not clear that taking 0-to-1 interaction estimates from LE 0 and 1-to-2 interaction estimates from LE 1 is a good idea. Specifically, that choice for the ICS implies that a second or later refinement pass does not use any information from the real observation to improve the estimates of LE 0 and LE 1. The obvious alternative, which was used in the initialization passes, is always to take the 0-to-1 interaction estimate from LE 1 and the 1-to-2 interaction

estimate from LE 2. This alternative is shown in Figure 6.44. Furthermore, the possibilities illustrated in Figure 6.43 and 6.44 can be mixed. Further experimental work is required in order to make a choice here.

Finally, consider the desired case where S0 exists but the SNR of S0 is much lower than the SNR of S1, that is, the Wenckebach case. For an initial design, either of the estimator structures shown in Figures 6.43 or 6.44 could be used, with the addition of the S0 signature from each C0 chain. Subtractors (i.e. P3 and P4 chains) could be added to either design in an obvious manner. Again, experimental work is required.

So far we have concentrated on the initialization and refinement passes. However, in previous estimators we have often used special designs for the final pass of the estimator. For example, the difference between Figures 6.31 and 6.32 is of this type. For the initialization and refinement passes shown in Figure 6.43, we could consider a five pass estimator (Initialization #1, Initialization #2, Refinement, Termination #1, and Termination #2) where the two termination passes are shown in Figure 6.45. The desirable aspect of the two passes shown in Figure 6.45 is that the final submodel 0 estimate implies interactions which are used to compute the final submodel 1 estimate which in turn implies interactions which are used to compute the final submodel 2 estimate. Note that this is not enough to guarantee complete consistency because of the P2 chains. However, it does match the causality of the design model. Again, experimentation is necessary. Note that this is another example of ripple. However, in this case the ripple is

"down" the model (i.e. submodel 0 to submodel 1 to submodel 2), while in the prior case it was in the reverse direction.

With this example, we have attempted to demonstrate the generic nature of the P0-P4 chains and simultaneously exhibit the important roles played by the designer and by experimentation in constructing simple high-performance estimators.

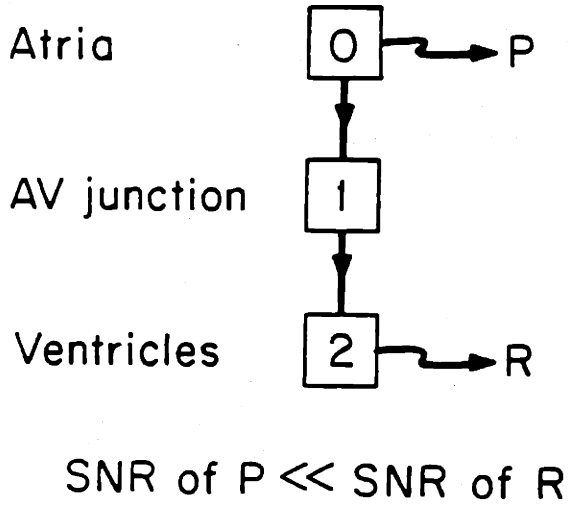
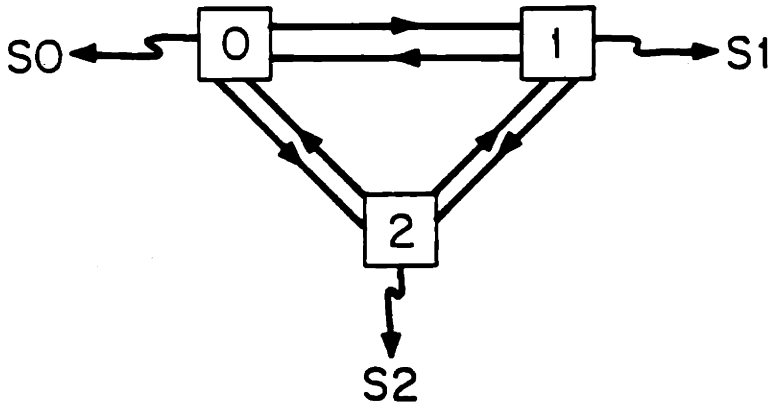


Figure 6.38 The Design Model.



SNR of SO ≈ SNR of S1 ≈ SNR of S2

Figure 6.39 The Symmetrized Design Model.

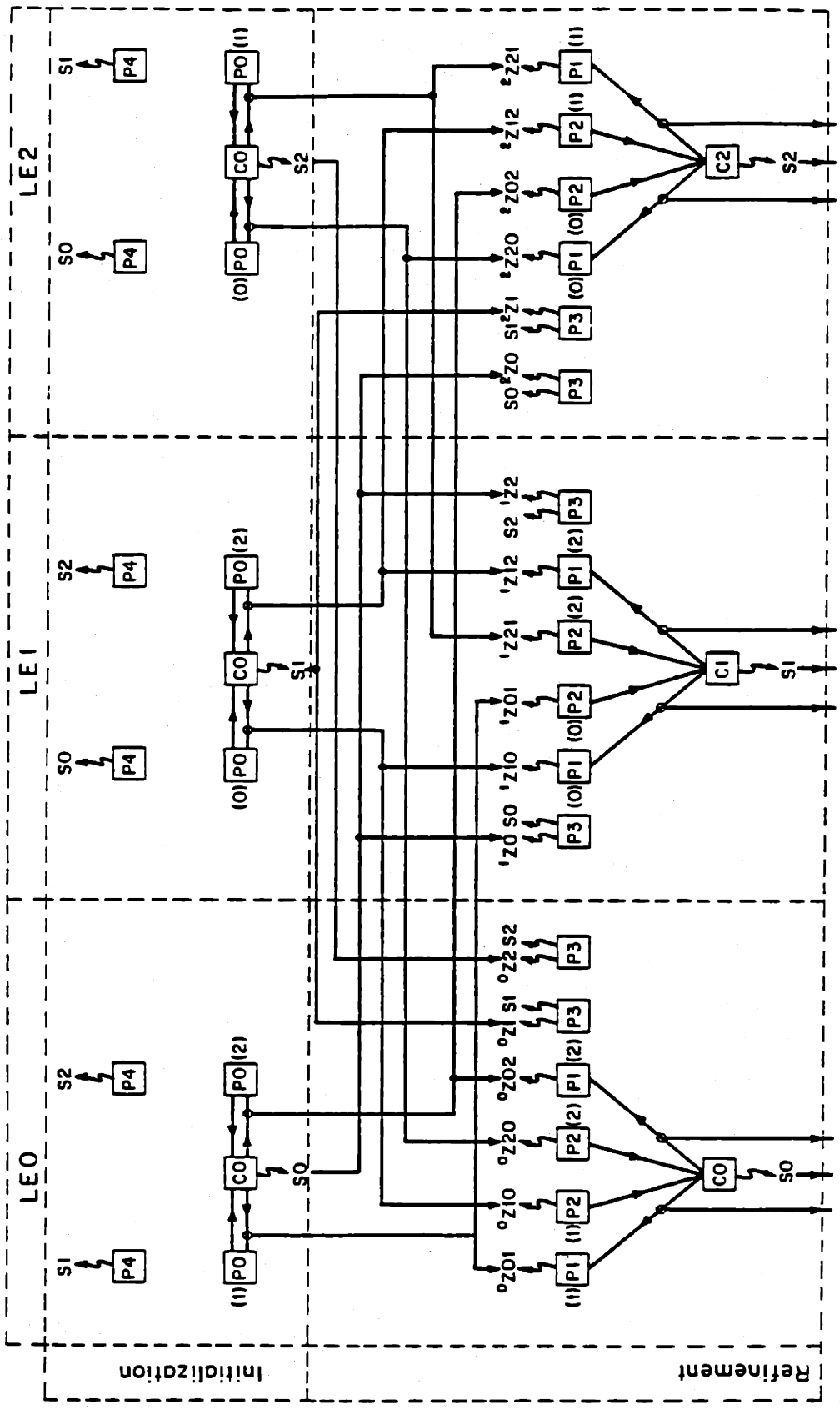


Figure 6.40 Estimator for the Symmetrized Design Model of Figure 6.39.

Section 6.23.

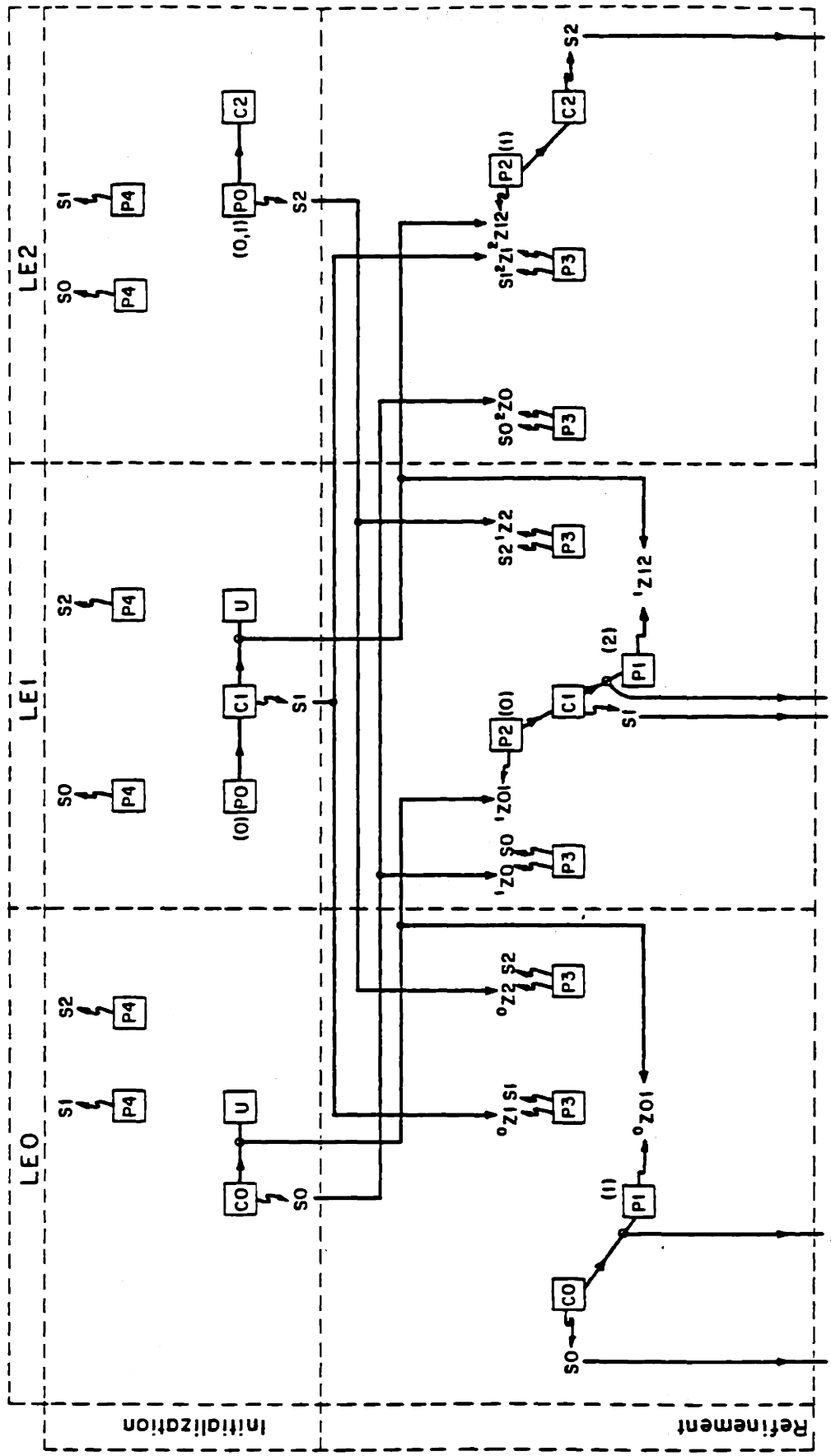


Figure 6.41 Simplification of the Estimator of Figure 6.40. The design model interactions are now unidirectional.

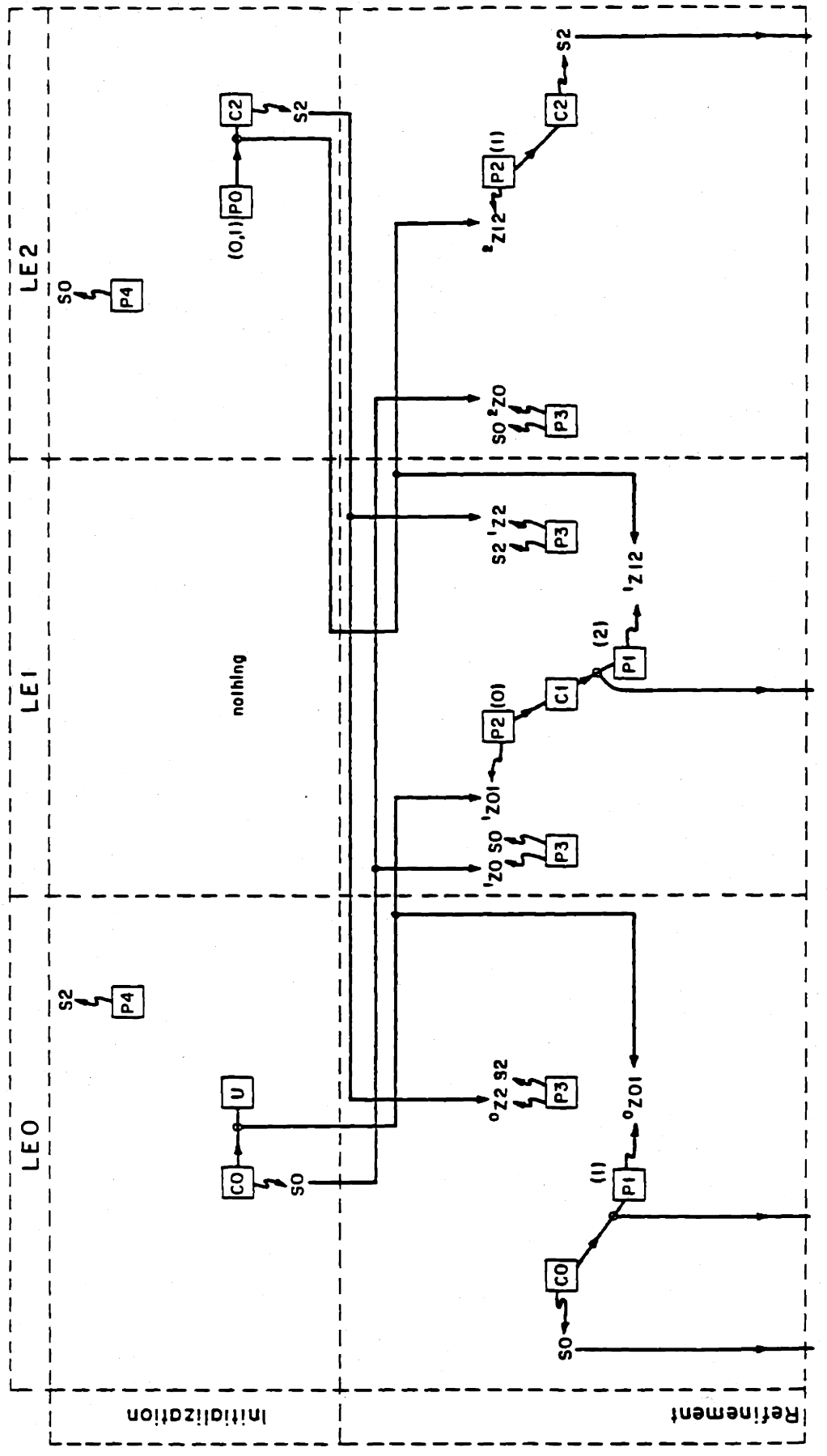


Figure 6.42 Simplification of the Estimator of Figure 6.41.
The S1 signature no longer exists.

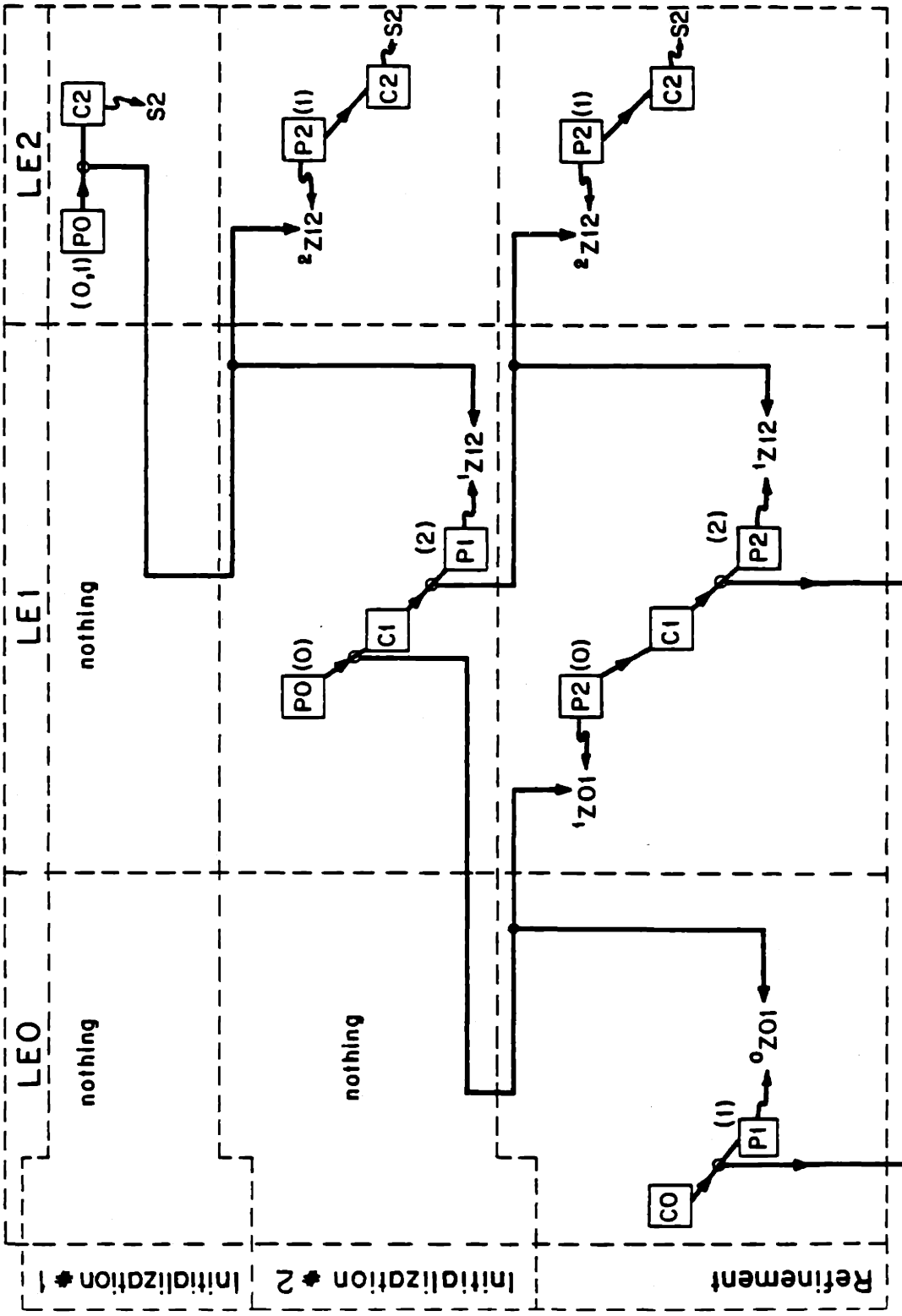


Figure 6.43 Simplification of the Estimator of Figure 6.42. The S0 signature no longer exists.

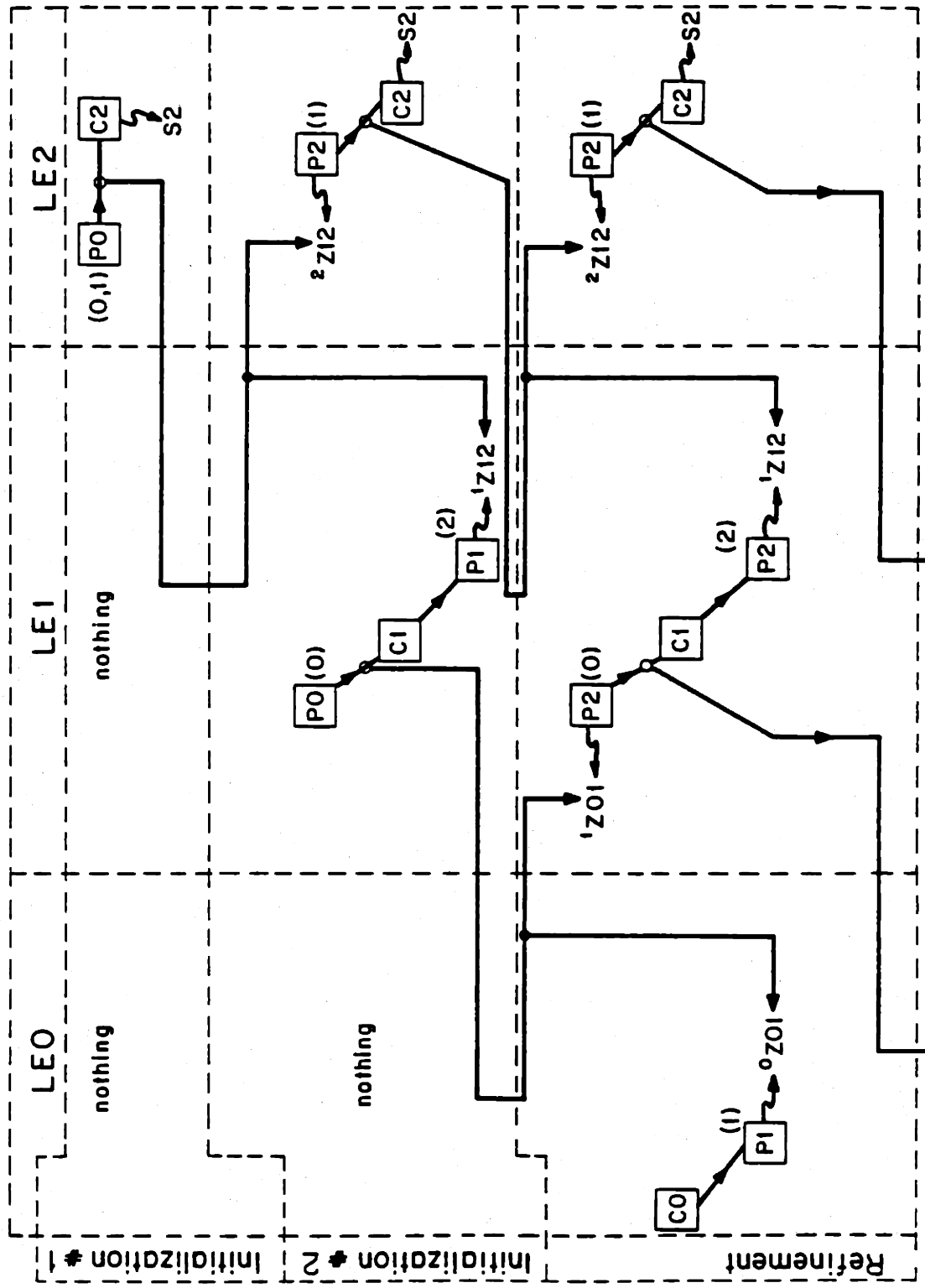


Figure 6.44 An Alternative to the Estimator of Figure 6.43. The interaction estimates are taken from different sources.

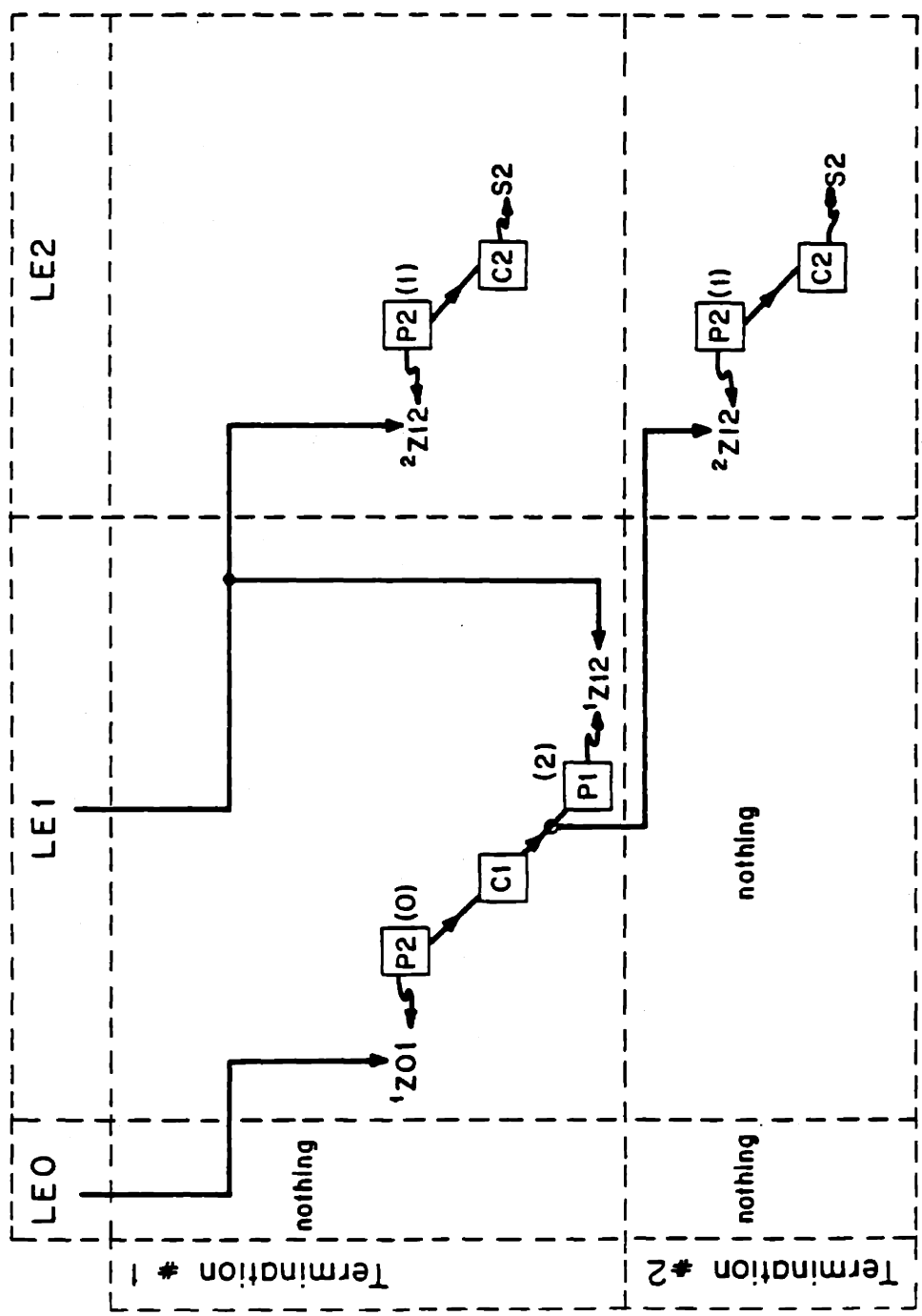


Figure 6.45 Possible Termination Passes for the Estimator of Figure 6.43.

Section 6.23.

6.24. Implementation Computation for the Nonoptimal and MAP-Based Estimators

In the suboptimal estimators described previously, each pass of each LE can be executed concurrently and, unlike many concurrent computations, the synchronization requirements are minimal. Furthermore, the ICS requires essentially no computation. For these reasons, we envision implementing these estimators on a multiple CPU computer,[2] with one CPU devoted to each pass of each LE. With this point of view, the appropriate measure of computation is the computation required by the computationally most intensive LE pass.[3]

The computation required by a Viterbi algorithm per stage[4] is essentially proportional to the average number of one-step tree extensions that must be considered at each stage. That is, at a given stage, the Viterbi algorithm tree has exactly one branch terminating on each accessible state. In order to find the minimum length extensions so that a branch terminates on every accessible state at the following stage, all positive probability transitions from accessible states at the current stage must be considered. We define a unit of computation to be the computation required to consider one such transition. The

---

[2] More precisely, we envision implementing these estimators on a multiple-instruction multiple-data stream computer.

[3] Excepting the results of the example, the remainder of this section assumes an understanding of the Viterbi algorithm.

[4] An efficient implementation of the Viterbi algorithm is applicable both to the global MAP estimator and to each LE pass of a suboptimal estimator. The absolute decrease in computation may differ, however, due to the differing number of submodels and the differing cardinality of the subprocess state spaces in each computation. The discussion here refers to the implementation that was actually used.

actual amount of computation represented by this unit changes depending, for example, on the dimension of the observation vector. However, such changes do not affect the comparison of the suboptimal and global MAP estimators.

In general, it is difficult to calculate which states are accessible at a given stage. That is, while it is easy to write down the relevant equations, the exact computation involves the entire global state space, and we have not found an approximation in terms of the subprocess state spaces. In the models we have considered, both design models and LEPMs, the situation is somewhat simplified because after a finite number of steps, the accessible states are exactly the persistent states of the global Markov chain. Even ignoring the initial transient, the calculation of the accessible states via the calculation of the persistent states must still be done in the global state space. For this reason, we only solve this problem for specific examples.

The calculation of the number of positive probability transitions from a given global state is identical to counting the number of nonzero elements in the corresponding row of the global tpm.[5] This number is called the branching ratio for that state. For ease in computation and interpretation, we now make two simplifying assumptions.

(1) Assume that if a transition in a particular subprocess is allowed under any impinging interaction, then it is allowed under all impinging interactions. This assumption decouples the subprocesses:

[5] Here we assume that a transition can be excluded only by a zero in the global tpm, i.e. not by the observations.

the number of possible global transitions is the product over the subprocesses of the number of possible subprocess transitions. Note that this assumption is pessimistic--at worst it overstates the number of possible transitions.

(2) Assume, with respect to the branching ratios, that accessible states are no different than any other states. Therefore, we will compute the total number of branches (i.e. the sum of the branching ratios over all accessible states) as the sum of the branching ratios over <u>all</u> states times the ratio of the number of accessible states to the number of total states.

With these two assumptions, the number of branches that must be considered after the finite initial transient is the number of persistent states times the product of the average branching ratios for each subprocess.

To understand this result in more detail, let $c(n,\pi_0)$ denote the number of branches that must be evaluating in going from stage n to stage n+1 when the initial-condition pmf is π_0. Similarly, let $A(n,\pi_0)$ be the set of accessible states at time n when the initial-condition pmf is π_0. Finally let b(x) be the branching ratio for global state x. Then,

$$c(n,\pi_0)=\sum_{x\in A(n,\pi_0)} b(x).$$

Applying the first assumption we can write b(x) as $\prod_{i=0}^{N-1} b^i(x^i)$ where x^i, $i\in\{0,\ldots,N-1\}$ are the subprocess states and $b^i(x^i)$, $i\in\{0,\ldots,N-1\}$ is the

subprocess branching ratio for subprocess state x^i. Therefore,

$$c(n,\pi_0) = \sum_{x \in A(n,\pi_0)} \prod_{i=0}^{N-1} b^i(x^i).$$

Applying the second assumption, where S is the global state space and S^i are the subprocess state spaces, gives

$$c(n,\pi_0) = \frac{\text{card}(A(n,\pi_0))}{\text{card}(S)} \sum_{x \in S} \prod_{i=0}^{N-1} b^i(x^i)$$

$$= \frac{\text{card}(A(n,\pi_0))}{\text{card}(S)} \sum_{x^0 \in S^0} \cdots \sum_{x^{N-1} \in S^{N-1}} \prod_{i=0}^{N-1} b^i(x^i)$$

$$= \text{card}(A(n,\pi_0)) \prod_{i=0}^{N-1} \left[\frac{1}{\text{card}(S^i)} \sum_{x^i \in S^i} b^i(x^i) \right]$$

$$= \text{card}(A(n,\pi_0)) \prod_{i=0}^{N-1} \bar{b}^i$$

where $\bar{b}^i$ is the average subprocess branching ratio for the i^{th} subprocess. Finally, consider only those n greater than the duration of the finite-length transient, giving

$$c = \text{card}(F) \prod_{i=0}^{N-1} \bar{b}^i$$

where $c(n,\pi_0)$ no longer depends on n or π_0 because the transient is past and where F is the set of persistent states.

Section 6.24.

Now consider two alternative estimators for the Wenckebach example of Section 3.8: the global MAP estimator and the suboptimal estimator of Section 6.23. First we consider the global MAP estimator. After augmenting the S^0 and S^2 subprocess state spaces in order to accommodate the multiple Markov-chain-cycle P and T wave signatures, the global state space has 31x58x24 = 43,152 states of which 551 are persistent. (Before augmentation, there are 30x38x15 = 26,100 states of which 421 are persistent). After augmentation, the average subprocess branching ratios are $\bar{b}^0 \simeq 1.23$, $\bar{b}^1 \simeq 1.07$, and $\bar{b}^2 \simeq 1.58$. Therefore, the computational cost c^{global} of the global MAP estimator is

$$c^{global} = 551x(1.23x1.07x1.58) = 1143.$$

Now consider the suboptimal estimator of Section 6.23. The computationally most expensive LE pass shown in any of Figures 6.43-6.45 is intuitively the refinement pass for LE 1 since this LE pass has the most and the largest chains. (Note that our estimates of computational cost do not differentiate between the cost of processing discrete pseudo-observations versus Gaussian observations). In the LEPM for this pass there are 58x3x3 = 522 states of which 160 are persistent. The average subprocess branching ratios are $\bar{b}^{C1} \simeq 1.07$, $\bar{b}^{P1} = 1$, and $\bar{b}^{P2} \simeq 1.33$. (Note that the average branching ratio for the P1 chain is exactly 1 because the P1 chain simply acts as a shift register memory for incoming interactions). Therefore, the computational cost c^{subopt} of the suboptimal estimator is

$$c^{subopt}=160x(1.07x1x1.33)=228.$$

Therefore, for this example, the suboptimal estimator is less costly than the global MAP estimator by a factor of 1143/228 = 5.01. These figures grow even more in favor of the suboptimal estimator as the number of subprocesses increases.

7. Conclusions and Future Research

In this chapter we summarize the contributions of this thesis and indicate several directions for future research.

7.1. Conclusions

The contributions of this thesis fall into two categories: modeling and signal processing.

We believe that the modeling ideas of this thesis are of interest both in the electrocardiogram application and more generally in any event-oriented signal-processing problem.

With respect to the electrocardiogram application, our desire was to perform rhythm classification at a detailed level, for example, the level of a ladder diagram description. For this purpose no existing electrocardiogram model was sufficient. We believe that the limitations of existing models stem from two problems: models devised to study physiological issues are too complex for signal processing while models created for signal processing lack sufficient internal detail. Therefore, in our models we have emphasized a balance between realistic physical mechanisms and practicality for signal processing. We believe that the models presented in Sections 3.7, 3.8, and 6.21 demonstrate that we can capture the physical mechanisms of arrhythmias in moderate detail with reasonably simple models. In this thesis we have not demonstrated that practical high-performance signal-processing algorithms can be constructed based on these models. However, preliminary simulation results, especially for the relatively realistic model of Section 6.21

are encouraging.

From a signal-processing point of view, the interesting aspects of the class of models described in this thesis are all related to the state space and tpm of the Markov chain. Specifically, the interesting ideas are

(1) the factorization of the state space into submodels,

(2) the definition of sparse interactions between different submodels,

(3) the definition of the tpm in terms of multiple tpms for each submodel by the assumption that submodel state transitions depend only on the previous submodel state and the impinging interactions, and

(4) the further decomposition of each submodel into networks of DLs and RDLs.

These ideas provide a way of structuring a huge Markov chain in a manner that seems natural for problems concerned with discrete events in asynchronous processes.

The work performed in the signal-processing area was strongly motivated by the electrocardiogram rhythm classification problem.

(1) We described a mathematical problem (the wave tracking problem) which is one way of formalizing certain transient rhythm classification problems.

(2) We discussed performance measures for the problem of (1). These performance measures were defined on discrete events--the initiation of waves in the ECG--because the type and time of the waves characterize the rhythm. The need to match true and estimated events in spite of phase shifts made the definition of these performance measures a complex issue. Additional performance measures concerned more with the internal workings of the estimators were also defined. These additional performance measures were also defined on discrete events--the inter-subprocess interactions--because the interactions are central in the local estimator philosophy that we have pursued.

(3) We described a second mathematical problem (MAP state-trajectory estimation) related to the problem of (1) and for which easily implementable solutions are known (the Viterbi algorithm). From the solution of this problem, a solution of (1) can be easily derived.

(4) In an attempt to reduce the computational and storage requirements of the implementation, to achieve a more robust estimator, to achieve an estimator more easily interpreted in electrocardiogram terms, and to shed light on problems for which optimal solutions are utterly intractable we developed a systematic procedure for the construction of suboptimal solutions to the problem of (2). These estimators emphasize the spatial decomposition in the model. (The ideas of time-scale and hierarchical levels are not used). These estimators provide a concrete example of ideas of aggregation, communication between cooperating processes, and constructing global

results from local computations. As such, they may be of interest independent of both the electrocardiogram application and the mathematical models.

7.2. Future Research

We divide the discussion of future research directions into two areas: the electrocardiogram application and the signal-processing ideas. Furthermore, the signal-processing directions are divided into unfinished issues within the present scheme and more fundamental alterations of the present scheme.

A practical electrocardiogram classification method based on the ideas of this thesis must address a number of issues.

(1) The heart rate is the inverse of the mean of the P-P interarrival time. This is specified by the Markov chain. In practice, one would not want to treat different heart rates as different persistent rhythm hypotheses. Furthermore, one would not want to model wide ranges of heart rate by using one very uncertain P-P interval pmf. Therefore, a means of estimating the current heart rate and adapting the Markov chain tpm is necessary.

(2) Similar to (1), time-varying morphology must be accounted for. Here, often-used ideas concerning libraries of templates may be useful.

In both (1) and (2), we believe that a useful perspective is to consider the time-variation to be a low frequency modulation of the algorithms described in this thesis which would be the high-frequency core of the

complete system.

(3) We have used very idealized noise models. Especially lacking have been models for bandlimited noise, for "shot noise" like phenomena, and for baseline shift phenomena. It is possible to model these phenomena within the framework described in this thesis. For example, Figure 7.1 shows a submodel for a baseline shift. However, it is possible that a better approach would be to model rhythms in detail as done here but use a dramatically less detailed model for the noise.

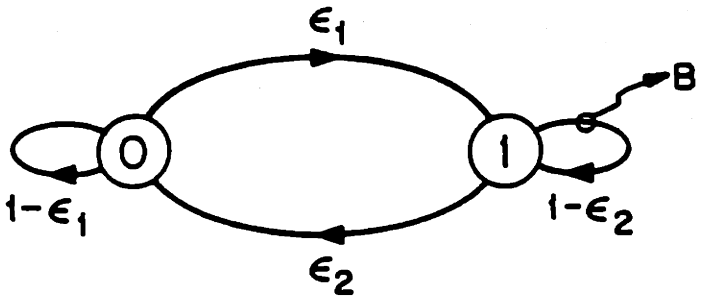
(4) The algorithms proposed in this thesis require a large amount of numerical computation and therefore careful implementations exploiting all of the structure (e.g. the submodel structure of the tpm) are necessary for a practical system.

(5) The persistent rhythm problem must also be solved. We conjectured (in Section 4.2) that a solution to the wave-tracking problem considered in this thesis could be used as an important component of a practical solution to the persistent rhythm problem.

(6) Additional rhythms must be considered (e.g. Wenckebach).

Within the present scheme, there are a number of open signal-processing issues.

(1) As mentioned previously, fast implementations of the Viterbi algorithm are necessary, implementations exploiting the structure of the problem (e.g. the submodel structure of the tpm) and using efficient programming techniques.



The submodel is autonomous and has $\varepsilon_1 << 1$, $\varepsilon_2 << 1$.

B (<u>baseline</u>) wave: 2 samples long, each sample having mean m_B and standard deviation σ_B.

The mean of the output of this submodel is a signal like

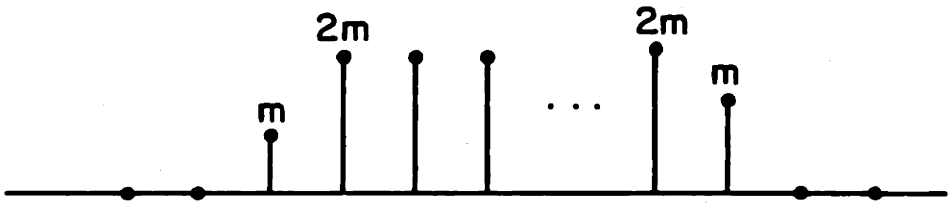


Figure 7.1 A Baseline-Shift Submodel.

(2) Methods for the calculation of the observation pmfs for the pseudo-observations without simulation.

(3) Methods for error analysis without simulation.

(4) Consideration of additional performance statistics. For example, the lag between the time n when the Viterbi algorithm receives the n^{th} observation and the time when the state-trajectory estimate is uniquely extended to time n. For normal rhythm, this should be one or two beats while for Wenckebach is might be as long as one or two Wenckebach cycles.

(5) Ideas for further development of the P0 chain:

(a) In the chain of Figure 4.7, understand the reason for the optimal value of p.

(b) Consider alternative choices of subprocess state space. Assuming that the number of transmitted interactions is to be preserved, then the only possibility is a less severe aggregation--i.e. a subprocess state space with larger cardinality in which residency in either of two (or more) different states results in the transmission of the same interaction.

(c) Consider alternative preserved statistics, including statistics defined on events external to the submodel being aggregated. (An example of a statistic defined on events external to the submodel being aggregated is the successful excitation statistic of Section 6.8).

(d) Have the aggregated submodel receive interactions even if the true submodel does not. (See the discussion in Section 6.7 where we note the correlation, in the DM1 class design models that we presented, between events in submodel 0 and submodel 1. Because of this correlation, an aggregated version of submodel 0 which receives interactions from submodel 1, in spite of the fact that the true submodel 0 does not, might yield superior performance).

(6) Many of the ideas in (5) are also relevant to the further development of the P4 chain.

(7) More generally, all of the ideas in (5) and (6) are concrete approaches to a general problem: how to aggregate local remainder models when the objective is signal processing performance.

(8) Find and examine situations in which the P3 and P4 chains (i.e. the subtractor chains) improve performance.

There are several obvious directions for modifying the present scheme in significant ways:

(1) Communicate time-varying likelihood ratios between LEs rather than pseudo-observations. (Note that a realization of a pseudo-observation is really just an index into a set of prestored, constant likelihood ratios--the pseudo-observation observation pmfs. Thus we are essentially proposing that these pmfs be time-varying and data dependent).

(2) Communicate more than one estimate between LEs. That is, for example, communicate the "best" known estimate and also a second estimate which, while not as probable, represents a radically different set of events. Such communication would allow the algorithm to hedge. Note that exchange of multiple estimates requires alterations in both the LE that transmits the information (so that it can compute the multiple estimates) and the LE that receives the information (so that it can make use of the multiple estimates).

(3) Use a different method of calculating state-trajectory estimates through the LEPM, that is, a replacement for the Viterbi algorithm. A replacement need not calculate the MAP estimate and might emphasize ideas of merging trajectories and/or extending only trajectories that seem likely candidates.

(4) Derive wave-tracking estimates directly rather than as a byproduct of state-trajectory estimates.

(5) Exploit the other structure in the model (i.e. the hierarchical levels and time-scale decompositions). The work described in this thesis only exploits the spatial decomposition.

(6) Develop a more fundamental approach to global reconstruction which might, for example, guarantee complete consistency.

8. Appendix A: The Association Rule

This appendix contains a detailed definition of the association rule. The association rule is a solution to the problem of matching N^t truth annotations with N^e estimated annotations where N^t and N^e are arbitrary. Let $Z_n=\{0,\ldots,n-1\}$. For the i^{th} truth annotation, $i\in Z_{N^t-1}$, let n_i^t be the time ($n_i^t \geq n_j^t$ if $i>j$) and α_i^t be the type. Similarly for the estimated annotations: $i\in Z_{N^e-1}$, n_i^e ($n_i^e \geq n_j^e$ if $i>j$), and α_i^e.

A matching is a map

$$f: Z_{N^e-1} \rightarrow Z_{N^t-1} \{\text{false positive}\}$$

which takes an estimated annotation to the corresponding truth annotation or false positive. The unmatched truth annotations are false negatives. The space of admissible matches, denoted F, is constrained only by the requirement for one-to-one matchings. Therefore

$$F=\{f: f(i)\neq f(j)\ \forall\ i,j \text{ unless } f(i)=\text{false positive}\}.$$

Let E(f) be the set of estimated annotations which are matched:

$$E(f)=\{i\in Z_{N^e-1} : f(i)\neq \text{false positive}\}.$$

Let T(f) be the set of truth annotations which are matched:

$$T(f)=f(E(f)).$$

The reward function

$$r:\ F\rightarrow R$$

is defined by

$$r(f)=\sum_{i\in E(f)} r_m(a_i^e, a_{f(i)}^t, n_i^e-n_{f(i)}^t)$$

$$+\sum_{\substack{i\in Z\\ N^e-1}}{}_{-E(f)}\, r_{fp}(a_i^e) + \sum_{\substack{i\in Z\\ N^t-1}}{}_{-T(f)}\, r_{fn}(a_i^t).$$

The sum over r_m deals with matched annotations while the sums in r_{fp} and r_{fn} deal with false positive and false negative annotations respectively. These functions are defined as follows:

(1) r_{fp} and r_{fn}:

$$r_{fp}(.)=r_{fn}(.)=0.$$

(2) r_m: if the signatures implied by the annotations do not overlap then

$$r_m(a^e, a^t, \tau)=-\infty$$

else

$$r_m(a^e,a^t,\tau)=\frac{\sum_n q(n;a^t)\hat{q}(n-\tau;a^e)}{\{\sum_n [q(n;a^t)]^2\}^{1/2}\{\sum_n [\hat{q}(n;a^e)]^2\}^{1/2}}.$$

(Here $q(n;a)$ is $q^i_{j,k}(n)$ for the appropriate i,j,k and likewise for $\hat{q}(n;a)$). (The definition of r_m actually employed is slightly different in three respects from the definition immediately above. However, at this point the motivation for these differences has not been provided. Therefore, the description of the differences is in Appendix B).

Define

$$v=\max_{f\in F} r(f)$$

and

$$F^*=\{f\in F: r(f)=v\}.$$

Because this optimization does not necessarily have a unique solution, F^* will typically contain more than one point. Therefore, define the final match f^* as the value selected from a lexicographic order over the elements of F^*.

This problem can be formulated as an integer programming problem. The problem has $(N^e+1)(N^t+1)-1$ variables denoted $x_{i,j}$, $i\in Z_{N^e+1}$, $j\in Z_{N^t+1}$ where the pair $i=N^e$, $j=N^t$ is excluded. Each $x_{i,j}$ is a binary variable.

For $i \in Z_{N^e}$, $j \in Z_{N^t}$, $x_{i,j}$ takes the value 1 if estimated annotation i is matched with truth annotation j. For $i=N^e$, $x_{i,j}$ takes the value 1 if truth annotation j is a false negative. Similarly for $j=N^t$, $x_{i,j}$ takes the value 1 if estimated annotation i is a false positive. An example of $x_{i,j}$ is shown in Figure A.1.

The requirement of one-to-one matchings implies that the sum of each of the first N^e rows and also each of the first N^t columns is 1. The final row (column) need not sum to 1 because there can be more than one false negative (false positive). Along with the usual integer programming constraint of $x_{i,j} \geq 0$, these N^e+N^t constraints also force each $x_{i,j}$ to be binary.

The cost z (which is just the negative of the previous reward r) can be written as a linear function of the $x_{i,j}$'s in the following manner:

		truth annotations											
		0	1	2	3	4	5	6	7	8	9	10	fp
estimated annotations	0	1											
	1		1										
	2			1									
	3												1
	4				1								
	5												1
	6						1						
	7												1
	8							1					
	9								1				
	fn	1				1			1		1		

Figure A.1 Example $x_{i,j}$ Array for the Integer Programming Problem.
fn is false negative, fp is false positive.

$$z = - \sum_{i \in Z_{N^e}} \sum_{j \in Z_{N^t}} r_m(a_i^e, a_j^t, n_i^e - n_j^t) x_{i,j}$$

$$+ - \sum_{i \in Z_{N^e-1}} r_{fp}(a_i^e) x_{i,N^t} + - \sum_{j \in Z_{N^t-1}} r_{fn}(a_j^t) x_{N^e,j}.$$

In summary, the integer programming formulation is:

$$\text{minimize } z = - \sum_{i \in Z_{N^e}} \sum_{j \in Z_{N^t}} r_m(a_i^e, a_j^t, n_i^e - n_j^t) x_{i,j}$$

$$+ - \sum_{i \in Z_{N^e-1}} r_{fp}(a_i^e) x_{i,N^t} + - \sum_{j \in Z_{N^t-1}} r_{fn}(a_j^t) x_{N^e,j}.$$

subject to

$$x_{i,j} \in Z, \; x_{i,j} \geq 0,$$

$$\sum_{i \in Z_{N^e}} x_{i,j} = 1, \; j \in Z_{N^t}$$

$$\sum_{j \in Z_{N^t}} x_{i,j} = 1, \; i \in Z_{N^e}.$$

For any reasonable length ECG, N^e and N^t are large and thus this is an enormous integer programming problem. For example, a one hour ECG might have 4×10^3 beats leading to 16×10^6 variables if each beat contributed just one annotation. However, the problem has a great deal of structure. For example, as illustrated in Figure A.1, the 1's in the $x_{i,j}$ matrix only occur in the final rows and columns and in a band near the diagonal. This behavior occurs because the truth and estimated annotations are listed in temporal order and r_m allows only annotations with overlapping signatures to be matched. Therefore, by testing the relative times of the truth and estimated annotations, a large fraction of the $x_{i,j}$ can be set to zero immediately.

Although there exists a vast literature concerning integer programming problems, we have not pursued this approach because it seemed tangential to the main thrust of the thesis. Rather, we have devised a simplier approximate solution which is described in Appendix C. All of the results presented in this thesis are based on approximately optimal matches calculated by this method.

9. Appendix B: Modifications to the Definition of r_m in the Association Rule

This appendix describes three modifications to the definition of r_m that was given in Appendix A. Recall that r_m is the reward for a match between a pair of annotations at some phase shift in the association rule. This modified definition is used to calculate the results presented in this thesis.

The first modification in the definition of r_m is the optional ability to require matches with zero phase shift. In this case, the portion of the definition of r_m stating

"if the signatures implied by the annotations do not overlap then

$$r_m(a^e, a^t, \tau) = -\infty"$$

is changed to

"if $\tau \neq 0$ then

$$r_m(a^e, a^t, \tau) = -\infty".$$

In addition to a duration and the mean and covariance functions, each annotation also carries a label such as "P" or "R" for P wave or R wave respectively. Two problems motivate the remaining two modifications in the definition of r_m from Appendix A.

Consider a model with just two different waves, a P wave and an

R wave. Assume that the morphology is fairly normal so that both P and R waves have the same polarity (i.e. the maximum mean values of the two waves have the same sign). Assume that the model is highly idealized so that both signatures are only one sample long. Then r_m will only take on the two values

$-\infty$ (if the two signatures do not overlap, i.e. a nonzero phase shift)

and

1 (if the two signatures overlap, i.e. occur at the same time).

But this is undesirable. It is obviously preferable to match $\hat{P}$ with P and $\hat{R}$ with R rather than $\hat{P}$ with R and $\hat{R}$ with P. To repair this problem, redefine r_m for certain conditions. Specifically, if r_m calculated from the Appendix A definition takes the value 1 and the first character of the labels for the pair of annotations are not the same, then redefine r_m to be 1-DELTA where DELTA is (approximately) the smallest single-precision floating-point number such that 1-DELTA (in single-precision floating-point arithmetic) is less than 1. We have used DELTA=10^{-6}.

Occasionally zero signatures that are deterministically coupled to nonzero signatures are introduced. These are useful when investigating phase errors. In these situations matches with zero phase shift are required. When r_m is evaluated at one (or a pair) of these annotations with zero signatures, it can only take on one of the two values

$-\infty$ (if the signatures do not overlap)

and

0 (if the signatures do overlap).

But this is undesirable because if the zero signature is to measure phase shift relative to (for example) P waves, it is preferable to match

a P wave with it rather than an R wave. Therefore r_m is redefined for certain conditions. Specifically, if r_m calculated from the Appendix A definition takes the value 0 and if either annotation label is the leading string of the other annotation label then redefine r_m as EPSILON where EPSILON is (approximately) the smallest positive single-precision floating-point number. We have actually used EPSILON=10^{-10} which is much larger.

10. Appendix C: The Suboptimal Matching Algorithm

This appendix contains a description of our approximate solution to the annotation matching problem of Section 4.3.1. The primary performance criteria considered in this thesis are false positive and false negative probabilities. For these probabilities, the approximate solution described in this appendix is pessimistic.

The solution is computed in several steps.

(1) Compute a maximal set of overlapping annotations with respect to the next unmatched estimated annotation. That is, enter the next unmatched estimated annotation as the first element in the set. Then, for each truth annotation that overlaps[1] with an estimated annotation in the set, add to the set that truth annotation and all further estimated annotations that overlap with that truth annotation. Repeat until no further truth annotations overlap with any estimated annotation in the set. In theory, this procedure may stop only when all truth and estimated annotations are in the set. However, in practice this does not occur because the annotations are relatively well separated in time. In further steps of the algorithm, only matches between truth and estimated annotations within the maximal set are considered. This does not represent an approximation because the annotation matching problem of Section 4.3.1 only allows matches between overlapping annotations.

---

[1] When we say that two annotations overlap, we actually mean that the mean of the signature for the earlier annotation overlaps the mean of the signature for the later annotation, where in both means, leading padding zeros are ignored. Padding with leading zeros was discussed in Section 3.4.

(2) For each estimated annotation, find (by exhaustive search through the maximal set) the truth annotation which gives the greatest reward and tentatively match the two. In the case where the estimated annotation overlaps with no truth annotations, the estimated annotation is left unmatched. Note that several estimated annotations can be tentatively matched with the same truth annotation. Such conflicts are resolved in later steps.

(3) For each truth annotation that is tentatively matched with several estimated annotations, allow only the tentative match that gives the greatest reward and return the remaining estimated annotations to the unmatched category.

(4)

(a) Repeat step (2) considering only estimated and truth annotations that are not tentatively matched. If no new tentative matches are created (i.e. all of the estimated annotations previously unmatched remain unmatched), then go to step (5). Otherwise, continue with steps (b) and (c).

(b) Repeat step (3).

(c) Return to step (a).

(5) Accept all existing tentative matches and label all unmatched estimated annotations as false positives and all unmatched truth annotations as false negatives. Return to step (1) to compute the next maximal set.

Note, in terms of Appendix A, that the resulting match is one-to-one and that the reward is finite (i.e. there are no $-\infty$ terms from matches between nonoverlapping annotations). However, a match with a larger reward may exist, though it seems unlikely in the specific examples we have considered in this thesis because the annotations in these examples are well-separated in time.

11. References

(Anderson, 1979)
B. D. O. Anderson and J. B. Moore, Optimal Filtering (Englewood Clifts, NJ: Prentice-Hall, Inc., 1979).

(Baker, 1974)
C. M. Baker and T. C. Pilkington, "Quasi-Linearization and Inverse Electrocardiology", IEEE Trans. on Biomed. Eng., vol. BME-21, no. 6, pp. 460-468, Nov., 1974.

(Baldwin, 1979)
A. F. Baldwin III and S. Rush, "Optimization of the Locations of Multiple-Dipole Heart Generators in a Simple Torso Model", Med. and Biol. Eng. and Comput. (G. B.), vol. 17, no. 5, pp. 569-577, Sept. 1979.

(Barr, 1969)
R. C. Barr and T. C. Pilkington, "Computing Inverse Solutions for an On-Off Heart Model", IEEE Trans. on Biomed. Eng., vol. BME-16, no. 3, pp. 205-214, July, 1969.

(Barr, 1970)
R. C. Barr, T. C. Pilkington, J. P. Boineau, and C. L. Rogers, "An Inverse Electrocardiographic Solution with an On-Off Model", IEEE Trans. on Biomed. Eng., vol. BME-17, no. 1, pp. 49-57, Jan., 1970.

(Barr, 1971)
R. C. Barr, M. S. Spach, and G. S. Herman-Giddens, "Selection of the Number and Positions of Measuring Locations for Electrocardiology", IEEE Trans. on Biomed. Eng., vol. BME-18, no. 2, pp. 125-138, March, 1971.

(Barr, 1977)
R. C. Barr, M. Ramsey III, and M. S. Spach, "Relating Epicardial to Body Surface Potential Distributions by Means of Transfer Coefficients Based on Geometry Measurements", IEEE Trans. on Biomed. Eng., vol. BME-24, no. 1, pp. 1-11, Jan., 1977.

(Borjesson, 1982)
P. O. Borjesson, O. Pahlm, L. Sornmo, and M.-E. Nygards, "Adaptive QRS Detection Based on Maximum A Posteriori Estimation", IEEE Trans. on Biomed. Eng., vol. BME-29, no. 5, pp. 341-351, May, 1982.

(Brody, 1972)
D. A. Brody and J. A. Hight, "Test of An Inverse Electrocardiographic Solution Based on Accurately Determined Model Data", IEEE Trans. on Biomed. Eng., vol. BME-19, no. 3, pp. 221-228, May, 1972.

(Chou, 1979)
T.-C. Chou, Electrocardiology in Clinical Practice (New York: Grune and Stratton, 1979).

(Cohn, 1982)
R. L. Cohn, S. Rush, and E. Lepeschkin, "Theoretical Analyses and Computer Simulation of ECG Ventricular Gradient and Recovery Waveforms", IEEE Trans. on Biomed. Eng., vol. BME-29, no. 6, pp. 413-423, June, 1982.

(Cox, 1972)
J. R. Cox Jr., F. M. Nolle, and R. M. Arthur, "Digital Analysis of the Electroencephalogram, the Blood Pressure Wave, and the Electrocardiogram", Proc. IEEE, vol. 60, no. 10, pp. 1137-1164, Oct., 1972.

(Cuffin, 1977)
B. N. Cuffin and D. B. Geselowitz, "Studies of the Electrocardiogram Using Realistic Cardiac and Torso Models", IEEE Trans. on Biomed. Eng., vol. BME-24, no. 3, pp. 242-252, May, 1977.

(Cuffin, 1978)
B. N. Cuffin, "Use of Electric and Magnetic Data to Determine Electric Sources in a Volume Conductor", Ann. Biomed. vol. 6, no. 3, pp. 173-193, 1978.

(Feldman, 1971)
C. L. Feldman, P. G. Amazeen, M. D. Klein, and B. Lown, "Computer Detection of Ventricular Ectopic Beats", Computers and Biomedical Research, vol. 3, pp. 666-674, 1971.

(Feldman, 1977)
C. L. Feldman and M. Hubelbank, "Cardiovascular Monitoring in the Coronary Care Unit", Med. Instrum. (USA), vol. 11, no. 5, pp. 288-292, Sept.-Oct., 1977.

(Forney, 1973)
G. D. Forney Jr., "The Viterbi Algorithm", Proc. IEEE, vol. 61, no. 3, pp. 268-278, March, 1973.

(Gersch, 1970)
W. Gersch, D. M. Eddy, and E. Dong Jr., "Cardiac Arrhythmia Classification: A Heart-Beat Interval-Markov Chain Approach", Comput. and Biomed. Res. (USA), vol. 4, pp. 385-392, 1970.

(Gersch, 1975)
W. Gersch, P. Lilly, and E. Dong Jr., "PVC Detection by the Heart-Beat Interval Data-Markov Chain Approach", Comput. and Biomed. Res. (USA), vol. 8, pp. 370-378, 1975.

(Geselowitz, 1967)
D. B. Geselowitz, "On Bioelectric Potentials in an Inhomogeneous Volume Conductor", Biophys. J., vol. 7, pp. 1-11, 1967.

(Geselowitz, 1979)
D. B. Geselowitz, "Magnetocardiography--An Overview", IEEE Trans. on Biomed. Eng., vol. BME-26, no. 9, pp. 497-504, Sept., 1979.

(Grove, 1978)
T. M. Grove, V. K. Murthy, G. A. Harvey, and L. J. Haywood, "Comparison of R-R Interval Prediction Models", Med. Instrum. (USA), vol. 12, no. 5, pp. 293-295, Sept.-Oct., 1978.

(Gustafson, 1975)
D. E. Gustafson, A. S. Willsky, and J.-Y. Wang, "Cardiac Arrhythmia Detection and Classification Through Signal Analysis", The Charles Stark Draper Lab., Inc., Cambridge, MA 02139, Technical Report R-920, July, 1975.

(Gustafson, 1977)
D. E. Gustafson, A. S. Willsky, J.-Y. Wang, M. C. Lancaster, and J. H. Triebwasser, "A Statistical Approach to Rhythm Diagnosis of Cardiograms", Proc. IEEE, vol. 65, no. 5, pp. 802-804, May, 1977.

(Gustafson, 1978a)
D. E. Gustafson, A. S. Willsky, J.-Y. Wang, M. C. Lancaster, and J. H. Triebwasser, "ECG/VCG Rhythm Diagnosis Using Statistical Signal Analysis, Part I: Identification of Persistent Rhythms", IEEE Trans. on Biomed. Eng., vol. BME-25, no. 4, pp. 344-353, July, 1978.

(Gustafson, 1978b)
D. E. Gustafson, A. S. Willsky, J.-Y. Wang, M. C. Lancaster, and J. H. Triebwasser, "ECG/VCG Rhythm Diagnosis Using Statistical Signal Analysis, Part II: Identification of Transient Rhythms", IEEE Trans. on Biomed. Eng., vol. BME-25, no. 4, pp. 353-361, July, 1978.

(Gustafson, 1978c)
D. Gustafson, J. Wang, W. Kessel, A. Akant, A. Willsky, S. Mitter, P. Doerschuk, and S. Zisk, "Investigation of Signal Analysis Techniques for ECG/VCG Classification, Vol. I: Analytical and Numerical Studies", Final Report to USAFSAM, Brooks AFB, Texas 78235 for contract F33615-77-C-0606, Scientific Systems, Inc., Cambridge, MA, Report S2I TR 78-2, Jan., 1978.

(Gustafson, 1979)
D. Gustafson, J. Wang, S. Gelfand, J. Rood, A. Willsky, and S. Mitter, "Computerized ECG Interpretation, Vol. I: Analytical and Numerical Studies", Final Report to USAFSAM, Brooks AFB, Texas 78235 for contract F33615-78-C-0615, Scientific Systems, Inc., Cambridge, MA, Report S2I Tr 79-2, April, 1979.

(Gustafson, 1980)
D. E. Gustafson, M. E. Womble, and I. C. Schick, "A P-wave Detection Method Using Statistical Signal Analysis Techniques", Proc. of the 1980 Engineering Foundation Conf., Computerized Interpretation of the Electrocardiogram, G. D. Tolan and T. A. Pryor (eds.), 27 April-2 May, 1980, Asilomar Conference Grounds, Pacific Grove, Calif. (New York: Engineering Foundation, 1980), pp. 15-22.

(Gustafson, 1981)
D. E. Gustafson, J.-Y. Wang, and A. S. Willsky, "Cardiac Rhythm Interpretation Using Statistical P and R Wave Analysis", Frontiers of Engineering in Health Care, Houston, Sept., 1981.

(Haywood, 1970)
L. J. Haywood, V. K. Murthy, G. A. Harvey, and S. Saltzberg, "On-line Real Time Computer Algorithm for Monitoring the ECG Waveform", Comput. and Biomed. Res., vol. 3, pp. 15-25, 1970.

(Haywood, 1977)
L. J. Haywood, V. K. Murthy, G. A. Harvey, and T. M. Grove, "Comparison of Models for R-R Interval Prediction in ECG Monitoring", Proc. of the First Annual Symp. of Computer Applications in Medical Care, IEEE Computer Society, Washington, D.C., Oct. 1977, pp. 281-283.

(Hosaka, 1976)
H. Hosaka, D. Cohen, B. N. Cuffin, G. Milan, and B. M. Horacek, "The Effect of the Torso Boundaries on the Magnetocardiogram", J. Electrocardiology vol. 9, no. 4, pp 418-425, 1976.

(Hristov, 1971)
H. R. Hristov, G. B. Astardjian, and C. H. Nachev, "An Algorithm for the Recognition of Heart Rate Disturbances", Med. and Biol. Eng. (G. B.), vol. 9, no. 3, pp. 221-227, May, 1971.

(Katz, 1977)
A. M. Katz, Physiology of the Heart (New York: Raven Press, 1977).

(Lainiotis, 1973)
D. G. Lainiotis and S. K. Park, "On Joint Detection, Estimation and System Identification: Discrete Data Case", Int. J. of Control, vol. 17, no. 3, pp. 609-633, 1973.

(LeBlanc, 1973)
A. R. LeBlanc and F. A. Roberge, "Present State of Arrhythmia Analysis by Computer", Canadian Medical Assoc. J., vol. 108, pp. 1239-1251, May, 1973.

(Lux, 1978)
R. L. Lux, C. R. Smith, R. F. Wyatt, and J. A. Abildskov, "Limited Lead Selection for Estimation of Body Surface Potential Maps in Electrocardiology", IEEE Trans. on Biomed. Eng., vol. BME-25, no. 3, pp. 270-276, May, 1978.

(Marcus, 1982)
J. N. Marcus, "A Matched Filter Detector of QRS Complexes in Electrocardiograms", SB Thesis, Dept. of Electrical Engineering and Computer Science, MIT, Cambridge, MA, May, 1982.

(Marriott, 1977)
H. J. L. Marriott, Practical Electrocardiology (6th edition; Baltimore: Williams and Wilkins, 1977).

(Martin, 1972)
R. O. Martin and T. C. Pilkington, "Unconstrained Inverse Electrocardiology: Epicardial Potentials", IEEE Trans. on Biomed. Eng., vol. BME-19, no. 4, pp. 276-285, July, 1972.

(Martin, 1975)
R. O. Martin, T. C. Pilkington, and M. N. Morrow, "Statistically Constrained Inverse Electrocardiology", IEEE Trans. on Biomed. Eng., vol. BME-22, no. 6, pp. 487-492, Nov., 1975.

(McFee, 1972)
R. McFee and G. M. Baule, "Research in Electrocardiology and Magnetocardiology", Proc. IEEE, vol. 60, no. 3, pp. 290-321, March, 1972.

(Miller, 1974)
W. T. Miller and D. B. Geselowitz, "Use of Electric and Magnetic Data to Obtain a Multiple Dipole Inverse Cardiac Generator--Spherical Model Study", Ann. Biomed., vol. 2, no. 4, pp. 343-360, 1974.

(Miller, 1978a)
W. T. Miller and D. B. Geselowitz, "Simulation Studies of Electrocardiogram, Part 1: Normal Heart", Circul. Res., vol. 43, no. 2, pp. 301-315, Aug., 1978.

(Miller, 1978b)
W. T. Miller and D. B. Geselowitz, "Simulation Studies of Electrocardiogram, Part 2: Ischemia and Infarction", Circul. Res., vol. 43, no. 2, pp. 315-323, Aug., 1978.

(Moe, 1966)
G. K. Moe and C. Mendez, "Simulation of Impulse Propagation in Cardiac Tissue", Ann. NY Acad., vol. 128, pp. 766-771, 1966.

(Murthy, 1979)
I. S. N. Murthy, M. R. Rangaraj, K. J. Udupa, and A. K. Goyal, "Homomorphic Analysis and Modeling of ECG Signals", IEEE Trans. on Biomed. Eng., vol. BME-26, no. 6, pp. 330-344, June, 1979.

(Oliver, 1977)
G. C. Oliver, K. L. Ripley, J. P. Miller, and T. F. Martin, "A Critical Review of Computer Arrhythmia Detection" in Computer Electrocardiography: Present Status and Criteria, L. Prody (ed.) (Mount Kisco, NY: Futura Press, 1977), pp. 319-360.

(Plonsey, 1966)
R. Plonsey, "Limitations on the Equivalent Cardiac Generator", Biophys. J., vol. 6, pp. 163-173, 1966.

(Plonsey, 1969)
R. Plonsey, Bioelectric Phenomena (New York: McGraw-Hill Book Co., 1969).

(Plonsey, 1971)
R. Plonsey, "The Biophysical Basis for Electrocardiography", CRC Critical Reviews in Bioengineering, vol. 1, no. 1, pp. 1-48, Oct., 1971.

(Plonsey, 1974)
R. Plonsey, "The Formulation of Bioelectric Source-field Relationships in Terms of Surface Discontinuities", J. of the Franklin Institute, vol. 297, no. 5, pp. 317-324, May, 1974.

(Plonsey, 1979)
R. Plonsey, "Fundamentals of Electrical Processes in the Electrophysiology of the Heart" in Advances in Cardiovascular Physics vol. 1: Theoretical Foundations of Cardiovascular Processes, D. N. Ghista, E. Van Vollenhoven, W.-J. Yang, and H. Revl (eds.) (Basel: S. Karger AG, 1979), pp. 1-28.

(Proceedings of the IEEE Computers in Cardiology Conference, 1974-1984+)
Proc. of the 1974 Computers in Cardiology Conference, K. M. Kempner (ed.), Oct. 2-4, 1974, National Institutes of Health, Bethesda, Maryland, USA (New York: IEEE, 1974).
Proc. of the 1975 Computers in Cardiology Conference, K. M. Kempner (ed.), Oct. 2-4, 1975, Thoraxcentrum, Rotterdam, The Netherlands (New York: IEEE, 1975).
Proc. of the 1976 Computers in Cardiology Conference, H. G. Ostrow and K. L. Ripley (eds.), Oct. 7-9, 1976, Washington University, St. Louis, Missouri, USA (New York: IEEE, 1976).
Proc. of the 1977 Computers in Cardiology Conference, H. G. Ostrow and K. L. Ripley (eds.), Sept. 29-Oct. 1, Thoraxcentrum, Rotterdam, The Netherlands (New York: IEEE, 1977).
Proc. of the 1978 Computers in Cardiology Conference, K. L. Ripley and H. G. Ostrow (eds.), Sept. 12-14, 1978, Stanford University, Stanford, California, USA (New York: IEEE, 1978).
Proc. of the 1979 Computers in Cardiology Conference, K. L. Ripley and H. G. Ostrow (eds.), Sept. 26-28, 1979, Cardiology Center, University of Geneva, Geneva, Switzerland (New York: IEEE, 1979).
Proc. of the 1980 Computers in Cardiology Conference, K. L. Ripley and H. G. Ostrow (eds.), Oct. 22-24, 1980, Williamsburg, Virginia, USA (New York: IEEE, 1980).
Proc. of the 1981 Computers in Cardiology Conference, K. L. Ripley (ed.), Sept. 23-25, 1981, Florence, Italy (New York: IEEE, 1982).
Proc. of the 1982 Computers in Cardiology Conference, K. L. Ripley (ed.), Oct. 12-15, 1982, Seattle, Washington, USA (New York: IEEE, 1983).

(Richardson, 1971)
J. M. Richardson, L. J. Haywood, V. K. Murthy, and R. E. Kalaba, "A Decision Theoretic Approach to the Detection of ECG Abnormalities, Part II: Approximate Treatment of the Detection of Ventricular Extrasystoles", Math. Biosci. (USA), vol. 12 nos. 1-2, pp. 97-103, Oct. 1971.

(Richardson, 1976)
J. M. Richardson, V. K. Murthy, and L. J. Haywood, "Nonlinear Markoff Models for Electrocardiographic R-R Interval Analysis", Math. Biosci. (USA), vol. 29, pp. 299-307, 1976.

(Rosenberg, 1972)
R. M. Rosenberg, C. H. Chao, and J. Abbott, "A New Mathematical Model of Electrical Cardiac Activity", Math. Biosci. (USA), vol. 14, nos. 3-4, pp. 367-394, Aug., 1972.

(Rudy, 1979)
Y. Rudy and R. Plonsey, "The Eccentric Spheres Model as the Basis for a Study of the Role of Geometry and Inhomogeneities in Electrocardiography", IEEE Trans. on Biomed. Eng., vol. BME-26, no. 7, pp. 392-299, July, 1979.

(Rush, 1971)
S. Rush, "Inhomogeneities as a Cause of Multiple Peaks of Heart Potential on the Body Surface: Theoretical Studies", IEEE Trans. on Biomed. Eng., vol. BME-18, no. 2, pp. 115-124, March, 1971.

(Schloss, 1971)
H. S. Schloss, "Computation of Solutions to the Inverse Problem of Electrocardiology", Comput. Biol. and Med. (G. B.), vol. 1, no. 3, pp. 193-198, April, 1971.

(Schluter, 1981)
P. S. Schluter, "The Design and Evaluation of a Bedside Cardiac Arrhythmia Monitor", PhD Thesis, Dept. of Electrical Engineering and Computer Science, MIT, Cambridge, MA, Sept., 1981.

(Shah, 1977)
P. M. Shah, J. M. Arnold, N. A. Haberern, D. T. Bliss, K. M. McClelland, and W. B. Clarke, "Automatic Real Time Arrhythmia Monitoring in the Intensive Coronary-Care Unit", Am. J. Card., vol. 39, no. 5, pp. 701-708, May 4, 1977.

(Smith, 1982)
J. M. Smith, "Finite Element Model of Ventricular Dysrhythmias", MS Thesis, Dept. of Electrical Engineering and Computer Science, MIT, Cambridge, MA, Aug., 1982.

(Sornmo, 1981)
L. Sornmo, P. O. Borjesson, M.-E. Nygards, and O. Pahlm, "A Method for Evaluation of QRS Shape Features Using a Mathematical Model for the ECG", IEEE Trans. on Biomed. Eng., vol. BME-28, no. 10, pp. 713-717, Oct., 1981.

(Spach, 1979)
M. S. Spach, W. T. Miller, E. Millerja, R. B. Warren, and R. C. Barr, "Extracellular Potentials Related to Intracellular Action Potentials During Impulse Conduction in Anisotropic Canine Cardiac-Muscle", Circul. Res., vol. 45, no. 2, pp. 188-204, Aug., 1979.

(Strand, 1973)
E. M Strand, "Considering the R-R Interval as a Stationary Timeseries: a Systems Approach", Proc. of the 26th Annual Conf. on Eng. in Med. and Biol., Minneapolis, Minn., USA, 30 Sept.--4 Oct. 1973 (Arlington, Va., USA: Alliance for Engng. in Medicine and Biology, 1973), p. 228.

(Thiry, 1974)
P. S. Thiry and R. M. Rosenberg, "On Electrophysiological Activity of the Normal Heart", J. of the Franklin Institute, vol. 297, no. 5, pp. 377-396, May, 1974.

(Thiry, 1975)
P. S. Thiry, R. M. Rosenbery, and J. A. Abbott, "A Mechanism for the Electrocardiogram Response to Left Ventricular Hypertrophy and Acute Ischemia", Circul. Res., vol. 36, no. 1, pp. 92-104, Jan., 1975.

(Thomas, 1979)
L. J. Thomas Jr., K. W. Clark, C. N. Mead, K. L. Ripley, B. F. Spenner, and G. C. Oliver Jr., "Automated Cardiac Dysrhythmia Analysis", Proc. IEEE, vol. 67, no. 9, pp. 1322-1337, Sept., 1979.

(Tripp, 1979)
J. H. Tripp, "Theory of the Magnetocardiogram" in Advances in Cardiovascular Physics, vol. 1: Theoretical Foundations of Cardiovascular Processes, D. N. Ghista, E. Vanvollenhoven, W. J. Yang, and H. Real (eds.) (Basel: S. Karger AG, 1979), pp. 29-46.

(Tsui, 1975)
E. T. Tsui and E. Wong, "Sequential Approach to Heart-Beat Rhythm Classification", IEEE Trans. on Info. Theory, vol. IT-21, no. 5, pp. 596-599, Sept., 1975.

(Uijen, 1979)
G. J. H. Uijen, J. P. C. de Weerd, and A. J. H. Vendrik, "Accuracy of QRS Detection in Relation to the Analysis of High Frequency Components in the Electrocardiogram", Med. and Biol. Eng. and Comput. (G. B.), vol. 17, no. 4, pp. 492-502, July, 1979.

(Vinke, 1977)
R. V. H. Vinke, A. C. Arntzenius, P. H. Huisman, H. E. Kulbertus, H. J. Ritsema van Eck, J. J. Schipperheijn, and M. L. Simoons, "Evaluation of a Computer Model of Ventricular Excitation", Proc. Computers in Cardiology 1977, pp. 605-610, 1977.

(Wang, 1976)
J.-Y. Wang, "Application of Signal Analysis Techniques to Cardiac Arrhythmia Detection and Classification", MS Thesis, Dept. of Aeronautics and Astronautics, MIT, Cambridge, MA, Jan., 1976.

(White, 1976)
C. C. White, "Note on a Markov Chain Approach to Cardiac-Arrhythmia Classification", Comput. Biomed. Res., vol. 9, no. 6, pp. 503-506, 1976.

(Wikswo, 1979)
J. P. Wikswo Jr., J. A. V. Malmivuo, W. H. Barry, M. C. Leifer, and W. M. Fairbank, "The Theory and Application of Magnetocardiography" in Advances in Cardiovascular Physics, vol. 2: Cardiograms: Theory and Application, D. N. Ghista, E. Van Vollenhoven, W.-J. Yang, and H. Reul (eds.) (Basel: S. Karger AG, 1979), pp. 1-67.

(Willsky, 1976)
A. S. Willsky and H. L. Jones, "A Generalized Likelihood Ratio Approach to the Detection and Estimation of Jumps in Linear Systems", IEEE Trans. on Automatic Control, vol. AC-21, no. 1, pp. 108-112, Feb., 1976.

(Willsky, 1982)
A. S. Willsky, "Some Solutions, Some Problems, and Some Questions", Laboratory for Information and Decision Systems, MIT, Cambridge, MA, Report LIDS-P-1168 (revised), Jan., 1982.

(Zloof, 1973)
M. Zloof, R. M. Rosenberg, and J. Abbott, "A Computer Model for Atrioventricular Blocks", Math. Biosci. (USA), vol. 18, pp. 87-117, 1973.