Technical Report No. 192
3674-16-T

THE THEORY OF SIGNAL DETECTABILITY:
EXTENSION TO THE DOUBLE COMPOSITE HYPOTHESIS SITUATION

by

Ronald L. Spooner

Approved by: T. G. Birdsall

for

COOLEY ELECTRONICS LABORATORY
Department of Electrical Engineering
The University of Michigan
Ann Arbor, Michigan

Contract No. Nonr-1224(36)
Office of Naval Research
Department of the Navy
Washington, D.C. 20360

April 1968

Reproduction in whole or in part is permitted for any purpose of the U. S. Government.
ABSTRACT

THE THEORY OF SIGNAL DETECTABILITY:
EXTENSION TO THE DOUBLE COMPOSITE HYPOTHESIS SITUATION

by Ronald Lee Spooner

The theory of signal detectability is extended to situations in which each of the two possible hypotheses are composite. The specific problem considered is that of detecting signals mixed with noise for examples in which uncertainties exist in both the signal and noise processes. The uncertain process parameters are considered to be both time invariant and time varying during the detection interval. The approach to the detection problem is Bayesian, and the initial uncertainties concerning the signal and noise processes are expressed in terms of a priori probability distributions. The time of a terminal decision is not assumed to be specified prior to the observation interval.

A general optimum processor for the doubly composite detection situation is developed. This processor is based on the overall optimal qualities of the likelihood ratio. The operation of the receiver is realized in a sequential manner for processes in which an m-th order Markov conditional dependence relationship is assumed to exist. The sequential processing technique results in a receiver design that provides practical memory requirements for arbitrary observation interval lengths and also exhibits adaptive or learning characteristics. In particular, the receiver displays classification outputs which present
probabilistic opinions of the signal and noise processes. These classification outputs are inherently linked to the detection output by a dependence relationship conditional to the true hypothesis.

Applications of the general theory to both the time invariant and time varying cases are considered. In particular, numerical results are obtained for the detection of signals known exactly and signals known except for amplitude in additive noise of unknown level. Sequential optimum receiver realizations are developed and their performance is evaluated in terms of ROC curves. In addition, the adaptive nature of the sequential receiver is demonstrated both analytically and by digital computer simulation.

For the time varying situation, the general theory is applied to develop a sequential optimum receiver for the detection of a certain signal in noise with uncertain and time varying parameters. The sequential realization results in a receiver that utilizes a practical memory size and attempts to track the varying parameters of the noise. These general results are then applied to the problem of detecting a signal in noise of varying level. The performance for this latter case is evaluated in terms of ROC curves.
FOREWORD

The research reported in this dissertation was supported in part by the Office of Naval Research, Department of Navy, under Contract No. Nonr-1224(36), "Acoustic Signal Processing."
ACKNOWLEDGMENTS

The author wishes to express his appreciation to the members of his doctoral committee, Professor T. G. Birdsall, Professor A. B. Clarke, Professor E. L. McMahon, Professor R. A. Volz, and his chairman, Professor L. F. Kazda, for their consultation and guidance during the preparation of this work. The author is especially grateful to his chairman, Professor L. F. Kazda, for his continued encouragement, guidance, and support throughout the author's university studies. In addition, the close association and direction of Professor T. G. Birdsall during the research and writing of this thesis have been invaluable.

The author also appreciates the support of the Acoustic Branch of the Office of Naval Research under whose auspices this work was performed.

Thanks are due to Mrs. Carol Wray for her dedication in the typing of both the rough and final draft copies.

Additional thanks must go to Mr. David Jaarsma for his excellent assistance in the computer programming.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>iii</td>
</tr>
<tr>
<td>FOREWORD</td>
<td>v</td>
</tr>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>vi</td>
</tr>
<tr>
<td>LIST OF ILLUSTRATIONS</td>
<td>ix</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>xv</td>
</tr>
<tr>
<td>LIST OF SYMBOLS</td>
<td>xvi</td>
</tr>
<tr>
<td>LIST OF APPENDIXES</td>
<td>xx</td>
</tr>
<tr>
<td>CHAPTER I: INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>1.1 Basic Problem</td>
<td>1</td>
</tr>
<tr>
<td>1.2 Previous Literature</td>
<td>2</td>
</tr>
<tr>
<td>1.3 Contributions of the Present Work</td>
<td>5</td>
</tr>
<tr>
<td>1.4 Method of Attack</td>
<td>6</td>
</tr>
<tr>
<td>1.5 Organization of Material</td>
<td>7</td>
</tr>
<tr>
<td>CHAPTER II: BACKGROUND</td>
<td>9</td>
</tr>
<tr>
<td>2.1 General Theory</td>
<td>9</td>
</tr>
<tr>
<td>2.2 Simple Hypothesis</td>
<td>12</td>
</tr>
<tr>
<td>2.3 Composite Signal Hypothesis</td>
<td>14</td>
</tr>
<tr>
<td>2.4 Double Composite Hypothesis</td>
<td>17</td>
</tr>
<tr>
<td>CHAPTER III: SEQUENTIAL DEVELOPMENT</td>
<td>21</td>
</tr>
<tr>
<td>3.1 Introduction</td>
<td>21</td>
</tr>
<tr>
<td>3.2 Notation</td>
<td>22</td>
</tr>
<tr>
<td>3.3 Sequential Development</td>
<td>23</td>
</tr>
<tr>
<td>3.4 Receiver Operation</td>
<td>30</td>
</tr>
<tr>
<td>CHAPTER IV: TIME-ININVARIANT UNCERTAINTIES</td>
<td>36</td>
</tr>
<tr>
<td>4.1 Introduction</td>
<td>36</td>
</tr>
<tr>
<td>4.2 Problem Statement and Notation</td>
<td>36</td>
</tr>
<tr>
<td>4.3 General Receiver Realization: Time-Invariant Case</td>
<td>38</td>
</tr>
</tbody>
</table>
## TABLE OF CONTENTS (Cont.)

4.4 Signals of Known Amplitude in Noise of Unknown Power
   4.4.1 Problem Formulation 43
   4.4.2 Receiver Design 44
   4.4.3 Receiver Operation 47
   4.4.4 Review of Receiver Evaluation Techniques 54
   4.4.5 Receiver Evaluation 67
   4.4.5.1 Effect of Noise Uncertainties 73
   4.4.5.2 Comparison with Suboptimum Receiver 102
   4.4.5.3 Sequential Operation 104

4.5 Signals of Unknown Amplitude in Noise of Unknown Level
   4.5.1 Problem Formulation 127
   4.5.2 Receiver Design 127
   4.5.3 Receiver Operation 130
   4.5.4 Receiver Evaluation 133
   4.5.4.1 Optimum Receiver Performance 134
   4.5.4.2 Comparison with Suboptimum Receiver 135
   4.5.4.3 Sequential Operation 146

CHAPTER V: TIME-VARYING UNCERTAINTIES 160
   5.1 Introduction 160
   5.2 Problem Statement and Notation 160
   5.3 General Receiver Realization 162
      5.3.1 Processing Equation 163
      5.3.2 Receiver Operation 171
   5.4 Signals of Known Amplitude in Noise of Varying Level 172
      5.4.1 Problem Formulation 172
      5.4.2 Distribution Function 174
      5.4.3 Receiver Realization 177
      5.4.4 Receiver Evaluation 186

CHAPTER VI: SUMMARY 195
   6.1 Summary and Conclusions 195
   6.2 Future Work 200

REFERENCES 218

DISTRIBUTION LIST 220
<table>
<thead>
<tr>
<th>Figure</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Illustration of the suboptimum receiver formulated by Taylor.</td>
<td>4</td>
</tr>
<tr>
<td>2.1</td>
<td>Illustration of the basic detection problem.</td>
<td>10</td>
</tr>
<tr>
<td>3.1</td>
<td>Sequential receiver realization based on Table 3.1.</td>
<td>32</td>
</tr>
<tr>
<td>4.1</td>
<td>Sequential receiver realization based on Table 4.1.</td>
<td>42</td>
</tr>
<tr>
<td>4.2</td>
<td>Graphs of Gamma probability density functions with parameters b and c.</td>
<td>48</td>
</tr>
<tr>
<td>4.3</td>
<td>Graphs of Gamma probability density functions with parameters m and v.</td>
<td>49</td>
</tr>
<tr>
<td>4.4</td>
<td>Sequential receiver for known signal unknown noise power level</td>
<td>55</td>
</tr>
<tr>
<td>4.5</td>
<td>Typical plot of Δ-circle radii versus Δ</td>
<td>59</td>
</tr>
<tr>
<td>4.6</td>
<td>Typical graph of constant log-likelihood ratio contours with parameter values $E_s = 2$, $m = 1$, $v = 1$.</td>
<td>61</td>
</tr>
<tr>
<td>4.7</td>
<td>Typical graph of constant log-likelihood ratio contours with parameter values $E_s = 2$, $m = 1$, $v = 2$.</td>
<td>62</td>
</tr>
<tr>
<td>4.8</td>
<td>Typical graph of constant log-likelihood ratio contours with parameter values $E_s = 2$, $m = 1$, $v = 3$.</td>
<td>63</td>
</tr>
<tr>
<td>4.9</td>
<td>Normal ROC curves with parameter d.</td>
<td>70</td>
</tr>
<tr>
<td>4.10</td>
<td>Two non-normal ROC curves parameterized by $d_e$.</td>
<td>71</td>
</tr>
<tr>
<td>4.11</td>
<td>Probability density functions of the likelihood ratio under SN and N</td>
<td>72</td>
</tr>
<tr>
<td>Figure</td>
<td>Title</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>----------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>4.12</td>
<td>The ROC curves for the optimum receiver, SKE-NUL case, with parameters $E_s = 1.0$, $c = 0.5$, $b = -0.5$, $0.0$, $1.0$, $2.0$, $3.0$, $4.0$.</td>
<td>84</td>
</tr>
<tr>
<td>4.13</td>
<td>The ROC curves for the optimum receiver, SKE-NUL case, with parameters $E_s = 1.0$, $c = 1.0$, $b = -0.5$, $0.0$, $1.0$, $2.0$, $4.0$.</td>
<td>85</td>
</tr>
<tr>
<td>4.14</td>
<td>The ROC curves for the optimum receiver, SKE-NUL case, with parameters $E_s = 2.0$, $c = 0.5$, $b = -0.5$, $0.0$, $0.5$, $1.0$, $1.5$, $2.0$.</td>
<td>86</td>
</tr>
<tr>
<td>4.15</td>
<td>The ROC curves for the optimum receiver, SKE-NUL case, with parameters $E_s = 2.0$, $c = 1.0$, $b = -0.5$, $0.0$, $1.0$, $2.0$, $3.0$, $4.0$</td>
<td>87</td>
</tr>
<tr>
<td>4.16</td>
<td>Detectability versus constance parameter, $c$, for SKE-NUL case with parameters $E_s = 1.0$, $b = -0.5$, $0.0$, $1.0$, $2.0$, $4.0$, $8.0$.</td>
<td>90</td>
</tr>
<tr>
<td>4.17</td>
<td>Detectability versus constance parameter, $c$, for SKE-NUL case with parameters, $E_s = 2.0$, $b = -0.5$, $0.0$, $1.0$, $2.0$, $3.0$, $4.0$.</td>
<td>91</td>
</tr>
<tr>
<td>4.18</td>
<td>The ROC curves for the optimum receiver, SKE-NUL case, with parameters $E_s = 1.0$, $v = 1.0$, $m = .707$, $1.0$.</td>
<td>93</td>
</tr>
<tr>
<td>4.19</td>
<td>The ROC curves for the optimum receiver, SKE-NUL case, with parameters $E_s = 1.0$, $v = 4.0$, $m = 1.414$, $2.0$, $4.0$.</td>
<td>94</td>
</tr>
<tr>
<td>4.20</td>
<td>The ROC curves for the optimum receiver, SKE-NUL case, with parameters $E_s = 1.0$, $v = 8.0$, $m = 2.0$, $2.83$, $4.0$, $6.0$.</td>
<td>95</td>
</tr>
<tr>
<td>4.21</td>
<td>The ROC curves for the optimum receiver, SKE-NUL case, with parameters $E_s = 2.0$, $v = 1.0$, $m = 0.707$, $1.0$, $1.414$, $2.0$.</td>
<td>96</td>
</tr>
<tr>
<td>4.22</td>
<td>The ROC curves for the optimum receiver, SKE-NUL case, with parameters $E_s = 2.0$, $v = 4.0$, $m = 1.414$, $2.0$, $2.83$, $4.0$.</td>
<td>97</td>
</tr>
<tr>
<td>4.23</td>
<td>The ROC curves for the optimum receiver, SKE-NUL case, with parameters $E_s = 2.0$, $v = 8.0$, $m = 1.0$, $1.414$, $2.0$, $2.83$, $4.0$, $5.0$.</td>
<td>98</td>
</tr>
<tr>
<td>Figure</td>
<td>Title</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>4.24</td>
<td>Detectability versus mean constance value, $m$, for SKE-NUL case with parameters, $E_S = 1.0$, $v = 0.0, 1.0, 2.0, 4.0, 8.0$.</td>
<td>99</td>
</tr>
<tr>
<td>4.25</td>
<td>Detectability versus mean constance value, $m$, for SKE-NUL case with parameters, $E_S = 2.0$, $v = 0.0, 1.0, 2.0, 4.0, 8.0$.</td>
<td>100</td>
</tr>
<tr>
<td>4.26</td>
<td>Comparison of optimum and suboptimum receiver performance, SKE-NUL case, with parameters $E_S = 2.0$, $v = 2.0$, $m = 0.5, 1.0, 2.0$.</td>
<td>105</td>
</tr>
<tr>
<td>4.27</td>
<td>Comparison of optimum and suboptimum receiver performance, SKE-NUL case, with parameters $E_S = 2.0$, $v = 4.0$, $m = 1.414, 2.0, 2.83, 4.0$.</td>
<td>106</td>
</tr>
<tr>
<td>4.28</td>
<td>Comparison of optimum and suboptimum receiver performance, SKE-NUL case, with parameters $E_S = 2.0$, $v = 8.0$, $m = 1.414, 2.0, 2.83, 4.0$.</td>
<td>107</td>
</tr>
<tr>
<td>4.29</td>
<td>Simulation runs, SKE-NUL case.</td>
<td>110</td>
</tr>
<tr>
<td>4.30</td>
<td>Simulation runs, SKE-NUL case.</td>
<td>112</td>
</tr>
<tr>
<td>4.31</td>
<td>Simulation runs, SKE-NUL case.</td>
<td>114</td>
</tr>
<tr>
<td>4.32</td>
<td>Conditional ROC curves as a function of observation time, SKE-NUL case, with parameters $m = 6$, $v = 4$.</td>
<td>117</td>
</tr>
<tr>
<td>4.33</td>
<td>Conditional ROC curves as a function of observation time, SKE-NUL case, with parameters $m = 6$, $v = 8$.</td>
<td>118</td>
</tr>
<tr>
<td>4.34</td>
<td>Conditional ROC curves as a function of observation time, SKE-NUL case, with parameters $m = 4$, $v = 4$.</td>
<td>119</td>
</tr>
<tr>
<td>4.35</td>
<td>Conditional ROC curves as a function of observation time, SKE-NUL case, with parameters $m = 4$, $v = 8$.</td>
<td>120</td>
</tr>
<tr>
<td>4.36</td>
<td>Conditional ROC curves as a function of observation time, SKE-NUL case, with parameters $m = 2.58$, $v = 4$.</td>
<td>121</td>
</tr>
<tr>
<td>Figure</td>
<td>Title</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>-----------------------------------------------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>4.37</td>
<td>Conditional ROC curves as a function of observation time, SKEA-NUL case, with parameters ( m = 2.58, v = 8 ).</td>
<td>122</td>
</tr>
<tr>
<td>4.38</td>
<td>Optimum receiver realization for signal of uncertain amplitude, noise of uncertain level.</td>
<td>129</td>
</tr>
<tr>
<td>4.39</td>
<td>Typical contours of constant log-likelihood ratio for uncertain signal amplitude and uncertain noise level with parameters ( v_\theta = 1.333, c = 1.0 ).</td>
<td>131</td>
</tr>
<tr>
<td>4.40</td>
<td>Typical contours of constant log-likelihood ratio for uncertain signal amplitude and uncertain noise level with parameters ( v_\theta = 3.0, c = 1.0 ).</td>
<td>132</td>
</tr>
<tr>
<td>4.41</td>
<td>The ROC curves for the optimum receiver, SKEA-NUL case, with parameters ( E_s = 1.0, m = 1.0, v = 1.0, v_\theta = 0.0, 1.33, 2.08, 2.80, 5.33 ).</td>
<td>136</td>
</tr>
<tr>
<td>4.42</td>
<td>The ROC curves for the optimum receiver, SKEA-NUL case, with parameters ( E_s = 1.0, m = 1.0, v = 2.0, v_\theta = 0.0, 1.33, 2.08, 2.80, 5.33 ).</td>
<td>137</td>
</tr>
<tr>
<td>4.43</td>
<td>The ROC curves for the optimum receiver, SKEA-NUL case, with parameters ( E_s = 1.0, m = 1.0, v = 4.0, v_\theta = 0.0, 1.33, 2.08, 5.33 ).</td>
<td>138</td>
</tr>
<tr>
<td>4.44</td>
<td>The ROC curves for the optimum receiver, SKEA-NUL case, with parameters ( E_s = 1.0, m = 1.0, v = 8.0, v_\theta = 0.0, 1.33, 2.80, 5.33 ).</td>
<td>139</td>
</tr>
<tr>
<td>4.45</td>
<td>Equivalent detectability vs. ( v_\theta ), SKEA-NUL case, with parameters ( E_1 = 1.0, m = 1.0, v = 0.5, 1.0, 2.0, 4.0, 8.0 ).</td>
<td>140</td>
</tr>
<tr>
<td>4.46</td>
<td>The ROC curves for the optimum receiver, SKEA-NUL case, with parameters ( E_s = 2.0, m = 1.0, v = 0.5, v_\theta = 0.0, 1.33, 2.08, 2.80, 5.33 ).</td>
<td>141</td>
</tr>
<tr>
<td>Figure</td>
<td>Title</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>-------------------------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>4.47</td>
<td>The ROC curves for the optimum receiver, SKEA-NUL case, with parameters $E_s = 2.0$, $m = 1.0$, $v = 1.0$, $v_\theta = 0.0$, $1.33$, $2.08$, $2.80$, $5.33$.</td>
<td>142</td>
</tr>
<tr>
<td>4.48</td>
<td>The ROC curves for the optimum receiver, SKEA-NUL case, with parameters $E_s = 2.0$, $m = 1.0$, $v = 2.0$, $v_\theta = 0.0$, $1.33$, $2.08$, $2.80$, $5.33$.</td>
<td>143</td>
</tr>
<tr>
<td>4.49</td>
<td>The ROC curves for the optimum receiver, SKEA-NUL case, with parameters $E_s = 2.0$, $m = 1.0$, $v = 4.0$, $v_\theta = 0.0$, $1.33$, $2.08$, $2.80$, $5.33$.</td>
<td>144</td>
</tr>
<tr>
<td>4.50</td>
<td>Equivalent detectability vs. $v_\theta$, SKEA-NUL case, with parameters $E_s = 2^2.0$, $m = 1.0$, $v = 0.5$, $1.0$, $2.0$, $4.0$.</td>
<td>145</td>
</tr>
<tr>
<td>4.51</td>
<td>Comparison of receiver performance, SKEA-NUL case, with parameters $E_s = 1.0$, $v = 1.0$, $m = 1.0$, $v_\theta = 0.0$, $1.33$, $2.80$, $5.33$.</td>
<td>146</td>
</tr>
<tr>
<td>4.52</td>
<td>Comparison of receiver performance, SKEA-NUL case, with parameters $E_s = 1.0$, $m = 1.0$, $v = 4.0$, $v_\theta = 0.0$, $1.33$, $2.80$, $5.33$.</td>
<td>147</td>
</tr>
<tr>
<td>4.53</td>
<td>Comparison of receiver performance, SKEA-NUL case, with parameters $E_s = 1.0$, $m = 1.0$, $v = 8.0$, $v_\theta = 0.0$, $1.33$, $2.80$, $5.33$.</td>
<td>148</td>
</tr>
<tr>
<td>5.54</td>
<td>Simulation runs, SKEA-NUL case.</td>
<td>150</td>
</tr>
<tr>
<td>5.55</td>
<td>Simulation runs, SKEA-NUL case.</td>
<td>152</td>
</tr>
<tr>
<td>5.56</td>
<td>Conditional ROC curves as a function of observation time, SKEA-NUL case, with parameters $m = 1$, $v = 2$, and $\theta_2 = 5.5$, $\theta_1 = 0.5$.</td>
<td>156</td>
</tr>
<tr>
<td>5.57</td>
<td>Conditional ROC curves as a function of observation time, SKEA-NUL case, with parameters $m = 1$, $v = 2$, and $\theta_2 = 5.5$, $\theta_1 = 0.1$.</td>
<td>157</td>
</tr>
<tr>
<td>4.58</td>
<td>Conditional ROC curves as a function of observation time, SKEA-NUL case, with parameters $m = 1$, $v = 4$ and $\theta_2 = 5.5$, $\theta_1 = 0.5$.</td>
<td>158</td>
</tr>
<tr>
<td>Figure</td>
<td>Title</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>----------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>4.59</td>
<td>Conditional ROC curves as a function of observation time, SKEA-NUL case, with parameters $m = 1$, $v = 4$ and $\theta_2 = 5.5$, $\theta_1 = 0.1$.</td>
<td>159</td>
</tr>
<tr>
<td>5.1</td>
<td>General sequential optimum receiver realization for slowly varying noise process.</td>
<td>170</td>
</tr>
<tr>
<td>5.2</td>
<td>Typical probability density function.</td>
<td>175</td>
</tr>
<tr>
<td>5.3</td>
<td>Typical conditional probability density function.</td>
<td>175</td>
</tr>
<tr>
<td>5.4</td>
<td>Sequential optimum receiver realization for slowly varying noise process with discrete form.</td>
<td>185</td>
</tr>
<tr>
<td>5.5</td>
<td>Optimum receiver performance, varying noise level case with parameters $\sigma^2 = 1.5$, $\mu = 6.0$.</td>
<td>190</td>
</tr>
<tr>
<td>5.6</td>
<td>Optimum receiver performance, varying noise level case with parameters $\sigma^2 = 3.0$, $\mu = 6.0$.</td>
<td>191</td>
</tr>
<tr>
<td>5.7</td>
<td>Optimum receiver performance, varying noise level case with parameters $\sigma^2 = 8.0$, $\mu = 6.0$.</td>
<td>192</td>
</tr>
<tr>
<td>5.8</td>
<td>A priori distribution functions as a function of the variance $\sigma^2$.</td>
<td>193</td>
</tr>
<tr>
<td>B.1</td>
<td>Some additional optimum receiver ROC curves for the SKE-NUL case with parameter $E_s = 1.0$.</td>
<td>208</td>
</tr>
<tr>
<td>B.2</td>
<td>Some additional optimum receiver ROC curves for the SKE-NUL case with parameter $E_s = 1.0$.</td>
<td>210</td>
</tr>
<tr>
<td>Table</td>
<td>Title</td>
<td>Page</td>
</tr>
<tr>
<td>-------</td>
<td>-----------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>3.1</td>
<td>Summary of General Sequential Equations</td>
<td>31</td>
</tr>
<tr>
<td>4.1</td>
<td>Sequential Receiver Design Equations - Time-Invariant Case</td>
<td>41</td>
</tr>
<tr>
<td>5.1</td>
<td>Sequential Receiver Design Equations - Time-Varying Case</td>
<td>169</td>
</tr>
<tr>
<td>5.2</td>
<td>Summary of Sequential Optimum Receiver Equations for Varying Noise Process</td>
<td>184</td>
</tr>
</tbody>
</table>
**LIST OF SYMBOLS**

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot;A&quot;</td>
<td>the response, &quot;signal and noise present&quot;</td>
</tr>
<tr>
<td>$A_1$</td>
<td>range of discrete probability density interval</td>
</tr>
<tr>
<td>&quot;B&quot;</td>
<td>the response, &quot;noise alone present&quot;</td>
</tr>
<tr>
<td>$B$</td>
<td>matrix $[\beta_{ij}]$</td>
</tr>
<tr>
<td>$b$</td>
<td>parameter of the a priori noise constance density</td>
</tr>
<tr>
<td>$b_k$</td>
<td>parameter of an updated noise constance density function</td>
</tr>
<tr>
<td>$c$</td>
<td>parameter of the a priori noise constance density</td>
</tr>
<tr>
<td>$c_k$</td>
<td>parameter of an updated noise constance density function</td>
</tr>
<tr>
<td>$D$</td>
<td>normalizing constant</td>
</tr>
<tr>
<td>DCH</td>
<td>Double Composite Hypothesis</td>
</tr>
<tr>
<td>$d'$</td>
<td>detectability parameter on normal ROC</td>
</tr>
<tr>
<td>$d$</td>
<td>detectability parameter on normal ROC, $d = (d')^2$</td>
</tr>
<tr>
<td>$d_e$</td>
<td>equivalent detectability parameter on non-normal ROC</td>
</tr>
<tr>
<td>$E_s$</td>
<td>total signal energy at end of observation interval</td>
</tr>
<tr>
<td>$E[\cdot]$</td>
<td>expected value of $[\cdot]$</td>
</tr>
<tr>
<td>$f(\cdot</td>
<td>\cdot)$</td>
</tr>
<tr>
<td>$g(\cdot</td>
<td>\cdot)$</td>
</tr>
<tr>
<td>$g_k$</td>
<td>parameter of an updated noise constance density function</td>
</tr>
<tr>
<td>$g_k(\cdot</td>
<td>\cdot)$</td>
</tr>
</tbody>
</table>
### LIST OF SYMBOLS (Cont.)

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>J</td>
<td>Jacobian</td>
</tr>
<tr>
<td>k</td>
<td>integer</td>
</tr>
<tr>
<td>(f(\cdot))</td>
<td>the likelihood ratio of ((\cdot))</td>
</tr>
<tr>
<td>(f(\cdot</td>
<td>\cdot))</td>
</tr>
<tr>
<td>M</td>
<td>normalizing constant</td>
</tr>
<tr>
<td>m</td>
<td>mean value of the a priori noise constance density</td>
</tr>
<tr>
<td>N</td>
<td>the hypothesis noise alone; also noise ensemble</td>
</tr>
<tr>
<td>n</td>
<td>noise vector</td>
</tr>
<tr>
<td>P(\cdot)</td>
<td>the probability of ((\cdot))</td>
</tr>
<tr>
<td>r</td>
<td>circle radius</td>
</tr>
<tr>
<td>(r_k)</td>
<td>random variable which is a function of the observations</td>
</tr>
<tr>
<td>S</td>
<td>signal ensemble</td>
</tr>
<tr>
<td>SN</td>
<td>the hypothesis signal and noise</td>
</tr>
<tr>
<td>SKE-NUL</td>
<td>signal known exactly in noise of unknown level</td>
</tr>
<tr>
<td>SKEA-NUL</td>
<td>signal known except for amplitude in noise of unknown level</td>
</tr>
<tr>
<td>s</td>
<td>signal vector</td>
</tr>
<tr>
<td>(s_i)</td>
<td>signal sample</td>
</tr>
<tr>
<td>T</td>
<td>observation interval</td>
</tr>
<tr>
<td>(t_k)</td>
<td>random variable which is a function of the observations</td>
</tr>
<tr>
<td>(u_k)</td>
<td>random variable which is a function of the observations</td>
</tr>
<tr>
<td>(v_k)</td>
<td>random variable which is a function of the observations</td>
</tr>
<tr>
<td>v</td>
<td>variance of the a priori noise constance density</td>
</tr>
<tr>
<td>(v_\theta)</td>
<td>variance of the uniform distribution of (\theta)</td>
</tr>
</tbody>
</table>
**LIST OF SYMBOLS (Cont.)**

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>( W )</td>
<td>receiver bandwidth</td>
</tr>
<tr>
<td>( X_k )</td>
<td>observation up to time ( k\tau )</td>
</tr>
<tr>
<td>( x )</td>
<td>observation vector</td>
</tr>
<tr>
<td>( x_k )</td>
<td>observation during the time ((k - 1)\tau \leq t \leq k\tau)</td>
</tr>
<tr>
<td>( y )</td>
<td>transformed observation vector</td>
</tr>
<tr>
<td>( y_c )</td>
<td>variable equal to the correlation of the observations and a known signal</td>
</tr>
<tr>
<td>( \alpha_i )</td>
<td>discrete probability density level</td>
</tr>
<tr>
<td>( \alpha_{i,k} )</td>
<td>discrete probability density level after ( k ) observations</td>
</tr>
<tr>
<td>( \alpha_k )</td>
<td>vector of ( \alpha_{i,k} )</td>
</tr>
<tr>
<td>( \beta )</td>
<td>threshold value</td>
</tr>
<tr>
<td>( \beta_{1,j} )</td>
<td>discrete conditional probability density level</td>
</tr>
<tr>
<td>( \Gamma )</td>
<td>( \gamma ) parameter space</td>
</tr>
<tr>
<td>( \Gamma(\cdot) )</td>
<td>Gamma function, ( \Gamma(\alpha) = \int_0^\infty z^{\alpha-1} \exp[-z] dz )</td>
</tr>
<tr>
<td>( \gamma )</td>
<td>noise parameter vector; also noise constance</td>
</tr>
<tr>
<td>( \gamma_k )</td>
<td>square root of the noise constance value during interval ((k-1)\tau \leq t \leq k\tau)</td>
</tr>
<tr>
<td>( \Delta )</td>
<td>( \Delta = \exp[\beta/(b + 2)] )</td>
</tr>
<tr>
<td>( \Theta )</td>
<td>( \theta ) parameter space</td>
</tr>
<tr>
<td>( \theta )</td>
<td>signal parameter vector; also signal amplitude</td>
</tr>
<tr>
<td>( \theta_1 )</td>
<td>lower end point of uniform ( \theta ) distribution</td>
</tr>
<tr>
<td>( \theta_2 )</td>
<td>upper end point of uniform ( \theta ) distribution</td>
</tr>
<tr>
<td>( \Lambda_k )</td>
<td>diagonal matrix of ( \lambda_k ) values</td>
</tr>
</tbody>
</table>

xviii
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda_1(z)$</td>
<td>$= \int_{A_1} (2\pi)^{-\frac{1}{2}} t e^{-z^2t^2} , dt$</td>
</tr>
<tr>
<td>SN</td>
<td>vector of $\lambda$ values conditional to SN</td>
</tr>
<tr>
<td>$\lambda_k$</td>
<td>mean value of the time-varying noise process</td>
</tr>
<tr>
<td>$\mu(\cdot)$</td>
<td>probability distribution of ( )</td>
</tr>
<tr>
<td>$\rho$</td>
<td>correlation coefficient of the time-varying noise process</td>
</tr>
<tr>
<td>$\sigma^2$</td>
<td>variance of the time-varying noise process</td>
</tr>
<tr>
<td>$r$</td>
<td>basic observation interval</td>
</tr>
<tr>
<td>$\Phi(z)$</td>
<td>$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp \left[-\frac{t^2}{2}\right] , dt$</td>
</tr>
<tr>
<td>$\omega_k$</td>
<td>discrete probability density levels after k observations</td>
</tr>
<tr>
<td>$\langle \cdot, \cdot \rangle$</td>
<td>inner product of two vectors</td>
</tr>
</tbody>
</table>
# LIST OF APPENDIXES

<table>
<thead>
<tr>
<th>Appendix</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Derivation of Equivalent Detectability Expression</td>
<td>202</td>
</tr>
<tr>
<td>B</td>
<td>Additional ROC Curves: Ske-nul Case</td>
<td>207</td>
</tr>
<tr>
<td>C</td>
<td>Performance of the Likelihood Correlator</td>
<td>212</td>
</tr>
<tr>
<td>D</td>
<td>Derivation of Design Equation for Skea-nul Case</td>
<td>215</td>
</tr>
</tbody>
</table>
CHAPTER I
INTRODUCTION

1.1. Basic Problem

The problem of extracting meaningful information from receptions contaminated by random interference arises in many diverse areas such as those of communication, control, psychology, and medicine. In some cases, such as in the reception of speech, the goal is to obtain the best possible reconstruction of the original transmitted waveshape from the reception. In many other cases, however, the only information needed is to decide the presence or absence of a signal in the reception and there is no particular need to reconstruct the signal. Such an example might be the determination of the presence or absence of a submarine in a specified area utilizing measurements obtained via an underwater hydrophone array. This latter type of problem is the area of concern in this work.

The detection of signals buried in noise is basically a statistical problem because of the random or statistical nature of the noise. In the early 1950's a theoretical formulation of this problem incorporating its statistical nature was developed by several authors as an application of statistical decision theory (Refs. 1 - 3). This theory of signal detection emphasized the making of the best possible decision as its primary goal and encompassed in its formulation the design and evaluation of optimum receivers for detecting the presence or absence of signals in noise.
These latter aspects of signal detection theory are the major aspects of interest in this work.

The abstract or mathematical theory of signal detection is related to the physical world through the nature of the signals to be detected and the nature of the background environment or noise in which these signals are received. The occurrence of uncertainties in either or both of these quantities affects the design of optimum receivers and their performance. In the general case, receiver design is usually simplest at either of the two extremes of uncertainty, i.e., where both the signal and noise are known exactly (the noise in a statistical sense) or where both of these quantities are completely random. Receiver performance, on the other hand, generally bears a direct relationship with certainty. The more certainty which exists concerning the nature of the signals to be detected and the noisy environment, the greater the performance the optimum receiver is able to achieve.

1.2. Previous Literature

Most of the literature which has appeared concerning the detection of signals in noise has dealt with various aspects of signal detection theory in which relative amounts of uncertainty have been assumed to exist in the signals to be detected while the noise statistics are assumed to be completely known. The detection of periodic signals of unknown amplitude and phase serves as one example (Refs. 4–6). Here the receiver is uncertain as to the exact amplitude and phase of the basic periodic waveshape whose presence or absence is to be detected in the noisy reception. Another example is that of time uncertainty. Known signal waveshapes of unknown epoch which are buried in noise are to be detected (Refs. 7, 8). Consideration of problems of this type is
quite reasonable since they represent, to varying degrees, idealized situations encountered in actual practice.

A number of authors have considered the problem of detecting signals in noise when uncertainties exist in the noise process as well as in the signal process. Taylor (Ref. 9), for example, considered the problem of detecting a known signal pulse of unknown amplitude in noise of an unknown power spectrum. The receiver design which he formulated is illustrated in Fig. 1.1. It consists of a number of parallel ideal band-pass filters equally spaced over the frequency spectrum of the input signal. The outputs of these filters are then processed by ideal limiters whose subsequent outputs are correlated with the known input signal waveshape, weighted by gain factors $W_i$, and summed. This final summation is then used as the decision variable. In Taylor's work no attempt was made to achieve optimum detection, only to optimize the parameter values in the basic design presented in Fig. 1.1.

The work reported by Thomas and Williams (Ref. 10) concerned the detection of signals in noise with time-varying uncertainties. Here again, no attempt was made to obtain optimum receiver realizations, and only suboptimum receivers were considered. In their particular work, two suboptimum detection systems were analyzed and the performances compared in an effort to determine the effect of the time-varying uncertainties on the relative performance.

The foundation of the work presented here is found in the theory of signal detectability as presented by Peterson, Birdsall and Fox (Ref. 2). In this theory the concept of likelihood ratio plays a central role, and it has been shown that for a wide range of criteria in which
Fig. 1.1. Illustration of the sub-optimum receiver formulated by Taylor
correct decisions are considered "good" and incorrect decisions are considered "bad" decisions based on the likelihood ratio yield optimum performance. For this reason, the likelihood ratio also plays a dominant role in this present work. An additional basis for the work in this thesis is to be found in the work by Nolte (Ref. 11) concerning the design of optimum sequential receivers. In this latter effort the likelihood ratio concept is also of central importance.

1.3. Contributions of the Present Work

In this work, the various aspects of the theory of signal detection are extended to situations in which uncertainties exist not only in the signals to be detected but also in the background environment in which these signals appear. From a practical standpoint, this type of situation is encountered quite frequently. In passive detection of underwater sound (sonar), for example, the time varying nature of the ocean as the noisy background medium implies that even exact knowledge of the noise statistics during one observation interval can at best yield only an uncertain opinion of these statistics during a later observation interval. In addition, measurements taken in an effort to estimate these statistics may not be entirely useful since a signal free measurement direction may not be available, thus leading to uncertainty as to whether the measurements consist of signal and noise or noise alone. Another example is found in stationary detection situations in which even though the basic statistical nature of the noise is known, various parameters of this random process, such as noise power density, may not be known exactly. Both of these types of uncertainties are considered in this work.

The consideration of uncertainties in both the noise and signal processes within the framework of the theory of signal detectability leads
to the formulation of a Double Composite Hypothesis (DCH) detection problem as will be seen later. As mentioned above, the practical aspect of this type of formulation is that a more realistic modeling of many physical situations can be achieved. To the author's knowledge, optimum solutions of this type of problem have not been considered before, except in pure mathematical generalities.

The general solution of the Double Composite Hypothesis case and its application to specific examples result in optimum receiver design and evaluation which enable conclusions to be reached concerning the effect of signal and noise uncertainties on the detectability of signals in noise. For example, from the consideration of a known signal to be detected in Gaussian noise of unknown noise power density, one concludes that for equal false alarm and miss probabilities, the form of the optimum receiver is independent of the a priori opinion of the noise parameter. Other conclusions are given as they occur in the text.

Consideration of the theory of signal detectability has previously been confined almost completely to the single composite hypothesis situation wherein uncertainties exist only in the signal hypothesis and not in the noise hypothesis. The basic contributions of the present work then are the extension of this theory to a consideration of sequential operating procedures that exhibit adaptive characteristics for the Double Composite Hypothesis situation, and the application of the theory to specific situations in an effort to determine the effect of composite uncertainties on the detectability of signals in noise.

1.4. Method of Attack

In this work, the concept of receiver design for the Double Composite Hypothesis case is approached from the viewpoint profitably
used by Nolte and extended to this case. This approach seeks optimum receiver realizations which operate in a sequential fashion, and consequently, exhibit adaptive or learning features. One of the beneficial characteristics of this type of receiver is the availability of the learned output for use in parameter classification and/or estimation in addition to the classical detection output. In addition, for the case of time-varying uncertainties where no terminal observation time is specified, the use of the sequential realization often results in a fixed-memory device; whereas a receiver designed in classical manner in general requires a time-growing memory.

The approach used throughout the work adheres to the Bayesian philosophy. In other words, the assumption is made that consistent a priori opinions concerning the nature of the signals and noise involved in the detection problem are held by the receiver designer and that, furthermore, these opinions can be expressed in terms of probability distributions. The modification of these opinions, in the light of evidence, gained through observations or otherwise, is made according to Baye's Rule. A more comprehensive discussion of the Bayesian approach may be found in the report by Breipolil and Koschmann (Ref. 12) and the paper by Edwards, Lindman, and Savage (Ref. 13).

1.5. Organization of Material

In Chapter II the basic background material is presented. In addition, at the end of Chapter II, the Double Composite Hypothesis detection problem is formulated. In Chapter III a sequential realization of the likelihood ratio is derived and its implications in terms of adaptive optimum receiver realizations are presented and discussed. In Chapter IV a nontime-varying situation and specific applications are presented in
an effort to consider the effect of this type of uncertainty on the detection
of signals in noise. Chapter V considers a time-varying situation in which
a higher degree of uncertainty exists.
CHAPTER II

BACKGROUND

2.1. General Theory

In the previous chapter, it was indicated that the concept of signal detection as formulated by Peterson, Birdsall and Fox (Ref. 2) forms the foundation for this work. It is appropriate, therefore, that this theory, which is now termed classical fixed-time detection theory, be reviewed. The work of Nolte is an extension of this basic theory to detection situations in which the response time is unspecified and is considered in context.

The basic problem of signal detection is best considered by reference to Fig. 2.1 below. The receiver is presented with an observation \( x(t) \) during a time interval \( t_0 \leq t \leq T + t_0 \). This observation may consist of signal and noise or just noise alone. On the basis of this observation the receiver must make the decision as to which of these two conditions was present. The two conditions are mutually exclusive since it is assumed that the signal is either present during the entire interval or absent during the entire interval. From a decision theory viewpoint this problem may be viewed as requiring the receiver to decide which one of two mutually exclusive hypotheses, signal and noise, or noise alone, occurred during the observation interval. In this present work, it is assumed that the signal is added to the noise so that mathematically the detection problem may be represented as

\[
\text{hypothesis SN: } x(t) = s(t) + n(t)
\]
hypothesis N: \( x(t) = n(t) \)

Thus the receiver must decide which one of the mutually exclusive hypotheses, SN or N, held during the observation interval.

![Diagram of the basic detection problem](image)

Fig. 2.1. Illustration of the basic detection problem

The classical solution of this problem involves the specification of receiver design, the realization of receiver design, and the evaluation of receiver performance. For the situation described above, a fixed-time solution is desired. That is, at the time \( T + t_0 \) the receiver is required to make a binary decision as to which hypothesis is present. This is in contrast to an unspecified response time detection situation where the time at which the receiver is to make a decision remains arbitrary.

In order to consider the specification of receiver designs which are optimum in some sense, the concept of what is meant by optimum must be included in the theory. Since the signal detection problem consists of two mutually exclusive hypotheses, either SN or N is present during the observation interval, the choosing of a particular hypothesis by the receiver results in two possible kinds of decisions; correct decisions and incorrect decisions. When the actual input consists of signal and noise, these decisions
are termed a detection and a miss respectively. If we define "A" to mean the response, "signal is present," and "B" to mean the response, "signal is absent," then a detection results when the actual input is signal and noise and the receiver correctly responds "A"; correspondingly, a miss results when the receiver responds "B." Similarly, when the actual input consists of noise alone, there are correct and incorrect decisions. The situation of responding "A" when noise alone is present is generally termed a false alarm. The specification of optimum receiver designs is based on these four possible condition-response pairs. Relative values and costs may be associated with these pairs and for a particular problem, a risk criterion or performance criterion is specified in terms of them and the a priori probabilities. The optimum receiver is then determined with respect to the particular criterion which has been chosen. Two examples are the Weighted Combination Criterion and the Neyman-Person Criterion. In the former, the objective is to make decisions so that the difference between the probability of detection and a weighted version of the probability of false alarm is maximized. In other words, the objective is to maximize

\[ P(\text{"A"}|\text{SN}) - w P(\text{"A"}|\text{N}) \]

where

\[ P(\text{"A"}|\text{SN}) \] is the probability of detection

\[ P(\text{"A"}|\text{N}) \] is the probability of false alarm

and

\[ w \] is a positive constant weighting factor expressing the relative cost of false alarms in relation to the value of detections.
For the Neyman-Person criterion the objective is to make decisions so that the false alarm probability is maintained below a preset value and at the same time the detection probability is maximized. Hence, the objective is to maximize $P(\text{"A"} \mid \text{SN})$ with the constraint

$$P(\text{"A"} \mid \text{N}) \leq m$$

It should be kept in mind that the detection problem is basically a statistical one and, therefore, the results are statistical in nature. In other words, the optimum performance is an optimum expected or average performance with respect to the various values, costs, and possible decisions.

In the early 1960's Birdsall (Ref. 14) utilized the methods of statistical decision theory to generalize the proof that optimum receiver design should be specified in terms of the likelihood ratio of the input observation. In this work it was shown that for a wide class of performance criteria, including those described above, receivers which make decisions on the basis of the likelihood ratio of the observation result in optimum performance. In fact, the class of performance criteria for which the likelihood ratio is optimum may generally be thought of as that class for which correct responses are considered "good" and incorrect responses considered "bad" as expressed in terms of the various values, costs, and possible decisions discussed above.

2.2. Simple Hypothesis

The likelihood ratio of the input observation, $x(t)$, is defined as the ratio of the probability density function of the observation $x(t)$ under the condition that signal and noise is present, to the
probability density function of \( x(t) \) provided that noise alone is present. This function of the input observation \( x(t) \) may be expressed as

\[
\ell(x) = \frac{f(x|SN)}{f(x|N)}
\]

where \( \ell(x) \) is the likelihood ratio for \( x(t) \), and \( f(x|SN) \) and \( f(x|N) \) are the two conditional density functions conditioned on the occurrence of the two possible hypotheses, \( SN \) or \( N \), respectively. The input observation \( x(t) \) is a random process, and as such, may be represented by numerous mathematical methods. If, for example, this input is timelimited and (Fourier Series) bandlimited, then by the sampling theorem it may be characterized by \( 2WT \) time samples and represented by the vector \( (x_1, x_2, \ldots x_{2WT}) \) where \( W \) is the bandwidth over which the input observation is defined and \( T \) is the total duration of the observation. Under this condition, the likelihood ratio for the fixed-time detection situation is given by

\[
\ell(x_1, x_2, \ldots, x_{2WT}) = \frac{f(x_1, x_2, \ldots, x_{2WT}|SN)}{f(x_1, x_2, \ldots, x_{2WT}|N)}
\]

where \( f(x_1, x_2, \ldots, x_{2WT}|SN) \) is the joint density of the \( 2WT \) observation samples under the condition that signal and noise are present and \( f(x_1, x_2, \ldots, x_{2WT}|N) \) is the joint density of the \( 2WT \) observation samples under the condition that noise alone is present. An example of a detection problem of this type is the case of detecting a known signal in additive white Gaussian noise. The formal solution of this problem results in a logarithm of the likelihood ratio of the form

\[1\]

\[1\] The logarithm is a monotone function of the likelihood ratio and therefore also provides optimum detection.
\[ \ell_n \left( x_1, x_2, \ldots, x_{2^{WT}} \right) = \sum_{i=1}^{2^{WT}} x_i s_i - \frac{s_i^2}{2} \]

where the known signal is represented by the signal sample vector, \((s_1, s_2, \ldots, s_{2^{WT}})\). From the mathematical form of this result it is evident that the receiver may be implemented as the familiar cross-correlator.

In this simple example the realization of the receiver as a cross-correlator followed directly from the form of the likelihood ratio. In the general situation, however, the realization of receiver design from the likelihood ratio is, at best, an art, since many different realizations are generally possible. The final design chosen, therefore, usually reflects both the ingenuity and experience of the designer and other additional engineering considerations such as complexity, cost, etc. However, if this final design is a realization of the likelihood ratio, or any monotonic function thereof, it will yield optimum performance.

The performance of receivers, both optimum and suboptimum, is evaluated in terms of their receiver operating characteristic (ROC) curves. An individual ROC curve is a plot of the probability of detection versus the probability of false alarm for the given receiver as a function of the threshold settings. The use of ROC curves provides a measure of the performance of the optimum receiver, and in addition, allows the comparison of suboptimum receivers with this optimum performance. The many additional properties of ROC curves are well discussed in the work by Birdsall (Ref. 15).

2.3. Composite Signal Hypothesis

In the work discussed above, it was considered that both the signal and noise statistics were known exactly at the receiver and the
likelihood ratio provided the means for determining in an optimum manner which of the two hypotheses SN or N was present during the observation. This type of problem in which both processes are completely known is termed a simple hypothesis problem (Ref. 5). Signal detection has also been extended by numerous people to situations in which one or more parameters of the signal process are unknown to the receiver. Under this condition, the hypothesis of signal and noise is composite and the general problem is termed a composite signal hypothesis problem as contrasted with the case where both signal and noise are known exactly. Examples of this type of problem which have appeared in the literature are numerous (Refs. 16 - 19). The detection of signal known except for amplitude, known except for carrier phase, and known except for epoch are among the many cases which have been considered. A more complex example of a composite signal hypothesis problem is given in the work of Nolte (Ref. 1) where signals composed of known components but of unknown time of occurrence are encountered.

The difference between the simple hypothesis problem and that of the composite signal hypothesis problem is that for the former a single known signal or an ensemble of signals with no parameter of interest is present at the receiver under the hypothesis SN while for the latter an ensemble of possible signals is present under this hypothesis. The presence of signal uncertainty in the composite detection problem implies that the possibility of "learning" the signal parameters exists as observations are taken. In addition, from a practical viewpoint, the composite nature assumed for the SN hypothesis allows the modeling of a larger class of physical problems, the design of optimum receivers again involving the realization of the likelihood ratio.
For the case of a composite signal hypothesis, the conditional likelihood ratio is averaged over each possible signal of the ensemble, i.e.,

$$\ell(x) = \int_S \ell(x|s) \, d\mu(s) \quad (2.1)$$

where $\ell(x|s)$ is the likelihood of the observation $x$ conditional to exact knowledge of the signal, and $\mu(s)$ is the a priori probability distribution defined over all possible signals in the ensemble. The averaging of a conditional likelihood ratio for a composite signal hypothesis situation follows directly from consideration of the total likelihood ratio

$$\ell(x) = \frac{f(x|SN)}{f(x|N)} \quad (2.2)$$

From the definition of the composite signal hypothesis, the numerator of this expression can be written as

$$f(x|SN) = \int_S f(x|s, SN) \, d\mu(s) \quad (2.3)$$

where $f(x|s, SN)$ is the probability density of the observation $x$ conditional to the hypothesis SN, and exact knowledge of the signal. If Eq. 2.2 is now substituted into Eq. 2.3 the result is

$$\ell(x) = \frac{1}{f(x|N)} \int_S f(x|s, SN) \, d\mu(s) \quad (2.4)$$

The denominator appearing on the right hand side of this equation is independent of the integration, and therefore can be taken inside the integral, yielding

$$\ell(x) = \int_S \left[ \frac{f(x|s, SN)}{f(x|N)} \right] \, d\mu(s) \quad (2.5)$$

It is evident that the term in brackets in the above expression is the
likelihood ratio of the observation conditional to exact knowledge of the signal. Thus if we define

$$\ell(x|s) = \frac{f(x|s, SN)}{f(x|N)} \quad (2.6)$$

then by substituting this expression in Eq. 2.5 the desired result is obtained.

$$\ell(x) = \int_{S} \ell(x|s) \, d\mu(s) \quad (2.7)$$

The implication of this result is that for a composite signal hypothesis problem the likelihood ratio receiver may be realized as a receiver which performs an averaging process over the ensemble of possible signals with respect to a conditional likelihood ratio.

2.4. Double Composite Hypothesis

In the composite signal hypothesis problem uncertainties exist in the signals to be detected, and the resulting optimum receiver designs are based on a conditional likelihood ratio averaged over the signal ensemble. If uncertainties also exist in the noise, then both the SN and N hypotheses are composite. Even for the situation in which the signal is known exactly, the uncertainties concerning the background noise still imply that each of the hypotheses SN and N is composite since noise is present during the occurrence of either hypothesis. For this case in which both of the hypotheses are composite the detection problem is termed a Double Composite Hypothesis problem. This type of problem is the concern of the present work.

For the Double Composite Hypothesis situation, optimum receiver design is still based on the likelihood ratio; however, in this
case the likelihood ratio receiver does not assume an "averaged" form as for the composite signal hypothesis problem discussed above. Instead it generally consists of two channels, one for each hypothesis, which perform computations conditional to these hypotheses and are then combined to form the total likelihood ratio. This result is easily seen from the form of the total likelihood ratio for this situation.

Starting with the likelihood ratio

$$f(x) = \frac{f(x|SN)}{f(x|N)} \quad (2.8)$$

one may use the definition of the Double Composite Hypothesis problem to express the denominator as

$$f(x|N) = \int_N f(x|n, N) \, d\mu(n) \quad (2.9)$$

In this latter expression $f(x|n, N)$ is the probability density of the observation $x$ conditional to the hypothesis $N$ and exact knowledge of the noise process and the quantity $\mu(n)$ is the a priori probability distribution defined over the noise process.

In a similar manner, the numerator of Eq. 2.8 may be expressed as

$$f(x|SN) = \int_S \int_N f(x|n, s, SN) \, d\mu(s, n) \quad (2.10)$$

where $f(x|n, s, SN)$ is the probability density of the observation conditional to the hypothesis $SN$ and exact knowledge of the signal and noise processes. The joint probability distribution $\mu(s, n)$ is the a priori distribution of the signal and noise processes (which in general may not be independent). It is evident that even if the signal is known exactly,
f(x|SN) must be averaged over the ensemble of possible noise processes since the noise is present under either hypothesis.

Utilizing Eqs. 2.9 and 2.10, the total likelihood ratio of Eq. 2.8 may be expressed as

\[
\ell(x) = \frac{\int_{S} \int_{N} f(x|n, s, SN) \, d\mu(s, n)}{\int_{N} f(x|n, N) \, d\mu(n)} \tag{2.11}
\]

In this latter expression for the total likelihood ratio, the occurrence of an averaging process under both of the possible hypotheses implies that the receiver is no longer concerned with performing an averaging process on a conditional likelihood ratio as was the case in the composite signal hypothesis situation. Instead, the receiver is usually constrained to operate in a dual channel mode, each channel operating on a separate hypothesis. This feature will become apparent in the work which follows.

The above discussion indicates that the theory of signal detection, whose central concept is the likelihood ratio, inherently allows the inclusion of a priori knowledge or opinions in its formulation. For any given situation the existence of a priori opinions by the designer is difficult to doubt and this knowledge is fully utilized by the theory. In addition, the inclusion of the a priori knowledge allows the consideration of optimum detection for a wide range of physical problems.

It was mentioned previously that the existence of uncertainties in the signal and noise processes involved in the detection problem implies the possibility of a learning feature in the optimum receiver. The learning or adaptive feature is fully exploited by the use of optimum
receivers which operate in a time-sequential manner as will be shown later on. It should be noted however that this learning process is of a conditional nature. In other words, the learning output is conditional to knowledge of the true hypothesis (N or SN) and therefore is utilized in conjunction with the detection output of the receiver. An additional point concerning the adaptive or learning feature of sequential receivers is the consideration of memory requirements. Classical detection theory is a full memory theory; that is, it is assumed that the optimum receiver is capable of storing all the information required in the computation of the likelihood ratio. In many unspecified response time detection problems a classical realization of the optimum receiver results in a memory requirement which grows with increasing time. Through the use of sequential receivers, which exhibit the adaptive or learning feature, this requirement for an unlimited memory may be eliminated, thus satisfying a very practical requirement for receiver realization.
CHAPTER III

SEQUENTIAL DEVELOPMENT

3.1. Introduction

In this chapter a sequential realization of the likelihood ratio is developed for the Double Composite Hypothesis detection problem. The results of this development serve as the basis for the design of a detection processor which operates sequentially in time. This receiver is optimum since a receiver which realizes the likelihood ratio is an optimum device, and therefore, a receiver which realizes the same likelihood ratio in a sequential manner is also an optimum device.

However, in addition to being an optimum detector, the sequential receiver also exhibits some additional characteristics which are quite beneficial. One of these characteristics is the "adaptive" nature of the sequential receiver. Without pursuing a rigorous definition of adaptivity (a term or concept which generally means somewhat different things to different people) it will be shown that the sequential receiver exhibits one of the most commonly referred to characteristics of an adaptive device, that of learning or updating. Thus, in addition to providing a detection output, the sequential receiver for the Double Composite Hypothesis situation also contains a learning section which may be used for classification and/or estimation of initially uncertain parameters.

Another useful characteristic of the sequential receiver is its ability to operate as a fixed memory device for many situations in which a classical detector would exhibit a time-growing memory. This
type of situation arises in detection problems where the response time of the receiver is not specified a priori. In other words, the receiver is allowed to operate over long averaging times and although it must always be capable of making a decision, no time is established a priori as to when this decision must be made. Additional discussion of sequential receiver designs and their relation to memory size may be found in the report by Nolte (Ref. 11) in which sequential likelihood receivers were developed for the composite signal hypothesis case.

3.2. Notation

In order to develop the sequential realization of the likelihood ratio for the Double Composite Hypothesis case the following notation will be adopted. The receiver operates in a sequential fashion; in every $\tau$ seconds the observations in the past $\tau$ seconds are processed. The quantity $X_k$ is used to denote $x(t)$ for the interval $0 \leq t \leq k\tau$, and $x_k$ denotes $x(t)$ for $(k-1)\tau \leq t \leq k\tau$. In other words, a capital letter indicates the observation from the beginning of time to "now" (i.e., $k\tau$) and the lower case letter refers to the present observation which is to be processed as a unit. Sequentially (in time) the receiver updates its decision output which is used to decide whether the signal was or was not present during the entire observation, $X_k$. In addition, the receiver also updates a set of parameter probability density functions conditioned on the presence or absence of signal in the total observation, $X_k$.

The design of the sequential receiver is based on the likelihood ratio, and therefore, it is inherently linked to the classical fixed time detection theory. This point is brought out by Nolte, and perhaps should be reiterated here. The use of a classical receiver in an unspecified response
time case requires the receiver to compute the likelihood ratio of the total observation from the time zero to the present time. As observation time is increased this receiver is then required to compute a new likelihood ratio based on the observations from time zero up to this new final time. In the case of composite hypothesis problems where the observations are not generally independent this new likelihood ratio is not obtainable from the previous likelihood ratio. The result therefore is a receiver which computes a series of likelihood ratios, each of which is valid for an observation beginning at zero with only one of them being utilized for making the decision at the time $t_k$. In general, operation of this type requires a memory storage which must increase with time.

The sequential receiver on the other hand maintains a stored likelihood ratio and a set of probability densities which reflect the state of the system, in other words, which retain all past information relevant to the decision. These quantities are updated on the basis of new observations and in this manner provide an updated likelihood ratio which contains all information relevant to a decision at any particular time.

3.3 Sequential Development

In this section the sequential form of the likelihood ratio is derived. The concept of conditional mth order dependence (sometimes called the mth order Markov) is utilized in this derivation, and therefore, it should be explained at the outset. Let us consider the conditional density of the total observation, $X_k$, conditioned on exact knowledge of the signal and noise processes and either hypothesis SN or N. These densities are expressed symbolically in Eq. 3.1

$$f(X_k | s, n, \frac{SN}{N})$$

(3.1)
Using the laws of conditional probability densities Eq. 3.1 may be rewritten in the following form:

\[
\left( X_k | s, n, \frac{SN}{N} \right) = f\left( x_k | X_{k-1}, s, n, \frac{SN}{N} \right) f\left( X_{k-1} | s, n, \frac{SN}{N} \right)
\]  

(3.2)

where \( x_k \) represents the last small observation interval.

We then say that the observations \( X_k \) are conditionally mth order dependent if the two quantities represented by \( f\left( x_k | X_k, s, n, \frac{SN}{N} \right) \) in Eq. 3.2 can be expressed as

\[
f\left( x_k | X_{k-1}, s, n, \frac{SN}{N} \right) = f\left( x_k | x_{k-1}, \ldots, x_{k-m}, s, n, \frac{SN}{N} \right)
\]  

(3.3)

where \( x_{k-1}, \ldots, x_{k-m} \) denote the observations taken during the intervals, \( (t_{k-m-1}, t_{k-1}) \). For example, in the case where \( m = 0 \), the observation \( x_k \) is conditionally independent of all past observations.

Turning now to the sequential derivation, consider the likelihood ratio, \( \ell(X_k) \), which by definition is given by

\[
\ell(X_k) = \frac{f(X_k | SN)}{f(X_k | N)}
\]  

(3.4)

The numerator and denominator of this expression for the likelihood ratio may be expressed in the following expanded forms, respectively, by invoking the definitions of conditional probabilities.

\[
f(X_k | SN) = f(x_k | X_{k-1}, SN) f(X_{k-1} | SN)
\]  

(3.5)

\[
f(X_k | N) = f(x_k | X_{k-1}, N) f(X_{k-1} | N)
\]  

(3.6)
If we now substitute the expressions of Eqs. 3.5 and 3.6 into Eq. 3.4, the likelihood ratio becomes

$$f(X_k) = \frac{\frac{f(x_k | X_{k-1}, SN)}{f(x_k | X_{k-1}, N)}}{\frac{f(x_{k-1} | SN)}{f(x_{k-1} | N)}} \quad (3.7)$$

The form of this latter expression lends itself to a natural definition of the likelihood ratio of the observation \(x_k\) conditioned on all the past observations \(X_{k-1}\). Thus, we define the conditional likelihood ratio of the observation \(x_k\) as

$$f(x_k | X_{k-1}) = \frac{f(x_k | X_{k-1}, SN)}{f(x_k | X_{k-1}, N)} \quad (3.8)$$

and using this definition the likelihood ratio of the total observation \(X_k\) becomes

$$f(X_k) = f(x_k | X_{k-1}) \cdot f(X_{k-1}) \quad (3.9)$$

At this point note that we have partially fulfilled our ultimate goal of obtaining a sequential realization of the likelihood ratio for the total observation \(X_k\). As Eq. 3.9 indicates, the total likelihood ratio has been expressed as a product of likelihood ratios, one which contains the decision information prior to the observation \(x_k\), and the other which depends upon \(x_k\), directly. It is now this latter term with which we wish to work.

From Eq. 3.8 we see that \(f(x_k | X_{k-1})\) is expressed as a ratio of two conditional probability densities. In what follows, only the numerator of this expression will be used, with similar results following for the denominator.

By rearranging Eq. 3.5 it follows that the numerator of Eq. 3.8 may be expressed as

25
\[ f(x_k | x_{k-1}, SN) = \frac{f(x_k | SN)}{f(x_{k-1} | SN)} \quad (3.10) \]

For a Double Composite Hypothesis problem the numerator of this latter expression is a composite probability density function which is the result of averaging over the signal and noise processes. Thus, \( f(x_k | SN) \) may be expressed as

\[ f(x_k | SN) = \int_S \int_N f(x_k | s, n, SN) \, g(s, n | SN) \, ds \, dn \quad (3.11) \]

where \( f(x_k | s, n, SN) \) is the density of the observation \( X_k \) conditioned on exact knowledge of the signal and noise processes, and \( g(s, n | SN) \) is the joint probability density for the signal and noise processes. The total expression is the average of the conditional density of \( X_k \) over all the possible signal and noise processes which could occur. Note that we have expressed the general situation in which the noise and the signal ensemble processes may not be independent. If we now assume that conditional \( m \)-th order dependence holds, then the conditional density for \( X_k \), \( f(x_k | s, n, SN) \), which appears in the integrand of Eq. 3.11, may be expressed as

\[ f(x_k | s, n, SN) = f(x_k | x_{k-1}, \ldots, x_{k-m}, s, n, SN) \, f(x_{k-1} | s, n, SN) \quad (3.12) \]

If we now substitute Eq. 3.12 into 3.11, we have

\[ f(x_k | SN) = \int_S \int_N f(x_k | x_{k-1}, \ldots, x_{k-m}, s, n, SN) \, f(x_{k-1} | s, n, SN) g(s, n | SN) \, ds \, dn \quad (3.13) \]

If this latter expression is now substituted into Eq. 3.10, the following result is obtained for the conditional density of \( x_k \) conditioned on the
hypothesis SN,

\[
f(x_k | X_{k-1}, \text{SN}) = \frac{1}{f(X_{k-1} | \text{SN})} \int \int_{SN} f(x_k | x_{k-1}, \cdots, x_{k-m}, s, n, \text{SN}) \\
\cdot f(X_{k-1} | s, n, \text{SN}) g(s, n | \text{SN}) \, ds \, dn
\]  

(3.14)

The quantity \( f(X_{k-1} | \text{SN}) \) appearing in Eq. 3.14 is independent of the variables of integration and, therefore, may be taken inside the integration signs. Thus Eq. 3.14 becomes

\[
f(x_k | X_{k-1}, \text{SN}) = \int \int_{SN} f(x_k | x_{k-1}, \cdots, x_{k-m}, s, n, \text{SN}) \\
\cdot \frac{f(X_{k-1} | s, n, \text{SN})}{f(X_{k-1} | \text{SN})} g(s, n | \text{SN}) \, ds \, dn
\]  

(3.15)

A natural definition for an updated probability density is apparent from the form of Eq. 3.15. This density, which might have been obtained by a pure Bayes' type manipulation, is given by

\[
g_{k-1}(s, n | \text{SN}) = \frac{f(X_{k-1} | s, n, \text{SN})}{f(X_{k-1} | \text{SN})} g(s, n | \text{SN})
\]  

(3.16)

and represents the updated or a posteriori joint probability density of the signal and noise processes based on the observations up to the time \( t_{k-1} \). If Eq. 3.16 is substituted into Eq. 2.15, the expression for \( f(x_k | S_{k-1}, \text{SN}) \) reduces to the following form:

\[
f(x_k | X_{k-1}, \text{SN}) = \int \int_{SN} f(x_k | x_{k-1}, \cdots, x_{k-m}, s, n, \text{SN}) g_{k-1}(s, n | \text{SN}) \, ds \, dn
\]  

(3.17)

This latter equation is of the desired form since it expresses the conditional density of the observation \( x_k \) (under the hypothesis SN) in terms of an
updated joint probability density. If this expression is now divided by the similar expression which can be obtained for the conditional density of $x_k$ under the noise alone hypothesis, $N$, then the conditional likelihood ratio of the observation $x_k$ can be expressed in terms of updated joint probability densities for $s$ and $n$. As the expression stands, however, the updated joint probability density for $g_{k-1}(s, n | SN)$ is given in terms of the a priori joint probability density for $s$ and $n$ conditioned on the hypothesis $SN$. We now wish to express this updated joint probability density in terms of a previously updated joint probability density so that an updating process which proceeds sequentially in time can be determined. Consider Eq. 3.16 with subscript $k$ replaced by $k + 1$. The result is

$$g_k(s, n | SN) = \frac{f(X_k | s, n, SN)}{f(X_k | SN)} g(s, n | SN) \quad (3.18)$$

where the operation of increasing the subscript is valid since it merely represents an expression of Bayes' law after $X_k$ observations instead of $X_{k-1}$ observations.

If we now solve Eq. 3.16 for the density function $g(s, n | SN)$, and substitute the result into Eq. 3.18, then, upon rearranging, we obtain the following expression for the conditional density $g_k(s, n | SN)$:

$$g_k(s, n | SN) = \frac{f(X_{k-1} | SN)}{f(X_k | SN)} \left[ \frac{f(X_k | s, n, SN)}{f(X_{k-1} | s, n, SN)} \right] g_{k-1}(s, n | SN) \quad (3.19)$$

To reduce this latter expression to a more useful form we first write the conditional density $f(X_{k-1} | s, n, SN)$ in terms of the expanded form given by Eq. 3.12 and then substitute it into Eq. 3.19. In addition, we also express the density $f(X_k | SN)$ in the expanded form given below (which
follows directly from the laws of conditional probabilities) and substitute it into Eq. 3.19.

\[ f(X_k \mid SN) = f(x_k \mid X_{k-1} \mid SN) f(X_{k-1} \mid SN) \quad (3.20) \]

Upon performing these two substitutions Eq. 3.19 becomes

\[
\begin{align*}
g_k(s, n \mid SN) &= \frac{f(X_{k-1} \mid SN)}{f(x_k \mid X_{k-1}, SN) f(X_{k-1} \mid SN)} \\
& \quad \cdot \frac{f(x_k \mid x_{k-1}', \ldots, x_{k-m}', s, n, SN) f(X_{k-1} \mid s, n, SN)}{f(X_{k-1} \mid s, n, SN)} \\
&= g_{k-1}(s, n \mid SN)
\end{align*}
\]

which reduces to the following form after cancellation of the common factors:

\[ g_k(s, n \mid SN) = \frac{f(x_k \mid x_{k-1}', \ldots, x_{k-m}', s, n, SN)}{f(x_k \mid X_{k-1}', SN)} g_{k-1}(s, n \mid SN) \quad (3.22) \]

This latter expression is of the form which we desire. It expresses the joint probability density for \( s \) and \( n \) at the time \( k\tau \) in terms of the last observation taken, \( x_k \), the joint probability density which held at the time \((k - 1)\tau \) and the previous \( m \) observations.

In the work above only the SN hypothesis has been considered. However, the same set of operations could have been performed under the noise alone hypothesis, \( N \). The equations which result from this derivation are similar to the equations obtained above. The combination of these two sets of equations then represents the design equations for an optimum sequential receiver for the Double Composite
Hypothesis case. Before discussing the behavior of this sequential receiver, the pertinent operational equations are summarized in Table 3.1.

It is important to note that a sequential detector whose design is based on the physical manifestation of these equations is truly an optimum device since the equations themselves are merely a manipulation of the likelihood ratio of the total observation \( X_k \) and therefore guarantee it to be an optimum processor.

A block diagram of a sequential receiver based on an implementation of the equations given in Table 3.1 is given in Fig. 3.1. The form of this receiver involves two distinct channels, one relevant to the hypothesis SN and the other relevant to the hypothesis N. This dual channel form is a direct manifestation of the fact that we are dealing with Double Composite Hypothesis problems in which the likelihood ratio cannot be expressed directly as an average likelihood ratio. This point was discussed in Chapter II. In addition, it should be noted that the receiver design presented in Fig. 3.1. represents only one particular realization of the equations given in Table 3.1. In general, other designs based on the same equations are possible, however, we have chosen to present the one shown for illustration.

### 3.4. Receiver Operation

To discuss the operation of the receiver presented in Fig. 3.1 we consider that the \( x_k \) observation has just been taken. The basic operation of the receiver then proceeds as follows: Upon reception of the observation \( x_k \) the receiver computes the quantities \( f(x_k | X_{k-1}^{S}, SN) \) and \( f(x_k | X_{k-1}^{N}, N) \) by utilizing the stored probability densities \( g_{k-1}^{S}(s,n|SN) \) and \( g_{k-1}^{N}(n|N) \) and the stored values of the previous \( m \) observations.
\begin{align}
\ell(X_k) &= \ell(x_k | X_{k-1}) \ell(X_{k-1}) \tag{3.9} \\
\ell(x_k | X_{k-1}) &= \frac{f(x_k | X_{k-1}, SN)}{f(x_k | X_{k-1}, N)} \tag{3.8} \\
\int_{S} \int_{N} f(x_k | x_{k-1}', \ldots, x_{k-m}', s, n, SN) g_{k-1}(s, n | SN) ds \, dn \tag{3.17} \\
g_k(s, n | SN) &= \frac{f(x_k | x_{k-1}', \ldots, x_{k-m}', s, n, SN)}{f(x_k | X_{k-1}', SN)} \, g_{k-1}(s, n | SN) \tag{3.22} \\
f(x_k | X_{k-1}) &= \int_{N} f(x_k | x_{k-1}', \ldots, x_{k-m}', N) \, g_{k-1}(n | N) \, dn \tag{3.23} \\
g_k(n | N) &= \frac{f(x_k | x_{k-1}', \ldots, x_{k-m}', n, N)}{f(x_k | X_{k-1}', N)} \, g_{k-1}(n | N) \tag{3.24} \\
\end{align}

Table 3.1 Summary of general sequential equations

31
Fig. 3.1. Sequential receiver realization based on Table 3.1
The conditional likelihood ratio of the observation $x_k$ is then formed by computing the ratio of these two quantities. This conditional likelihood ratio of the observation $x_k$ is then multiplied by the stored value of the previous likelihood ratio $\ell(X_{k-1})$ to yield the new likelihood ratio $\ell(X_k)$ for the time $k$.

In addition, the receiver also computes the updated joint probability densities $g_k(s, n | SN)$ and $g_k(n | N)$ by utilizing the incoming observation $x_k$ and the stored values of the past $m$ observations. Note that if the problem at hand can be described by conditional independence (zero order dependence) $m$ observations need not be stored and the updating process depends only on the incoming observation $x_k$. These updated probability densities are then stored to be used for computing the likelihood ratio of the forthcoming observation. In this manner, then, the receiver operates in a sequential fashion accepting observations; computing likelihood ratios based on these observations and previously stored quantities; and then updating the stored quantities for use in computing the next likelihood ratio.

The adaptive nature of the device is easily seen from the block diagram. As observations are taken the receiver continually "adapts its opinion" concerning the joint probability density of $s$ and $n$; in other words, it attempts to "learn" about these processes. It should be noted that the learning process is a conditional one; in other words, the learned outputs are conditioned on whichever hypothesis the receiver user believes to be true. As indicated in the block diagram of the receiver, each channel computes an updated joint probability density based upon the observations taken which represents its "opinion" concerning the nature of the signal and noise processes.
It is evident that these learned opinions stored in the receiver memory and continually updated may be utilized for classification or estimation of the signal and noise processes.

Although a sequential realization of the likelihood ratio for the Double Composite Hypothesis situation has been developed here and a block diagram of a receiver implementing these equations has been presented, it must be noted that the actual implementation of a physical receiver with a fixed memory size may not always be possible. Two cases where this type of implementation is possible for simple hypothesis situations have been covered by Fralick (Ref. 20) and Roberts (Ref. 21). The major factor governing the ability to reduce an unlimited memory problem to one which can be solved with a fixed memory utilizing a sequential receiver involves the existence of a priori densities which are of a special class. The densities of this class are termed "reproducing densities," and their relation to memory size is an inherent aspect of all problems involving learning and classification. Reproducing densities and their relationship to fixed memory devices have been investigated by many people. For example, comprehensive discussion of this problem may be found in the work of Grettenberg (Ref. 22) in which necessary and sufficient conditions are given for a class of learning machines to be realizable using a fixed memory size independent of the number of observations.

The idea of reproducing density is basically quite simple. In general, a class of a priori densities is called reproducing with respect to an observation if an a priori density which is a member of the class causes the resulting a posteriori density function to be a member
of the class of density functions. This is most important when the mem-
bers of the class of functions are specified by a fixed finite number of
parameters. One often writes these densities in a specific "functional
form" that specifies their evaluation. Thus, if for a given
observation class, the a priori density is reproducing, then the functional
forms of the a posteriori densities obtained as additional observations are
taken are all the same (perhaps with different parameters) as the
original density function. The existence of this reproducing property
in turn implies that the associated sequential processor is realizable
with a fixed memory size, only changing parameter values as observa-
tions are taken.

One additional point should be mentioned concerning
sequential operation. It was discussed previously that in any fixed
time problem a sequential design is not necessary for the optimum
processor. However, from the form of the sequential detector
developed in this section it is apparent that the use of this type of
receiver, even in a fixed-time situation, might yield advantages; namely,
that of providing an output which may be utilized for classification or
estimation of the signal and noise processes. This classification or
estimation is conditional, of course, on whichever hypothesis, signal
and noise or noise alone, was present during the observation interval.
CHAPTER IV

TIME-INVARIANT UNCERTAINTIES

4.1. Introduction

In this chapter the problem of the detection of a signal with a finite number of unknown parameters embedded in noise with uncertain parameters is considered in detail. The uncertainties which exist in the signal and noise processes are assumed to be time-invariant during the observation interval and are expressed as random parameters of these processes. The existence of these uncertainties in both the signal and noise processes implies a detection problem of the Double Composite Hypothesis type. In fact, the existence of uncertainties in only the noise process is sufficient to imply a detection problem of this type.

The study of this class of problems is motivated from numerous practical situations. For example, in many practical communication systems the signals to be detected are approximately known in waveshape but unknown in amplitude because of signal fading within the transmission channel. The noise process on the other hand may often be modeled as statistically known except for some unknown parameter value such as noise power level.

In the work following the general class of problems is first formulated utilizing the results of Chapter III. This effort is then followed by specific numerical applications.

4.2. Problem Statement and Notation

The existence of uncertainties in both the signal and noise processes
which are time-invariant during the observation interval of interest is a reasonable approximation to many physical situations. In addition, consideration of problems of this type leads directly to the consideration of Double Composite Hypothesis detection problems as noted above. In this section, a general class of such problems is formulated and the notation in the solution is developed.

The observations which are presented to the receiver are assumed to consist of either a signal added to noise or of noise alone. These receiver inputs are assumed to be defined for the entire observation interval, \( 0 \leq t \leq T \), and furthermore, are assumed to be (Fourier Series) limited to a band of frequencies of width \( W \). By the sampling theorem, these inputs may be represented as points in a \( 2WT \) dimensional space: the coordinates of a point being the value of the input function at the sample points \( t_i = i/2WT \), for \( 1 \leq i \leq 2WT \). With this representation the notation \( X_k \) denotes a receiver input, \((x_1, x_2, \ldots, x_k)\), where \( x_j \) denotes the \( j \)th sample value or coordinate.

The signal and noise processes are assumed to be known to the receiver except for a finite number of parameters appearing in each process. In the case of the signal process, these parameters are denoted by the vector \( \theta \) and in the case of the noise process by the vector \( \gamma \). These unknown parameters are characterized by their joint a priori distribution functions \( g(\theta) \) and \( g(\gamma) \) respectively, and are considered constant throughout the observation interval. In other words, during any particular observation interval the values of the random parameter vectors \( \theta \) and \( \gamma \) are time-invariant. In addition, \( \theta \) and \( \gamma \) are considered to be initially independent so that their joint a priori density function may be expressed as the product of their individual a priori density functions.
The final assumption is that the conditional joint density of the total observation $X_k$ (that is, the density of $X_k$ conditional to knowledge of the values of the random parameter vectors, $\theta$ and $\gamma$) satisfies an mth order Markov condition under both hypotheses SN and N. This condition was discussed in Chapter III. From this assumption of an mth order Markov condition it follows that

$$f(X_k | \theta, \gamma, SN) = f(x_k | x_{k-1}, \ldots, x_{k-m}, \theta, \gamma, SN) f(X_{k-1} | \theta, \gamma, SN)$$

(4.1)

under the hypothesis SN, and that

$$f(X_k | \gamma, N) = f(x_k | x_{k-1}, \ldots, x_{k-m}, \gamma, N) f(X_{k-1} | \gamma, N)$$

(4.2)

under the hypothesis N.

The formulation and notation for the class of time-invariant detection problems considered in this chapter is now complete. In the next section the results of Chapter III are utilized to derive a sequential realization of the optimum receiver for this class of problems.

4.3. General Receiver Realization: Time-Invariant Case

The realization of an optimum receiver for a given detection situation is based on the likelihood ratio of the received observation. In the case of general Double Composite Hypothesis detection problems it was shown in Chapter III that the likelihood ratio could be realized in a sequential manner. In this section, these results are applied to the class of detection problems formulated above, and an optimum receiver which operates in a sequential manner is realized. In the following sections more specific numerical applications are considered.
To assist in the development of the receiver realization the pertinent results of Chapter III, as summarized in Table 3.1, are partially repeated below. For simplicity only the equations conditioned on the hypothesis $\text{SN}$ are presented with the equations conditioned on the hypothesis $\text{N}$ following in a corresponding manner.

\begin{equation}
\ell(X_k) = \ell(x_k | X_{k-1}) \ell(X_{k-1}) \tag{3.9}
\end{equation}

\begin{equation}
\ell(x_k | X_{k-1}) = \frac{f(x_k | X_{k-1}, \text{SN})}{f(x_k | X_{k-1}, \text{N})} \tag{3.8}
\end{equation}

\begin{equation}
f(x_k | X_{k-1}, \text{SN}) = \int_{S} \int_{N} f(x_k | x_{k-1}, \ldots, x_{k-m}, s, n, \text{SN}) g_{k-1}(s, n | \text{SN}) \, dn \, ds \tag{3.17}
\end{equation}

\begin{equation}
g_k(s, n | \text{SN}) = \frac{f(x_k | x_{k-1}, \ldots, x_{k-m}, s, n, \text{SN})}{f(x_k | X_{k-1}, \text{SN})} \quad g_{k-1}(s, n | \text{SN}) \tag{3.22}
\end{equation}

In the problem formulation above the signal process was assumed to contain a finite number of unknown parameters denoted by the vector $\theta$. This implies that the averaging with respect to the entire signal ensemble appearing in Eq. 3.17 is equivalent to an average over the space of all $\theta$ values. If we denote this space of $\theta$ values by $\Theta$, then the signal vector $s$ appearing in Eqs. 3.17 and 3.16 may be replaced by the signal process parameter, $\Theta$. The average over the space $S$ is then replaced by an average over the space $\Theta$. With regard to the noise process, a corresponding statement can be made. The average over the noise ensemble in the equations above is replaced by an average
over the space of $\gamma$ vectors. We will denote this latter space by $\Gamma$.

These conclusions are incorporated into Eqs. 4.3 and 4.4.

\[
f(x_k | X_{k-1}, SN) = \int_{\theta} \int_{\Gamma} f(x_k | x_{k-1}', \cdots, x_{k-m}', \theta, \gamma, SN) g_{k-1}(\theta, \gamma | SN) d\theta d\gamma
\]

\[
g_{k}(\theta, \gamma | SN) = \frac{f(x_k | x_{k-1}', \cdots, x_{k-m}', \theta, \gamma, SN)}{f(x_k | X_{k-1}', SN)} g_{k-1}(\theta, \gamma | SN)
\]

These last two equations, together with the corresponding equations conditioned on the hypothesis $N$, represent the basic equations utilized in the optimum receiver design for the time-invariant case.

They are summarized in Table 4.1 and a block diagram of the receiver realization is illustrated in Fig. 4.1. The requirement of a dual channel receiver for Double Composite Hypothesis detection problems is evident in this block diagram.

The operation of the sequential receiver obtained in this section is entirely analogous to that of the general receiver fully discussed in Chapter III. Therefore, its operation is only briefly discussed here. With reference to Fig. 4.1 it is evident that upon reception of the $k$th observation sample the receiver computes the quantities $f(x_k | X_{k-1}', SN)$ and $f(x_k | X_{k-1}', N)$, utilizing the stored distributions of the parameter vectors $\theta$ and $\gamma$ and Eqs. 4.7 and 4.9. These quantities are then used to obtain the conditional likelihood ratio which in turn is used to obtain the current or updated likelihood ratio. At the same time, the $k$th observation sample is used to adapt or update the distributions of the signal and noise parameter vectors, $\theta$ and $\gamma$ (conditioned on each hypothesis). These new or updated density functions of the parameters are obtained by
\[ \ell(X_k) = \ell(x_k | X_{k-1}) \ell(X_{k-1}) \] (4.5)

\[ \ell(x_k | X_{k-1}) = \frac{f(x_k | X_{k-1}, SN)}{f(x_k | X_{k-1}, N)} \] (4.6)

\[ f(x_k | X_{k-1}, SN) = \int_\Theta \int_\Gamma f(x_k | x_{k-1}, \ldots, x_{k-m}, \theta, \gamma, SN) g_{k-1}(\theta, \gamma | SN) d\theta d\gamma \] (4.7)

\[ g_k(\theta, \gamma | SN) = \left[ \frac{f(x_k | x_{k-1}, \ldots, x_{k-m}, \theta, \gamma, SN)}{f(x_k | X_{k-1}, SN)} \right] g_{k-1}(\theta, \gamma | SN) \] (4.8)

\[ f(x_k | X_{k-1}, N) = \int_\Gamma f(x_k | x_{k-1}, \ldots, x_{k-m}, \gamma, N) g_{k-1}(\gamma | N) d\gamma \] (4.9)

\[ g_k(\gamma | N) = \left[ \frac{f(x_k | x_{k-1}, \ldots, x_{k-m}, \gamma, N)}{f(x_k | X_{k-1}, N)} \right] g_{k-1}(\gamma | N) \] (4.10)

Table 4.1. Sequential receiver design equations - time invariant case

Eqs. 4.8 and 4.10 and represent an explicit expression for the learning feature of the sequential receiver. It should be noted that even though the parameters \( \theta \) and \( \gamma \) are considered a priori independent, the updated joint density for these parameters, \( g_k(\theta, \gamma | SN) \), may not reflect this independence. In other words, for the general case this updated joint density function is not expressible as the product of the individual updated density
Fig. 4.1. Sequential receiver realization based on Table 4.1
functions, the result for any particular situation being dependent upon the form of the conditional density function.

The realization of the sequential receiver illustrated in Fig. 4.1 as a device with a fixed memory is dependent on the form of the a priori densities which describe the signal and noise process parameter vectors, \( \theta \) and \( \gamma \). As discussed in Chapter III, the desirability of a fixed memory size implies that these a priori densities be of a reproducible nature with respect to the observations. To be more specific, it is required that the functional form of the density \( g_K(\theta, \gamma|\text{SN}) \) be the same as the functional form of the density \( g_{k-1}(\theta, \gamma|\text{SN}) \), and similarly, for the updated density functions conditional to the hypothesis \( N \). If this latter condition is satisfied, then a fixed memory receiver is assured.

In this section, a sequential realization of an optimum receiver has been developed for Double Composite Hypothesis detection problems with time-invariant uncertainties. The nature of the problem led to a dual channel receiver in which each channel operates in a mode conditioned on one of the two possible hypotheses which might be present. In addition, the sequential realization of the receiver resulted in a processor which provides a learning or classification output. In the next sections more specific numerical applications are considered.

4.4. Signals of Known Amplitude in Noise of Unknown Power

In order to obtain specific results concerning the effect of Double Composite Hypothesis situations on our ability to detect signals in noise, more detailed examples than that presented in the previous sections must be considered. It is the application of the general theory to specific problems that is of interest in the design and evaluation of
detection receivers. The application of this theory, of course, implies computations utilizing specific mathematical functions and numerical constants. In particular, specific mathematical forms for the general density functions utilized in the previous section must be considered. The basic detection concepts do not depend on the particular form of these distributions; however, the detailed mathematics and operational principles of the solution do depend on these distributions.

4.4.1. Problem Formulation. In this section the results of the previous work are applied to the specific example of detection of a signal known exactly added to noise of unknown level (SKE-NUL). The noise is assumed to be white Gaussian noise of unknown but time-invariant noise power. In the mathematical operations which follow the reciprocal of the noise power, in other words the constance, is used and is denoted by the symbol $\gamma$. This notation is consistent with that previously used since it is assumed that the constance is the only random parameter existing in the noise process. The functional form of the a priori distribution chosen for $\gamma$ contains two parameters, and by proper choice of these parameters a wide range of uncertainty may be modeled. The form of the distribution was chosen (1) to obtain reproducibility with respect to the conditionally normal observations encountered, and (2) to provide a sufficient amount of mathematical tractability.

The functional form of the constance distribution is defined by the density function given in Eq. 4.11. This distribution is called the Gamma distribution.

$$g(\gamma) = D\gamma^b \exp \left[-c\gamma\right] \quad \text{for } \gamma \geq 0, \quad b > -1, \quad c > 0$$

$$g(\gamma) = 0 \quad \text{for } \gamma < 0$$

(4.11)
Since \( g(\gamma) \) is a density function, the normalizing constant \( D \) is determined so that the integral of the function over the entire range of \( \gamma \) is equal to one. Performing this computation results in a value for \( D \) given by

\[
D = c^{b+1} / \Gamma(b+1)
\]

where \( \Gamma(b+1) \) denotes the gamma function. This function is defined by

\[
\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} \exp[-x] \, dx
\]

The mean and variance of the Gamma distribution may be expressed in terms of the parameters \( b \) and \( c \). These quantities are easily computed and are given by the following equations. The symbols \( m \) and \( v \) denote the mean and variance respectively.

\[
m = \frac{b+1}{c}
\]

\[
v = \frac{b+1}{c^2}
\]

From these latter two equations, it is apparent that the mean and variance are each dependent on both of the distribution parameters, \( b \) and \( c \). For this reason, it is sometimes more convenient (and will be useful later) to express the density function directly in terms of its mean and variance. To do this, Eqs. 4.14 and 4.15 are first solved simultaneously to yield \( b \) and \( c \) as functions of \( m \) and \( v \). The results of these computations are

\[
b = \frac{m^2}{v} - 1
\]

and

45
\[ c = \frac{m}{v} \]  \hspace{1cm} (4.17)

If the values for \( b \) and \( c \) as functions of \( m \) and \( v \) are substituted into Eq. 4.11, the constance distribution parameterized by its mean and variance is obtained. This form of the distribution function is given in Eq. 4.18 below.

\[ g(\gamma) = M_\gamma \left( \frac{m^2}{v} - 1 \right) \exp \left[ -(m/v)\gamma \right], \gamma \geq 0 \]
\[ g(\gamma) = 0 \hspace{1cm} \gamma < 0 \]
\[ (4.18) \]

The normalizing constant \( M \) appearing in this latter equation is determined by substitution in Eq. 4.12 or by normalizing the integral of the density function over the range of \( \gamma \). Using either method, the result is

\[ M = \left( \frac{m}{v} \right)^{m^2/v} \left[ \Gamma \left( \frac{m^2}{v} \right) \right]^{-1} \]  \hspace{1cm} (4.19)

Graphs of the constance density function in terms of the parameters \( b \) and \( c \) are presented in Fig. 4.2. Graphs of this same density function parameterized by its mean and variance, \( m \) and \( v \), are also presented and appear in Fig. 4.3. Investigating these figures we note that as \( \gamma \to \infty \), \( g(\gamma) \to 0 \). In addition, for \( b > 0 \) \((m^2/v > 1)\) the function is zero at \( \gamma = 0 \) and has a maximum at \( \gamma = b/c \) \([\gamma = (m^2 - v)/m]\). This maximum value is given by \( D(b/c)^b \exp(-b) \) or in terms of the mean and variance by \( M[(m^2 - v)/m]^{(m^2-v)/v} \exp[-(m^2-v)/v] \).

From Figure 4.2 the effects of changes in the parameter \( c \) are evident. An
increase in the value of \( c \) causes the density to peak more predominantly at lower values of \( \gamma \). This result is expected from Eqs. 4.14 and 4.15 which indicate that an increase in the value of \( c \) reduces both the variance and the mean of the distribution. Similarly, increases in the value of \( b \) tend to broaden the density and increase the point at which the maximum value occurs. An increase in the mean value \( m \) causes the maximum value of the density function to move away from zero while an increase in the variance \( \nu \) causes the density function to broaden out. These latter features are incorporated into Fig. 4.3.

4.4.2. Receiver Design. Optimum receiver design for the Double Composite, time-invariant application presented in this section is based on the general theory developed in the previous section and summarized by the equations presented in Table 4.1. The likelihood ratio plays the dominant role in these equations.

In order to use the results of Table 4.1 the form of the general density functions appearing in Eqs. 4.5 through 4.10 must be determined. At present, the signal is known exactly and, therefore, the dependence on the signal parameter \( \theta \) which appears in these equations may immediately be suppressed. This leaves only the noise process parameter \( \gamma \). In addition, the consideration of a white Gaussian noise process (which we assume to have zero mean value) implies that the \( k \)th observation, \( x_k \), is normally distributed and conditionally independent under the condition that the constance value \( \gamma \) is known. Hence, under noise alone, the density of the \( k \)th observation conditioned on knowledge of the constance value is

\[
f(x_k | \gamma, N) = (\gamma / 2\pi)^{\frac{1}{2}} \exp[-x_k^2 \gamma / 2] \tag{4.20}
\]
Fig. 4.2. Graphs of gamma probability density functions with parameters $b$ and $c$
Fig. 4.3. Graphs of gamma probability density functions with parameters $m$ and $v$. 

$m = 1.414$, $v = 1.0$

$m = 2.0$, $m = 2.83$, $v = 1.0$

$m = 2.0$, $v = 0.5$

$v = 1.0$, $v = 2.0$
The effect of adding a known signal is to shift the mean value of the above density function. Therefore, when both signal and noise are present at the receiver input the conditional density of the kth observation is given by

\[ f(x_k | \gamma, SN) = \frac{1}{(2\pi)^{k/2}} \exp\left[-\frac{(x_k - s_k)^2}{\gamma/2}\right] \quad (4.21) \]

The process of updating the constance density function conditioned on the hypotheses SN or N is expressed by Eqs. 4.8 and 4.10 in Table 4.1. It is this updating process which allows us to break the total observation into several smaller cascaded observations and then calculate the likelihood ratio of each smaller observation in sequence. This procedure is summarized by the equations of Table 4.1. The practical implementation of this concept is contingent upon satisfying the closure property which then permits the updating process to be carried out by an adequate memory receiver. The closure of the constance density function under conditionally normal observations is demonstrated below.

Considering the conditional updating process after one observation as expressed by Eq. 4.10 with \( k = 1 \), we obtain the following density conditioned on the noise alone hypothesis.

\[
g_1(\gamma | N) = \frac{\gamma^{\frac{1}{2}} \exp\left(-x_1^2\gamma/2\right)}{\int_0^{\infty} \gamma^{\frac{1}{2} + b} \exp\left(-c\gamma - x_1^2\gamma/2\right) d\gamma} \quad \gamma \geq 0
\]

\[ = 0 \quad \text{otherwise} \quad (4.22) \]

Performing the integration indicated in the denominator of this latter expression and collecting terms we have the following form for Eq. 4.22.
\[ g_1(\gamma | N) = \frac{(c + x_1^2/2)^{b+3/2}}{\Gamma(b + 3/2)} \gamma^{b + 1/2} \exp \left[ -\left(c + x_1^2/2\right)\gamma \right] \]

\[ \gamma \geq 0 \]

= 0 otherwise

(4.23)

This operation demonstrates the closure of the a priori density under the conditionally normal observations which we are considering. The observation maps \( b \rightarrow b + 1/2 \) and \( c \rightarrow c + x_1^2/2 \). Thus, there is one deterministic parameter \( b \) and one random parameter \( c \). These results are easily generalized to obtain the conditional density after \( k \) observations. Performing the generalization we obtain the following form for the conditional density of \( \gamma \) after \( k \) observations

\[ g_k(\gamma | N) = \frac{(c_k)^{b_k+1}}{\Gamma(b_k + 1)} \gamma^{b_k} \exp (-c_k \gamma) \]

\[ \gamma \geq 0 \]

= 0 otherwise

(4.24)

where

\[ c_k = c + 1/2 \sum_{i=1}^{k} x_i^2 \]

(4.25)

\[ b_k = b + k/2 \]

(4.26)

The parameters \( c \) and \( b \) (without subscripts) have already been defined as the parameters of the distribution in the absence of observations.

The effect of the closure property is evident from the form of Eq. 4.24. As additional observations are taken, the mathematical
form of the conditional density function remains fixed (i.e., as that of a Gamma density) with only the parameters describing it changing. This result implies a receiver memory requirement which remains fixed in time and need only be of sufficient magnitude to maintain these various parameter values. In addition, the required parameter values are easily obtained in a sequential manner since from Eqs. 4.25 and 4.26 it follows that

$$c_k = c_{k-1} + x_k^2/2 \quad (4.27)$$

and

$$b_k = b_{k-1} + 1/2 \quad (4.28)$$

The updated density conditioned on SN follows directly from Eq. 4.8 of Table 4.1. It is obtained in a manner similar to the work above, with the closure property again being evident. Performing the operations we obtain for the conditional density after k observations

$$g_k(\gamma | SN) = \frac{(g_k)^{b_k+1}}{\Gamma(b_k + 1)} \gamma^{b_k} \exp(-g_k \gamma)$$

$$\gamma \geq 0$$

$$= 0 \quad \text{otherwise} \quad (4.29)$$

where

$$g_k = g_{k-1} + (x_1 - s_1)^2/2 \quad (4.30)$$

$$b_k = b_{k-1} + 1/2 \quad (4.31)$$

To obtain the sequential form of the likelihood processor we use Table 4.1 and the updating equations given by Eqs. 4.24 and
4.29. Rewriting Eq. 4.9 from Table 4.1 we have

\[ f(x_k | X_{k-1}, N) = \int_{\Gamma} f(x, | \gamma, N) \cdot g_{k-1}(\gamma | N) \, d\gamma \]  \hspace{1cm} (4.32)

The quantities on the right-hand side of this equation have all been previously calculated. Substituting these quantities into Eq. 4.32 we obtain

\[ f(x_k | X_{k-1}, N) = \int_{0}^{\infty} \left( \frac{\gamma}{2\pi} \right)^{-\frac{1}{2}} \exp \left( -x_k^2 \frac{\gamma}{2} \right) \cdot \left( \frac{c_{k-1}^{b_k-1+1}}{\Gamma(b_k-1+1)} \right) \cdot (\gamma)^{b_k-1} \exp(-c_{k-1}\gamma) \, d\gamma \]  \hspace{1cm} (4.33)

Performing the integration indicated in Eq. 4.33 we obtain the following form for \( f(x_k | X_{k-1}, N) \).

\[ f(x_k | X_{k-1}, N) = (2\pi)^{-\frac{1}{2}} \left( \frac{c_{k-1}^{b_k-1+1}}{\Gamma(b_k-1+1)} \right) \cdot \frac{\Gamma(b_k-1+3/2)}{[c_{k-1} + x_k^2/2]^{b_k-1+3/2}} \]  \hspace{1cm} (4.34)

The conditional density of the kth observation conditioned on all past observations and the hypothesis SN, \( f(x_k | X_{k-1}, SN) \), follows in a corresponding manner. Performing the calculation yields

\[ f(x_k | X_{k-1}, SN) = (2\pi)^{-\frac{1}{2}} \left( \frac{g_{k-1}^{b_k-1+1}}{\Gamma(b_k-1+1)} \right) \cdot \frac{\Gamma(b_k-1+3/2)}{[g_{k-1} + (x_k - s_k)^2/2]^{b_k-1+3/2}} \]  \hspace{1cm} (4.35)

As Eq. 4.6 indicates, the conditional likelihood ratio \( f(x_k | X_{k-1}) \) is obtained as a ratio of the conditional densities appearing
in Eqs. 4.33 and 4.35. Thus, using these equations and Eq. 4.6 we obtain the following form for the conditional likelihood ratio.

\[
\ell(x_k \mid X_{k-1}) = \left[ \frac{c_k - 1}{g_k - 1} \right]^{b_{k-1} + \frac{3}{2}} \left[ \frac{x_k / 2 + c_k - 1}{(x_k - s_k)^2 / 2 + g_k - 1} \right]^{b_{k-1}}
\]

(4.36)

The likelihood ratio of the total observation at time \(k\) is then obtained as the product of this conditional likelihood ratio (Eq. 4.36) with the likelihood ratio of the total observation up to time \(k-1\), as expressed in Eq. 4.5.

The sequential design of the likelihood processor is obtained from Eqs. 4.36 and 4.35. One particular design is shown in block diagram form in Fig. 4.4. For convenience, the logarithm of the likelihood ratio is presented as the decision variable in this receiver. As previously mentioned, however, this is equivalent to presenting the likelihood ratio since the logarithm is a monotonic function of its argument. The sequential operation of this receiver is evidenced by the feedback loops appearing in the figure.

4.4.3. Receiver Operation. In Fig. 4.4 the effect of the uncertainty in the noise process requiring essentially a dual section receiver is evident. Each section of the receiver operates on one of the possible hypotheses and their outputs are combined to form the decision variable. It was previously noted that this is a general characteristic of Double Composite Detection problems. The learning or classification feature of the receiver is inherently linked to the sequential calculation of the parameters \(b_k\), \(c_k\) and \(g_k\). Since the mathematical forms of the conditional density functions do not change with time, their
Fig. 4.4. Sequential receiver for known signal unknown noise power level
values after \( k \) observations are immediately determined by the \( k \)th values of these parameters.

Another important feature of the receiver operation is apparent when the loci of constant likelihood ratio are investigated. Any ultimate decision by the receiver is based on a comparison of the likelihood (or log-likelihood) ratio with a threshold value. Therefore, these loci determine, as a function of the threshold setting, the boundary in observation space between possible decisions. To aid in an investigation of the decision boundaries the contours of constant log-likelihood ratio are considered after two observations although the sequential operation of the receiver implies its capability to provide a detection decision after any number of observations. The results obtained for two observations are then generalized to an arbitrary number of observations.

The logarithm of the likelihood ratio after two observations is easily obtained from Eq. 4.36 and the logarithm of Eq. 4.5 given below.

\[
\ln f(x_1, x_2) = \ln [f(x_2/x_1)f(x_1)] \quad (4.37)
\]

Performing the calculations, the following result is obtained for the logarithm of the likelihood ratio after two observations.

\[
\ln f(x_1, x_2) = (b + 2) \ln \left[ \frac{x_1^2 + x_2^2 + 2c}{(x_1 - s_1)^2 + (x_2 - s_2)^2 + 2c} \right] \\
(4.38)
\]

In this expression we recall that \((s_1, s_2)\) is the signal vector, \((x_1, x_2)\) is the observation vector, and \(b\) and \(c\) are the parameters of the a priori \( \gamma \) distribution.
For a given threshold value $\beta$, the contours of constant $\ln f(x_1, x_2)$ are given by

$$(b + 2) \left[ \frac{x_1^2 + x_2^2 + 2c}{(x_1 - s_1)^2 + (x_2 - s_2)^2 + 2c} \right] = \beta$$

(4.39)

If both sides of Eq. 4.39 are divided by the quantity $b + 2$, and the exponential taken, the result is given by the following equation:

$$\frac{x_1^2 + x_2^2 + 2c}{(x_1 - s_1)^2 + (x_2 - s_2)^2 + 2c} = \Delta$$

(4.40)

In this equation the quantity $\Delta$ is defined as

$$\Delta = \exp \left[ \beta / (b + 2) \right]$$

(4.41)

Using algebraic manipulations Eq. 4.40 can be reduced to the following form.

$$(\Delta - 1) (x_1^2 + x_2^2) - 2\Delta(x_1 s_1 + x_2 s_2) + \Delta(s_1^2 + s_2^2) = -2c(\Delta - 1)$$

(4.42)

Dividing each side of this latter equation by the quantity $(\Delta - 1)$ yields

$$x_1^2 - 2 \left( \frac{\Delta}{\Delta - 1} \right) x_1 s_1 + x_2^2 - 2 \left( \frac{\Delta}{\Delta - 1} \right) x_2 s_2 + \left( \frac{\Delta}{\Delta - 1} \right) E_s = -2c$$

(4.43)

where $E_s$ is defined as

$$E_s = s_1^2 + s_2^2$$

(4.44)
We note that $E_s$ is a measure of the signal energy.

If we now complete the square of the quadratic terms on the right-hand side of Eq. 4.43 and rearrange terms, the result may be written in the following form.

$$\left[ x_1 - \Delta s_1 / (\Delta - 1) \right]^2 + \left[ x_2 - \Delta s_2 / (\Delta - 1) \right]^2 = \Delta E_s / (\Delta - 1)^2 - 2c$$

(4.45)

This latter equation is recognized immediately as the equation of a circle with radius

$$r = \left[ \Delta E_s / (\Delta - 1)^2 - 2c \right]^{\frac{1}{2}}$$

(4.46)

and center at the point

$$x_1 = \Delta s_1 / (\Delta - 1), \quad x_2 = \Delta s_2 / (\Delta - 1)$$

(4.47)

Thus, it follows from Eqs. 4.45, 4.46, and 4.47 that the loci of constant log-likelihood ratio after two observations are given by circles in observation space. The centers of these circles are functions of the signal vector $(s_1, s_2)$ and the "threshold" value $\Delta$. In addition, the centers lie along a line given by

$$x_2 = s_2 x_1 / s_1$$

(4.48)

and thus are oriented in the direction of the signal vector.

The circle radii are functions of the signal energy $E_s$, the "threshold" value $\Delta$, and the $\gamma$ density parameter $c$. Since the parameter $c$ is a function of both the mean $m$ and the variance $\nu$ of the a priori $\gamma$ density function (see Eq. 4.17), it follows that the circle radii are also functions of both of these quantities. In Fig. 4.5 a typical plot of the circle radii versus the quantity $\Delta$ is given. For values of $\Delta$ outside

58
of the range of \([\Delta_1, \Delta_2]\) the radii are imaginary and therefore the "circles" degenerate to points in the \((x_1, x_2)\) plane. This follows since log likelihood values outside of the range \([\Delta_1, \Delta_2]\) do not occur. For \(\Delta = 1\), the circle radius is infinite. However, from Eq. 4.47 the center point is also at infinity and therefore, the \(\Delta = 1\) locus becomes a straight line. In addition, it follows from Eq. 4.47 that when \(s_1\) and \(s_2\) are positive, a \(\Delta\) value in the range \([\Delta_1, 1]\) implies that the circle centers are in the negative quadrant of the \((x_1, x_2)\) plane while a \(\Delta\) value in the range \([1, \Delta_2]\) implies that they lie in the positive quadrant of this plane.

\[E_s = 2\]
\[c = 1\]

Fig. 4.5. Typical plot of \(\Delta\)-circle radii versus \(\Delta\)
In Figs. 4.6, 4.7, and 4.8 typical contours of constant log likelihood ratio are illustrated with the features discussed above incorporated in the diagrams. These diagrams are parameterized by the signal energy $E_S$ and the mean and variance of the $\gamma$ distribution, $m$ and $v$. The correspondence between the quantities $\beta$ and $\Delta$ appearing in the figures is given by Eq. 4.41. The relationship of the figures to receiver operation is as follows. For a given value of $\Delta$, the corresponding contour of constant log-likelihood ratio forms the boundary between mutually exclusive sets of observation-response pairs. For $\Delta \geq 1$, all those observation pairs within or on the boundary of the corresponding $\Delta$-circle are ascribed to the hypothesis SN, and all those outside of the $\Delta$-circle are ascribed to the hypothesis N. For the situation where $\Delta < 1$ the opposite situation is encountered; all those observation pairs falling within or on the boundary of the corresponding $\Delta$-circle are considered to arise from the hypothesis N while those outside of the $\Delta$-circle are ascribed to the hypothesis SN. In this manner the receiver ascribes a decision function to each point in the observation space.

For the special case of low false alarm probabilities (a highly desirable feature in many situations) the value of the true threshold $\beta$ is generally quite high so that $\Delta > 1$. In terms of the corresponding $\Delta$-circle, only those observation pairs within or on the boundary of this contour are ascribed to the hypothesis SN. In other words, an important feature of the receiver operation is an inherent automatic gain control (AGC) action in which observation pairs which result in large excursions from the expected signal are rejected. This type of operation seems to be in correspondence with most observation techniques in which unusually
Fig. 4.6. Typical graph of constant log-likelihood ratio countours with parameter values $E_s = 2$, $m = 1$, $v = 1$.
Fig. 4.7. Typical graph of constant log-likelihood ratio contours with parameter values $E_s = 2$, $m = 1$, $v = 2$
Fig. 4.8. Typical graph of constant log-likelihood ratio contours with parameter values $E_s = 2$, $m = 1$, $v = 3$ 

$\Delta = \exp(\beta / (b+2))$
large or spurious responses are disregarded. The important point is that this AGC action is not the simple type of AGC leading to triangular or conical $\Delta$ contours which is added to many receivers but is an inherent feature of the optimum receiver and results directly from the uncertainty which exists concerning the noise process. This result seems to indicate the theoretical role of AGC in optimum receiver design and in particular provides a theoretical basis for the hard limiters used by Taylor (Ref. 9) as discussed in Chapter I.

A comparison of contours in Figs. 4.6 through 4.8 with constant $\beta$ threshold values indicates another interesting result. This result is that as the variance or uncertainty of the $\gamma$ distribution, $v$, increases, contours with equal $\beta$ values become smaller and are centered more on the signal vector. Thus, as the noise process becomes less certain the effect of the receiver AGC action increases.

The overall result of this effect is to shift the coordinates of maximum log-likelihood ratio more toward the signal vector coordinates. To investigate this aspect of receiver operation, the coordinates of maximum log-likelihood ratio are determined below.

The coordinates of maximum log likelihood ratio are easily determined from Eq. 4.39. From this equation we see that the logarithm of the likelihood ratio attains its maximum value at the coordinate pair for which $\beta$ is a maximum. Reference to Eq. 4.41, however, shows that the coordinates for which $\beta$ is a maximum are also the coordinates for which $\Delta$ attains its maximum value. This maximum value for $\Delta$ is, from Fig. 4.5, $\Delta_2$, and at this value the corresponding $\Delta$-circle has a radius of zero magnitude. Therefore, substituting the value $\Delta = \Delta_2$ into Eq. 4.45 the following equation determining the coordinates of maximum

64
log-likelihood ratio is obtained.

\[ [x_1 - \Delta_2 s_1 / (\Delta_2 - 1)]^2 + [x_2 - \Delta_2 s_2 / (\Delta_2 - 1)]^2 = 0 \]

(4.49)

Both quantities on the left-hand side of Eq. 4.49 are positive, and therefore, for their sum to be zero each of them must equal zero. Equating each of these quantities to zero we obtain the coordinates of maximum log-likelihood ratio. They are given by

\[ x_1 = \frac{\Delta_2 s_1}{(\Delta_2 - 1)}, \quad x_2 = \frac{\Delta_2 s_2}{(\Delta_2 - 1)} \]

(4.50)

where

\[ \Delta_2 = 1 + \left( \frac{E_s}{4c} \right) + \sqrt{\left( \frac{E_s}{c} \right)^2 + \left( \frac{E_s}{2c} \right)^2} \]

(4.51)

From Eq. 4.50 it follows that the point of maximum log-likelihood lies on a line given by

\[ x_2 = \left\{ \begin{array}{c} s_2 \\ s_1 \end{array} \right\} x_1 \]

(4.52)

which passes through the origin and the coordinates of the signal vector \((s_1, s_2)\). As the quantity \((E_s / 4c) \to \infty\) the point of maximum log-likelihood ratio approaches the signal vector coordinates \((s_1, s_2)\). Thus the point of maximum likelihood, which generally corresponds to low false alarm probability values, lies in the vicinity of the signal vector \((s_1, s_2)\). This result is in contrast to the cross-correlation receiver in which maximum likelihood occurs as \(x, y \to \infty\).
The above results are easily generalized to the case of more than two observations. The discussion is entirely analogous with the $\Delta$-circles becoming $\Delta$-hyperspheres in the higher dimensional observation space. All of the results discussed above them follow directly. The generalization to higher dimensions is most easily accomplished by choosing a coordinate system in observation space such that the first unit vector is parallel to the known signal vector with coordinate $y_1 = \langle x, s \rangle / \sqrt{E_s}$ where $\langle \cdot , \cdot \rangle$ denotes the inner product of two vectors and $x$ and $s$ are the observation and signal vectors. The remaining coordinates are then chosen to form an arbitrary orthonormal basis, $y_2, y_3, \ldots, y_k$. Using this system it can be shown that the contours of constant log likelihood ratio are given by

\[
(y_1 - \frac{\Delta_k}{\Delta_k - 1} \sqrt{E_s})^2 + \sum_{i=2}^{k} y_i^2 = \frac{\Delta_k}{(\Delta_k - 1)^2} E_s - 2c
\]

where $E_s$ is the total signal energy and $b$ and $c$ describe the a priori knowledge of the $\gamma$ density. The parameter $\Delta_k$ is related to the log likelihood ratio value $\beta$ and the updated $\gamma$ density parameter $b_k = b + k/2$ by

\[
\Delta_k = \exp(\beta / b_k + 1)
\]

The interesting thing to note about this result is that as the number of observations $k$ becomes large $\Delta_k \to 1$. The implication of this result in terms of the hypersphere equation is that as $k$ becomes large the hyperspheres become hyperplanes and the receiver becomes a correlation process. The relationship of this result to the adaptive nature of the receiver is investigated in Sec. 4. 4. 5. 3.
To determine the effect of noise uncertainties on detection performance we must now turn to an evaluation of the performance of the optimum detector for the SKE-NUL case realized in this section. Before doing this, however, a brief review of the techniques involved in evaluating receiver performance is in order. This review is given in the next section with the evaluation of the optimum receiver following. The reader familiar with evaluation techniques may proceed directly to Sec. 4.4.5.

4.4.4. Review of Receiver Evaluation Techniques. The detection performance of a detection receiver is summarized on a receiver operating characteristic (ROC). The ROC is a convenient means of graphically displaying the performance of the receiver in a given situation involving signal, noise, and the receiver itself. The presence of noise in the decision implies that the possibility of error is always present. There are two kinds of errors, termed a false alarm and a miss respectively, which can occur in the basic decision process. Correspondingly, there are two kinds of correct responses, termed a correct detection and a correct rejection, which can occur (see Sec. 2.1). A false alarm results when the response is "signal present" ("A") and the input actually consists of noise alone. On the other hand, a miss occurs when the response is "signal absent" ("B") and the input is signal and noise. Similarly, a correct detection results when the input is signal and noise and the receiver correctly responds "signal present" ("A"). A correct rejection occurs when the response is "signal absent" ("B") and the input consists of noise alone. In a detection problem there are probabilities associated with these possible responses as a result of the interaction of the signal process, the noise process, and the receiver. The standard notation utilized to denote these probabilities is as follows.
P("A"|N) probability of false alarm
P("B"|SN) probability of a miss
P("A"|SN) probability of a correction detection
P("B"|N) probability of a correction rejection

where

"A" is the response "signal present"
"B" is the response "signal absent"
SN is the hypothesis "signal and noise"
N is the hypothesis "noise alone"

An important point to note is that the above probabilities are not independent. Indeed, the following relationships exist among them.

\[ P(\text{"A"} \mid \text{SN}) + P(\text{"B"} \mid \text{SN}) = 1 \quad (4.53) \]

\[ P(\text{"A"} \mid \text{N}) + P(\text{"B"} \mid \text{N}) = 1 \quad (4.54) \]

The dependence among the four probabilities as displayed by Eqs. 4.53 and 4.54 implies that all of the information which they contain may be conveyed by a consideration of only two of them, each conditioned on a different hypothesis. For example, all of the information may be conveyed by a plot of the probability of detection, \( P(\text{"A"} \mid \text{SN}) \), versus the probability of a false alarm, \( P(\text{"A"} \mid \text{N}) \) for all possible threshold settings on the receiver output. Such a plot is termed an ROC curve and, as discussed above, summarizes the detection performance of the receiver (Ref. 15.).

An ROC curve is called "normal" if it can be parameterized by the normal probability distribution function. This means that the probability of detection \( P(\text{"A"} \mid \text{SN}) \) and the probability of false alarm
\[ P('A' \mid N) \] can be written as

\[ P('A' \mid SN) = \Phi(\lambda + d'), \text{ when } P('A' \mid N) = \Phi(\lambda) \] (4.55)

where

\[ \Phi(\lambda) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\lambda} \exp \left( -\frac{t^2}{2} \right) \, dt \] (4.56)

Normal ROC curves arise whenever the natural logarithm of the likelihood ratio is normally distributed under N and SN with equal variances and means separated by the variance. From Eq. 4.55 it follows that an entire normal ROC curve may be characterized by the one parameter \(d'\). In addition, when ROC curves are plotted on log-log paper, as is frequently convenient, normal ROC curves are linearized.

Much of the utility of normal ROC curves lies in the fact that for signals in white Gaussian noise they provide an upper bound on detection performance. For example, if a signal of energy \(E\) is known exactly and the noise power per cycle per second is known to be \(N_0\), then the ROC for the optimum receiver is normal and the parameter \(d'\) has the value

\[ d' = \sqrt{\frac{2E}{N_0}} \] (4.57)

This "upper bound" property of normal ROC curves provides a basis for comparison of various ROC curves. In general, the ROC curves obtained in this work are not normal. However, they are approximately normal and therefore the parameter \(d'_{e}\), or more conveniently, \(d_{e} = (d'_{e})^2\), the "equivalent detectability" serves as a convenient quantitative measure
to characterize them. The equivalent detectability is defined as the equivalent normal \( d = (d')^2 \) measured at the intersection of the ROC with the negative diagonal.

![ROC Curve Diagram](image)

**Fig. 4.9.** Normal ROC curves with parameter \( d \)

The ROC's for the signal known exactly case for several values of the parameter \( d = (d')^2 \) are shown in Fig. 4.9. Figure 4.10 illustrates two non-normal ROC curves with the parameter \( d_e \) specified as the equivalent normal \( d \) measured at the intersection of the ROC with the negative diagonal.
In general, the evaluation of the performance of an optimum receiver requires the distribution of the likelihood ratio, or some monotonic function of it, under both signal mixed with noise and noise alone. These distributions are used to determine, as a function of threshold setting, the $P("A\mid SN)$ and $P("A\mid N)$ values needed to plot the ROC curve. For example, if the probability densities under the two hypotheses are those shown in Fig. 4.11, then for the threshold value shown, $P("A\mid SN)$ is given by the striped area and $P("A\mid N)$ is given by the cross-hatched area. The ROC is then determined as a function of all possible values.
of the threshold setting. In practice, the determination of the distribution functions of the likelihood ratio in an analytic form may be considerably difficult. One is usually able to specify the functions in general; however, the evaluation of the integrals involved frequently becomes difficult. At this point a digital computer is frequently used.

![Diagram of probability density functions of the likelihood ratio under SN and N](image)

Fig. 4.11. Probability density functions of the likelihood ratio under SN and N

Another approach to the evaluation of receiver performance is given by the use of digital computer simulation techniques. For this experimental approach the given detection situation including receiver operations is simulated on the digital computer and the signal with noise and noise alone density functions are sampled. Alternatively, use may be made of a fundamental theorem in decision theory which states that, "The likelihood ratio of the likelihood ratio is the likelihood ratio" or mathematically that (Ref. 15.)

\[ l'(l) = \frac{f(l | SN)}{f(l | N)} = l \]

(4.58)

The use of this theorem implies that only the noise alone density need be sampled since the density of the likelihood ratio under SN can be
obtained from the density of the likelihood ratio under N by multiplying by the likelihood ratio. In any case, the simulation approach, although an approximate one, has been found to provide extremely good accuracy with the results limited only by the number of computer runs feasible. Throughout this report receiver performance is evaluated utilizing both the analytical and computer simulation approaches.

4.4.5. Receiver Evaluation. In the previous section the basic concepts involved in evaluating receiver performance were reviewed. In this section the detection performance of the optimum receiver for the signal known exactly in noise of unknown level (SKE-NUL) case is determined. This evaluation of the detection performance is summarized in terms of receiver operating characteristics or ROC curves. The ROC curves are determined using techniques discussed in the previous section. The primary emphasis of the evaluation is placed on obtaining quantitative principles which are useful in general detection situations. The following points are emphasized:

1. Effect of noise uncertainties on the detection performance of the optimum receiver

2. Comparison of the detection performance of the optimum receiver and a suboptimum receiver

3. Operation of the sequential receiver realization of the optimum receiver and its "adaptive" characteristics.

4.4.5.1. Effect of Noise Uncertainties. To evaluate the effect of noise uncertainties on the performance of the optimum receiver developed in Sec. 4.3.2 a dc input signal was utilized. This waveform was used because, as will become evident, signal energy and not
signal waveshape affects receiver performance. The waveform was specified by equal sample values of amplitude $\sqrt{E_s/k}$ and the ROC curves were obtained for a total of 100 observations (i.e., $k = 100$) and two values of $E_s$; namely, $E_s = 1$ and $E_s = 2$. The procedure for determining the ROC curves consisted of determining false alarm and detection probabilities utilizing an IBM 7090 digital computer.

To consider the method used for obtaining these probabilities let us first consider the calculation of false alarm probability. The probability of false alarm is the probability of responding "A" when noise alone is the actual cause of the input. From before, this probability is denoted by $P('A'\mid N)$ where the response "A" occurs whenever the likelihood ratio of the input exceeds a critical value, denoted $\beta$. Therefore, we can write

$$P('A'\mid N) = P(\ell > \beta \mid N) \quad (4.59)$$

Thus, to calculate the false alarm probability the distribution of the likelihood ratio in noise alone, or equivalently, the distribution in noise of any quantity which is monotonic in the likelihood ratio is needed. The logarithm is a monotonic function of the likelihood ratio, and therefore, we can write Eq. 4.59 as

$$P('A'\mid N) = P(\ln \ell > \ln \beta \mid N) \quad (4.60)$$

Using Eqs. 4.35 and 4.36 the logarithm of the likelihood ratio after $k$ observations can be written as

$$\ln \ell = (b_k + 1) \ln \left( \frac{c_k}{g_k} \right) \quad (4.61)$$

where for the constant amplitude signal
\[ c_k = 2c + \sum x_i^2 \]  
(4.62)

and

\[ g_k = 2c + \sum x_i^2 - 2\sqrt{E_s/k} \sum x_i + E_s \]  
(4.63)

Let us now define the following two random variables.

\[ u_k = \sum x_i^2 \]  
(4.64)

\[ v_k = \sum x_i \]  
(4.65)

Using these two variables, Eqs. 4.62 and 4.63 can be expressed as

\[ c_k = 2c + u_k \]
and

\[ g_k = 2c + u_k - 2\sqrt{E_s/k} v_k + E_s \]

and therefore from Eq. 4.61 the logarithm of the likelihood ratio after \( k \) observations can be expressed in terms of the random variables \( u_k \) and \( v_k \). Hence, to determine the distribution of the logarithm of \( \ell \) conditional to noise we must first determine the joint distribution of \( u_k \) and \( v_k \) conditional to noise.

To determine the joint distribution of \( u_k \) and \( v_k \) conditional to noise and for an arbitrary value of \( k \), let us first consider the case for which \( k = 2 \). Under this condition

\[ u_2 = x_1^2 + x_2^2 \]  
(4.67)

and

\[ v_2 = x_1 + x_2 \]
The system of equations defined in Eq. 4.67 is a transformation of $E^2 - E^2$ which maps the $x_1, x_2$ plane onto the region

$$0 \leq u_2, \quad v_2^2 \leq 2u_2$$

The Jacobian of the transformation is given by

$$J = (2u_2 - v_2^2)^{-\frac{1}{2}} / 2$$ \hspace{1cm} (4.68)

and since it is always positive in the region over which it is defined, it is equal to its absolute value in this region. The inverse of the transformation has two solutions, namely,

$$x_1 = v_2 + \sqrt{2u_2 - v_2^2}, \quad x_2 = v_2 - \sqrt{2u_2 - v_2^2}$$

and

$$x_1 = v_2 - \sqrt{2u_2 - v_2^2}, \quad x_2 = v_2 + \sqrt{2u_2 - v_2^2}$$

Because this inverse transformation is double-valued, the Jacobian given in Eq. 4.68 must be multiplied by a factor of 2 when it is used to determine the joint density of $u_2$ and $v_2$.

The joint density of $x_1$ and $x_2$ conditional to $\gamma$ and $N$ is Gaussian since we have assumed conditional independence. Therefore, this joint density function is given by

$$f(x_1, x_2 | \gamma, N) = (2\pi)^{-1} \gamma \exp \left[ -\frac{(x_1^2 + x_2^2)\gamma}{2} \right]$$ \hspace{1cm} (4.69)

To determine the conditional density of $u_2$ and $v_2$, we substitute for the variables $x_1$ and $x_2$ in Eq. 4.69 using the Eq. 4.67. We then multiply the result by twice the absolute value of the Jacobian given Eq. 4.68.
The result of these operations is the joint density function of \( u_2 \) and \( v_2 \) conditional to \( \gamma \) and \( N \).

\[
f(u_2, v_2 | \gamma, N) = (2\pi)^{-1} \gamma^{(2u_2 - v_2^2)^{-\frac{1}{2}}} \exp[-u_2 \gamma / 2],
\]

\[
0 \leq u_2, \quad v_2^2 \leq 2u_2
\]

\[
= 0 \quad \text{otherwise}
\]

(4.70)

To obtain the joint density of \( u_2 \) and \( v_2 \) conditional only to \( N \), the density in Eq. 4.70 must be averaged over the a priori density of \( \gamma \) given by Eq. 4.11. Performing the required averaging yields

\[
f(u_2, u_1 | N) = C_2 (2u_2 - v_2^2)^{-\frac{1}{2}} (u + 2c)^{-(b+2)}
\]

\[
0 < u_2, \quad v_2^2 < 2u_2
\]

\[
= 0 \quad \text{otherwise}
\]

(4.71)

The quantity \( C_2 \) in this latter expression is the normalizing constant and is given by

\[
C_2 = (2c)^{b+1} (b + 2)
\]

To determine the joint density of \( u_k \) and \( v_k \) conditional to noise and for an arbitrary value of \( k \), we first assume that the following conditional form holds for \( k-1 \):

\[
f(u_{k-1}, v_{k-1} | \gamma, N) = D_{k-1}^{(k-1)/2} [(k-1)u_{k-1} - v_{k-1}^2]^{(k-4)/2} \exp[-u_{k-1} \gamma / 2].
\]

77
\[ 0 \leq u_{k-1}, \quad v_{k-1}^2 \leq (k-1)u_{k-1} \]

\[ = 0 \quad \text{otherwise} \]

(4.72)

where the normalizing constant \( D_{k-1} \) is given by

\[ D_{k-1} = \left[ 2^{(k-1)/2} \cdot \frac{(k-1)(k-3)}{2} \cdot \frac{\Gamma \left( \frac{k-2}{2} \right)}{\pi^{\frac{1}{2}}} \right]^{-1} \]

Our intention is to show that if this conditional form of the density is true for \( k-1 \), then it is true for \( k \), and, therefore, by induction, that it holds for all values of \( k \) (\( k \neq 1 \)). (Note that the case \( k = 2 \) was considered above.)

From the resultant conditional joint density we can then determine the joint density of \( u_k \) and \( v_k \) conditional to \( N \). This latter function is the result we desire.

From the definition of \( u_k \) and \( v_k \) we have

\[ u_k = u_{k-1} + x_k^2 \]

(4.73)

and

\[ v_k = v_{k-1} + x_k \]

(4.74)

In addition, we also define for computation purposes

\[ w = x_k \]

(4.75)

From the assumption of conditional independence, it follows that \( x_k \) is independent of \( u_{k-1} \) and \( v_{k-1} \) conditional to knowledge of the value of \( \gamma \). Hence, the conditional joint density of \( u_{k-1} \), \( v_{k-1} \), and \( x_k \) may be written as follows:

\[ f(u_{k-1}, v_{k-1}, x_k | \gamma, N) = f(u_{k-1}, v_{k-1} | \gamma, N) f(x_k | \gamma, N) \]

If we substitute the expressions for the density functions in the above
equation, we have

\[ f(u_{k-1}, v_{k-1}, x_k | \gamma, N) = (2\pi)^{-\frac{1}{2}} D_{k-1}^{\gamma k/2} [(k-1)u_k - v^2_{k-1}]^{(k-4)/2} \]

\[ \cdot \exp \left[ - (u_{k-1} + x_k^2) \gamma / 2 \right], \]

\[ 0 \leq u_{k-1}, v_{k-1} \leq (k-1)u_{k-1} \]

\[ = 0 \quad \text{otherwise} \]

(4.76)

The Jacobian of the transformation of \( E^3 \rightarrow E^3 \) defined by Eqs. 4.73, 4.74, and 4.75 is given by

\[ J = 1 \]

Using this value of the Jacobian together with the transformation equations and the joint density function given in Eq. 4.76, the joint density function of \( u_k, v_k \) and \( w \) conditional to \( \gamma \) and \( N \) is easily obtained. The result is

\[ f(u_k, v_k, w | \gamma, N) = G^{\gamma k/2} [(k-1)(u_k - v^2_k / k) - (k^2 w - v_k / k^2)^2]^{(k-4)/2} \exp \left[ - u_k \gamma / 2 \right] \]

\[ 0 < (k-1)(u_k - v^2_k / k) \]

\[ (k^2 w - v_k / k^2)^2 \leq (k-1)(u_k - v^2_k / k) \]

\[ = 0 \quad \text{otherwise} \]

(4.77)

The normalizing constant \( G \) is given by

\[ G = \left[ \pi 2^{k/2} (k-3)/2 \left( \Gamma \left( \frac{k-2}{2} \right) \right)^{-1} \right] \]
To obtain the conditional joint density for $u_k$ and $v_k$ we must integrate Eq. 4.77 with respect to the variable $w$. Expressing this integration explicitly we have

$$f(u_k, v_k | \gamma, N) = G \int_{w_1}^{w_2} \gamma^{k/2} \left[ \frac{(k-1)(u_k - v_k^2/k)}{w} - \left( \frac{k^{1/2}}{2} - v_k^{1/2}/k \right)^2 \right]^{(k-4)/2} \exp \left[ -u_k \gamma / 2 \right] dw \quad (4.78)$$

where the limits of integration, $w_1$ and $w_2$, are given by

$$w_{1,2} = \left( v_k^{1/2}/k^{1/2} \pm \sqrt{(k-1)(u_k - v_k^2/k)} \right) / k^{1/2}$$

To perform the indicated integration we make the following substitution:

Let

$$z = \left[ k^{1/2} w - v_k^{1/2} / k^{1/2} \right]^2 / (k-1)(u_k - v_k^2/k)$$

Using this substitution Eq. 4.78 becomes

$$f(u_k, v_k | \gamma, N) = G \gamma^{k/2} \left[ (k-1)(u_k - v_k^2/k) \right]^{(k-3)/2} \exp \left[ -u_k \gamma / 2 \right] \cdot \int_0^1 z^{-1/2} (1 - z)^{(k-4)/2} dz$$

$$0 \leq u_k, \quad v_k^2 \leq k u_k$$

$$= 0 \quad \text{otherwise} \quad (4.79)$$

The integral appearing in Eq. 4.79 is expressible in terms of Gamma functions. If we express it in this manner and then simplify (using the value of $G$ previously determined), Eq. 4.79 becomes

$$f(u_k, v_k | \gamma, N) = D_k \gamma^{k/2} \left[ k u_k - v_k^2 \right]^{(k-3)/2} \exp \left[ -u_k \gamma / 2 \right]$$

80
\[ 0 \leq u_k, \quad v_k^2 \leq k u_k \]

\[ = 0 \quad \text{otherwise} \quad (4.80) \]

where

\[ D_k = \left[ 2^{k/2} \frac{2^{(k-2)/2}}{\Gamma \left( \frac{k-1}{2} \right) \pi^{1/2}} \right]^{-1} \]

By comparing Eqs. 4.72 and 4.80 it is evident that the latter expression is of the same form as that assumed for \( k-1 \) (Eq. 4.72). Thus, the proof by induction is complete and Eq. 4.80 gives the joint density of \( u_k \) and \( v_k \) conditional to \( \gamma \) and \( N \) for an arbitrary value of \( k \) \((k \neq 1)\).

To obtain the joint density of \( u_k \) and \( v_k \) conditional to \( N \), Eq. 4.80 must be averaged over the a priori density for \( \gamma \) given by Eq. 4.11. Performing this calculation we have

\[ f(u_k, v_k | N) = C_k \frac{[k u_k - v_k^2]^{(k-3)/2}}{[u_k + 2c]^{b+1+k/2}} \]

\[ = 0 \quad \text{otherwise} \quad (4.81) \]

The normalizing constant \( C_k \) is given by

\[ C_k = \frac{(2c)^{b+1}}{\pi^{1/2} k^{(k-2)/2}} \frac{\Gamma(b+1+k/2)}{\Gamma(b+1) \Gamma\left(\frac{k-1}{2}\right)} \]

From Eq. 4.81 it is apparent that conditional to noise, the two random variables \( u_k \) and \( v_k \) are not independent. For the purpose of obtaining the ROC curves, it is much more convenient to have the likelihood ratio
expressed as a function of independent random variables. Therefore we define the following transformation between $u_k$, $v_k$ and the new variables $r_k$ and $t_k$.

$$ r_k = \frac{u_k}{(u_k + 2c)} $$

and

$$ t_k = v_k^2 / ku_k $$

(4.82)

Because of the region of definition of $u_k$ and $v_k$ (Eq. 4.81) it follows that

$$ 0 \leq t_k \leq 1 \quad , \quad 0 \leq r_k \leq 1 $$

(4.83)

The inverse of the transformation defined in Eq. 4.82 is given by

$$ u_k = 2cr_k / (1 - r_k) $$

(4.84)

$$ v_k = \pm \left( \frac{(2c_k r_k t_k)/(1 - r_k)}{1} \right)^{1/2} $$

The Jacobian of this transformation is

$$ J = c^{3/2} (2k)^{1/2} (1 - r_k)^{5/2} r_k^{1/2} t_k^{-1/2} $$

(4.85)

To obtain the joint density of $r_k$ and $t_k$ conditional to noise we must substitute Eq. 4.84 into Eq. 4.81 and then multiply the result by twice the Jacobian (Eq. 4.85) since the transformation is not single valued.

Performing these calculations we obtain

$$ f(r_k, t_k | \text{N}) = \left[ B(k/2, b + 1) B(1/2, \frac{k-1}{2}) \right]^{-1} r_k^{(k-2)/2} (1 - r_k)^b t_k^{-1/2} \cdot (1 - t_k)^{(k-3)/2} $$

82
\[ 0 \leq r_k \leq 1, \quad 0 \leq t_k \leq 1 \]

\[ = 0 \quad \text{otherwise} \quad \text{(4.86)} \]

The quantity \( B(\cdot, \cdot) \) is a Beta function.

From the form of Eq. 4.86 it is apparent that \( f(r_k', t_k'|N) \) can be written as a product of two density functions (of the Beta type), one for \( r_k \) and one for \( t_k' \). Therefore we conclude that \( r_k \) and \( t_k \) are independent.

To obtain the ROC curves the long-likelihood ratio was expressed in terms of the independent random variables, \( r_k \) and \( t_k' \), using Eqs. 4.61, 4.66 and 4.84. The corresponding distributions for these random variables were approximated on the digital computer using 500 (equal probability) points and a Monte Carlo approach was then used to determine the distribution of \( \ln f \) conditional to noise. The approach consisted of sampling from the distributions of the independent random variables, \( r_k \) and \( t_k' \), and then determining the log-likelihood ratio as a function of these values. The distribution for \( \ln f \) conditional to signal and noise was obtained using the fundamental theorem discussed in Section 4.4.4, and the ROC curves were obtained from these two distributions.

The ROC curves obtained by the method discussed above are presented in Figs. 4.12 through 4.15. Additional curves are given in Appendix B. These curves are parameterized by the constance distribution parameters \( b \) and \( c \) and the total signal energy \( E_s \). As mentioned previously \( k = 100 \) for all of the curves. The general appearance of the curves indicates that they are all approximately normal. For a given set of parameter values \( c \) and \( E_s \) the effect of the parameter value \( b \) on detectability can be seen. It is evident that detectability is directly related to the parameter \( b \); increasing values of \( b \)
Fig. 4.12. The ROC curves for the optimum receiver, SKE-NUL case, with parameters $E_s = 1.0$, $c = 0.5$, $b = -0.5$, 0.0, 1.0, 2.0, 3.0, 4.0
Fig. 4.13. The ROC curves for the optimum receiver, SKE-NUL case, with parameters $E_s = 1.0$, $c = 1.0$, $b = -0.5, 0.0, 1.0, 2.0, 4.0$
Fig. 4.14. The ROC curves for the optimum receiver, SKE-NUL case, with parameters $E_s = 2.0$, $c = 0.5$, $b = -0.5, 0.0, 0.5, 1.0, 1.5, 2.0$. 

86
Fig. 4.15. The ROC curves for the optimum receiver, SKE-NUL case, with parameters $E_s = 2.0$, $c = 1.0$, $b = -0.5, 0.0, 1.0, 2.0, 3.0, 4.0$.
implying increasing detectability. This result is consistent since the parameter \( b \) is directly proportional to the mean value of the constance parameter and, therefore, increasing \( b \) results in a higher expected signal-to-noise ratio. The influence of the parameter \( c \) on optimum receiver performance is easier to investigate if each ROC curve is read along the negative diagonal (i.e., read where the probabilities of each of the two possible kinds of errors are equal) and this measure of detectability, \( d_e \), plotted as a function of the parameter \( c \). In performing this procedure we are implicitly "normalizing" the ROC curves, however, it is apparent that the curves in Figs. 4.12 through 4.15 are already approximately normal. An alternative to this graphical method for obtaining a measure of \( d_e \) involves the determination of \( d_e \) as a function of the process parameters. In general, an analytic approach for obtaining \( d_e \) is very difficult; however, for the situation which we are considering an explicit function can be obtained. The details of the derivation are given in Appendix A. The result is shown to be

\[
d_e^{-\frac{1}{2}} = 2 \Phi^{-1}(\zeta) \tag{4.87}
\]

where

\[
\Phi(y) = \int_{-\infty}^{y} (2\pi)^{-\frac{1}{2}} \exp\left[-\frac{y^2}{2}\right] \, dy
\]

and

\[
\zeta = \int_{-\infty}^{\infty} \frac{(E_s/8c)^{1/2}}{\Gamma(\frac{\alpha + 1}{2})} \frac{\Gamma(\frac{\alpha}{2})}{\Gamma(\frac{\alpha}{2})} (\alpha 2\pi)^{-1/2} [1 + t^2/\alpha]^{-(\alpha+1)/2}
\]
with

\[ \alpha = 2(b + 1) \]

This expression for the detectability, \( d_e \), involves the use of the \( t \)-distribution function (Ref. 23) which is a well tabulated distribution. Using this relationship (or the graphical method discussed above) the quality of detection, \( d_e \), may be plotted as a function of the parameter \( c \). It is interesting to note that the expression for \( d_e \) depends only on the parameters \( E_s \), \( b \), and \( c \) and is independent of the number of observations \( k \). Thus, although shape of the ROC curves depends on the number of observations, the equivalent detectability measured along the negative diagonal is independent of this quantity for a given value of signal energy. It follows that any plot of \( d_e \) is valid for an arbitrary number of observations.

In Figures 4.16 and 4.17 graphs of \( d_e \) versus the \( \gamma \) density parameter \( c \) are plotted for the two values of \( E_s \) used in obtaining the ROC curves. Reference to these curves indicates that for given values of the parameters \( b \) and \( E_s \) detectability becomes quite pronounced at low values of \( c \). For a fixed value of the parameter \( b \), Eqs. 4.14 through 4.17 show that

\[ c = \frac{K}{m} \]

where \( K \) is a constant determined by the fixed value of \( b \). Therefore, for a fixed value of \( b \), lower values of \( c \) correspond to high expected values for the constance parameter which in turn implies high expected signal-to-noise ratios. Hence, the marked increase in detectability at low values of \( c \) is understandable.
Fig. 4.16. Detectability versus constance parameter, $c$, for SKE-NUL case with parameters $E_s = 1.0$, $b = -0.5$
$0.0, 1.0, 2.0, 4.0, 8.0$
Fig. 4.17. Detectability versus constance parameter, $c$, for SKE-NUL case with parameters, $E_s = 2.0$, $b = -0.5, 0.0, 1.0, 2.0, 3.0, 4.0$
The influence of the parameters $b$ and $c$ follows directly from their relationship to the mean and variance of the constance density function. However, these effects are somewhat difficult to interpret quantitatively and therefore, a set of ROC curves were obtained parameterized by the mean and variance, $m$ and $v$, of the constance distribution. These ROC's are plotted in Figs. 4.18 through 4.23 and their general appearance is the same as we have already seen. Additional ROC curves for this case are also given in Appendix B. To aid in the interpretation of this latter set of ROC's, plots of the equivalent detectability, $d_e$, versus the mean parameter $m$ were obtained. These plots are presented in Figs. 4.24 and 4.25 for the two values of signal energy. These latter diagrams were obtained graphically from the ROC curves and analytically from Eq. 4.59 using the transformations between the parameters $b$, $c$, and $m$, $v$ expressed in Eqs. 4.16 and 4.17. They present, in a very lucid manner, the effects of noise uncertainty on optimum detection performance. In the figures the detectability values for $v = 0$ are important since they represent an upper bound on the performance attainable by an optimum detector for the SKE-NUL case. As the expected value of the constance, $m$, increases, there is a corresponding increase in detectability. As $m$ continues to increase, the equivalent detectability begins to converge to that of the $v = 0$ bound. This result is even more enhanced by the inversion effect of the ROC curves which indicates higher detection probabilities as one moves away from the negative diagonal. Since the value of $d_e$ which has been plotted corresponds to the intersection of the ROC with the negative diagonal, the detectability for the inverted ROC's is actually higher at low false alarm values than that predicted on the basis of the "normal" characterization.
Fig. 4.18. The ROC curves for the optimum receiver, SKE-NUL case, with parameters $E_s = 1.0$, $v = 1.0$, $m = 0.707$, 1.0.
Fig. 4.19. The ROC curves for the optimum receiver, SKE-NUL case, with parameters $E_s = 1.0$, $v = 4.0$, $m = 1.414, 2.0, 4.0$.
Fig. 4.20. The ROC curves for the optimum receiver, SKE-NUL case, with parameters $E_s = 1.0$, $v = 8.0$, $m = 2.0$, 2.83, 4.0, 6.0
Fig. 4.21. The ROC curves for the optimum receiver, SKE-NUL case, with parameters $E_s = 2.0$, $v = 1.0$, $m = 0.707$, $1.0$, $1.414$, $2.0$
Fig. 4.22. The ROC curves for the optimum receiver, SKE-NUL case, with parameters $E_s = 2.0$, $v = 4.0$, $m = 1.414$, 2.0, 2.83, 4.0
Fig. 4.23. The ROC curves for the optimum receiver, SKE-NUL case, with parameters $E_s = 2.0$, $v = 8.0$, $m = 1.0, 1.414, 2.0, 2.83, 4.0, 5.0$.
Fig. 4.24. Detectability versus mean constance value, $m$, for SKE-NUL case with parameters, $E_s = 1.0$, $v = 0.0, 1.0, 2.0, 4.0, 8.0$
Fig. 4.25. Detectability versus mean constance value, $m$, for SKE-NUL case with parameters, $E_s = 2.0$, $v = 0.0, 1.0, 2.0, 4.0, 8.0$
The dependence of the equivalent detectability $d_e$ on the parameter $m$ can be explained by investigating the signal-to-noise ratio (SNR). From Eq. 4.57 and the definition of $\gamma$ as the constance or inverse of the noise power it follows that the signal-to-noise ratio is given by

$$SNR = E_s \gamma$$

(4.88)

where

$$E_s = \sum_{i=1}^{k} s_i^2$$

Thus for known signal energy but an uncertainty in the value of $\gamma$, the expected value of SNR is given by

$$E[SNR] = E_s E[\gamma]$$

or

$$E[SNR] = E_s m$$

(4.89)

Thus the expected value of the signal-to-noise ratio is proportional to $m$ and increases as $m$ increases. This result then causes an increase in $d_e$ as $m$ increases as evidenced by the behavior of the ROC curves.

From Eq. 4.88 the variance of SNR is given by

$$\text{var}(SNR) = E_s^2 \text{var}(\gamma)$$

or

$$\text{var}(SNR) = E_s^2 v$$

(4.90)

This direct dependence of the signal to noise ratio on the variance of the constance distribution $v$ is evident from the ROC curves. For fixed values of $m$ and $E_s$, an increase in the uncertainty parameter $v$ causes the detectability to initially decrease rapidly. The rate of this decrease
becomes smaller as the performance approaches the chance diagonal. From the shape of the curves, it is also apparent that at low expected SNR values (low values of m) greater values of uncertainty cause proportionately smaller decreases in detectability than at higher SNR values. Thus, at low signal-to-noise ratios the relative effect of initial noise uncertainty is greater than it is at higher signal-to-noise ratios. This is a significant result since in most detection situations we are concerned with operation at low signal to noise ratios.

4.4.5.2. Comparison with Suboptimum Receiver. In this section the performance of the optimum receiver for the SKE-NUL case is compared with the performance of a receiver which is suboptimum for this detection situation. It is interesting to make this comparison to determine the amount of performance increase which is attainable utilizing the optimum receiver. The suboptimum receiver considered here is a "likelihood correlator" which correlates the received observation with a stored reference value of the input signal and then computes the likelihood ratio of this quantity. This type of receiver was chosen because it represents a logical extension of a simple correlation type receiver.

To determine the detection output of the likelihood correlator we must determine the likelihood ratio of the correlation variable. Using the definitions of the previous section this variable may be defined as

$$y_c = \sqrt{E_s / k} v_k$$

for the dc input signal. The quantity \(\sqrt{E_s / k}\) is the amplitude of the signal samples and \(v_k\) was defined by Eq. 4.65 as

$$v_k = \sum_{1}^{k} x_i \quad (4.65)$$

The joint distribution of \(v_k\) and \(u_k\) conditional to noise was determined previously and is given by Eq. 4.81. If we integrate this distribution

102
with respect to $u_k$, the marginal distribution of $v_k$ conditional to $N$ is obtained. From this latter distribution, the distribution of $y_c$ conditional to $N$ is easily obtained by using the substitution $v_k = y_c \sqrt{E_s/k}$ and multiplying by the Jacobian of this transformation. The value of the Jacobian is easily seen to be $\sqrt{E_s/k}$. Performing these calculations yields the following conditional density for $y_c$

$$f(y_c | N) = \frac{\Gamma(b + 3/2)}{\sqrt{2\pi} c E_s \Gamma(b + 1)} \left( 1 + \frac{y_c^2}{2c E_s} \right)^{-\left(b+3/2 \right)}$$

(4.91)

Under the hypothesis SN the distribution for $y_c$ is the same as that given in Eq. 4.91 with a shift in the mean value from zero to a value equal to the signal energy $E_s$. Hence the distribution for $y_c$ conditioned on the hypothesis SN is given by

$$f(y_c | SN) = \frac{\Gamma(b + 3/2)}{\sqrt{2\pi} c E_s \Gamma(b + 1)} \left( 1 + \frac{(y_c - \sqrt{E_s})^2}{2c E_s} \right)^{-\left(b+3/2 \right)}$$

(4.92)

From the definition of the likelihood ratio, we now obtain the detection output of the likelihood correlator as the ratio of Eqs. 4.91 and 4.92. Taking this ratio the result is

$$f(y_c) = \left( \frac{y_c^2 + 2c E_s}{(y_c - \sqrt{E_s})^2 + 2c E_s} \right)^{(b+3/2)}$$

(4.93)

It is interesting to note that the two conditional distributions for $y_c$ (and hence the likelihood of $y_c$) are independent of the number of observations $k$. This result implies that the performance of the likelihood correlator is independent of the number of observations and only depends on the constance distribution and the signal energy (via $b$, $c$, and $E_s$).
In addition we note that the two distribution functions in Eqs. 4.91 and 4.92 are in the form of the \( t \)-distribution function (see Section 4.4.5.1).

The ROC performance curves for the correlator were obtained using the \( t \)-distribution function with the aid of the digital computer since prepared tables of the \( t \)-distribution were not in a convenient form. The methods used to obtain these curves are discussed in Appendix C.

The ROC curves for the likelihood correlator for the SKE-NUL case are presented in Figs. 4.26 through 4.28. In addition appropriate optimum receiver ROC's for \( k = 100 \) are included for comparison purposes. It is interesting to note that the detectability of both receivers is identical along the negative diagonal. This result is expected because, as we observed in Section 4.4.3, the operation of the optimum receiver becomes a correlation process for threshold values corresponding to the negative diagonal. This result seems to have great implications in communication systems in which binary signals are transmitted with equal probability. For systems of this type, the detection problem becomes one of deciding which of the two possible signals was sent and the comparator setting generally corresponds to the negative diagonal. Thus, the result we have obtained indicates that uncertainties in the noise do not affect the detection performance of the optimum receiver for this case. Additional inspection of the ROC's indicates that as one moves away from the negative diagonal the differential in performance increases. This differential becomes quite large at low false alarm probabilities and as the initial uncertainty level increases. It is in these regions that the performance of the optimum receiver becomes quite superior.

4.4.5.3. Sequential Operation. In Section 4.2 the classification outputs which result from realizing the optimum receiver in a

104
Fig. 4.26. Comparison of optimum and sub-optimum receiver performance, SKE-NUL case, with parameters $E_s = 2.0$, $v = 2.0$, $m = 0.5$, 1.0, 2.0
Fig. 4.27. Comparison of optimum and sub-optimum receiver performance, SKE-NUL case, with parameters $E_s = 2.0$, $v = 4.0$, $m = 1.414, 2.0, 2.83, 4.0$
Fig. 4.28. Comparison of optimum and sub-optimum receiver performance, SKE-NUL case, with parameters $E_s = 2.0$, $v = 8.0$, $m = 1.414$, 2.0, 2.83, 4.0.
sequential manner were pointed out. In this section we wish to inves-
tigate the sequential operation and the resulting classification outputs to
a greater extent. In particular we wish to determine how the detection
and classification outputs change with time as the receiver operates
sequentially. In addition, we also wish to investigate the "adaptive"
features of receiver performance as a function of observation time. To
evaluate these effects, conditional ROC curves were obtained as a
function of observation time. These curves yield the performance of
the receiver for a particular value of \( \gamma \) as contrasted with the average
performance with respect to all values of \( \gamma \) considered in the previous
sections. In this manner the performance for a particular sample run can
be investigated. This is in contrast to the average performance with
respect to many sample runs which were considered previously.

To perform the investigations discussed above, the detection
receiver was simulated on the digital computer and a number of trial
runs were made. The computer simulation consisted of approximating
the noise and signal-plus-noise distributions for given values of noise
level and using values from these distributions as inputs to the receiver.
Although a number of runs were made, the results of only a few are
presented since they were quite similar.

The selected results of the computer simulation runs are
presented in Figs. 4.29 through 4.31 and they illustrate the classification
outputs as a function of the observation time. The classification output
consisting of the probability density function of the noise constance,
conditioned on the particular set of observations obtained up to time \( t_k \)
(indexed by \( k \)) and the noise alone hypothesis, is presented in the first
column of these figures. This output represents the a posteriori density
of the noise constance conditional to a noise input to the receiver. It is a function of both time and the observation set. The second column is a graph of the mean and variance of the conditional distribution given in the first column. These quantities are plotted as a function of the number of observations.

The third column presents the probability density function of the noise constance. This density function is conditioned on the received observation set and the signal mixed with noise hypothesis. This classification output represents the a posteriori density of the noise constance as a function of both time and the observation set, and it is conditional to the hypothesis that signal and noise are present at the receiver input. The fourth column presents the mean and variance of this latter distribution and is plotted as a function of the observation time.

In the final column the sequential detection output is presented. This output is the value of the likelihood ratio at each time $t_k$ (indexed by k) for the particular set of observations obtained up to that time.

In sample Runs 1 through 4 the receiver input consisted of signal plus noise and in Runs 5 and 6 the input consisted of noise alone. The a priori constance density functions varied from run to run. For all of the runs, the expected signal-to-noise ratio (see Sec. 4.4.5.1) at the end of 100 observations was maintained at a value of 6.53 db. It should again be noted that all of the figures represent particular "sample values" of the general stochastic process as contrasted to average values.

From the results of the trial runs, the sequential operation of the receiver is apparent. As observation time increases, the proper classification output changes with time in such a manner that it approaches the true value of the noise constance. Because of the noise fluctuations
Simulation run no. 1, parameters $m = 6$, $v = 8$
Simulation run no. 2, parameters $m = 6$, $v = 8$

Fig. 4.29. Simulation runs, SKE-NUL case
Simulation run no. 3, parameters $m = 4$, $v = 4$
Simulation run no. 4, parameters $m = 4$, $v = 8$

Fig. 4.30. Simulation runs, SKE-NUL case
Simulation run no. 5, parameters $m = 2.83, \nu = 4$
Simulation run no. 6, parameters $m = 2.83$, $v = 8$

Fig. 4.31. Simulation runs, SKE-NUL case
from run to run, the time of build-up to the true constance value varies, however, this value is always approached after a sufficient number of observations have been taken. This fact is evidenced by the approach of the mean to the true value of the noise constance and the decrease of the variance to zero. The term "proper" is used in this discussion since it must be remembered that the two classification outputs are conditioned on the two hypotheses respectively. The choice of the correct classification output is governed by the value of the detection output, the likelihood ratio, presented in the final column. This choice may occasionally be incorrect, however, the resulting error is a function of the stochastic nature of the problem and not of failures in the receiver operation.

The conditional ROC curves as a function of observation time are presented in Figs. 4.32 through 4.37. These curves are parameterized by the values of the \( \gamma \) density parameters \( m \) and \( n \) used for the simulation runs. For each set of curves, the expected signal-to-noise ratio at the end of 100 observations was maintained at 6.53 db as discussed above.

From the ROC diagrams, the "learning" feature of the receiver becomes apparent. As the number of observations increases, the receiver performance asymptotically approaches the upper performance bound determined by the absence of noise uncertainty. This approach to the upper bound is evidenced by the approach of the ROC curves to normality (straight lines) as the number of observations increases. For a given number of observations (i.e., given value of signal-to-noise ratio) the best performance obtainable in the absence of noise uncertainty is given by a straight line (normal ROC curve) through the intersection of the ROC curves in Figs. 4.32 through 4.37 with the negative diagonal. This result follows since, in the absence of noise uncertainty, the optimum
Fig. 4.32. Conditional ROC curves as a function of observation time, SKE-NUL case, with parameters $m = 6, \nu = 4$
Fig. 4.33. Conditional ROC curves as a function of observation time, SKE-NUL case with parameters $m = 6$, $v = 8$
Fig. 4.34. Conditional ROC curves as a function of observation time, SKE-NUL case, with parameters $m = 4$, $v = 4$
Fig. 4.35. Conditional ROC curves as a function of observation time, SKE-NUL case, with parameters $m = 4$, $v = 8$
Fig. 4.36. Conditional ROC curves as a function of observation
time, SKE-NUL case, with parameters \( m = 2.58, v = 4 \)
Fig. 4.37. Conditional ROC curves as a function of observation time, SKE-NUL case, with parameters $m = 2.58$, $v = 8$
performance is given by a correlator (Section 2.2) and along the negative diagonal the operation of the sequential receiver reduces to that of a correlator (Section 4.3.3). Thus, we see that as the number of observations increases, the conditional performance of the sequential receiver approaches normality or, in other words, approaches the best performance attainable in the absence of noise uncertainty. This adaptive behavior occurs as the receiver attempts to learn the actual value of the noise level.

To consider this latter result more fully, let us consider the log-likelihood ratio after k observations as given by Eq. 4.61

$$\ln \ell = (b_k + 1) \ln \left( \frac{c_k}{g_k} \right)$$

(4.61)

We wish to show that as the number of observations becomes large, the distribution of $\ln \ell$ conditional to a particular value of $\gamma$ approaches normality as the computer simulation results indicate. To do this, let us first express Eq. 4.61 in terms of the variables $u_k$ and $v_k$ using the definitions expressed in Eq. 4.66. We have

$$\ln \ell = (b_k + 1) \ln \left[ \frac{u_k + 2c}{u_k - 2\sqrt{E_s/k} \cdot v_k + E_s + 2c} \right]$$

(4.94)

Dividing the numerator and denominator of this latter expression by $u_k + 2c$ we have

$$\ln \ell = -(b_k + 1) \ln \left[ 1 + \frac{E_s - 2\sqrt{E_s/k} \cdot v_k}{u_k + 2c} \right]$$

(4.95)

In order to investigate the asymptotic behavior of $\ell$ conditional to a
particular value of $\gamma$ let us define the new variables $p_k$ and $q_k$ by the transformation

$$p_k = \frac{\gamma u_k - \gamma v_k^2 / k - (k - 1)}{\sqrt{2(k - 1)}}$$

(4.96)

$$q_k = \frac{v_k}{k^{\frac{3}{2}}}$$

The inverse of this transformation is given by

$$u_k = \gamma^{-1} \sqrt{2(k - 1)} p_k + \gamma^{-1}(k - 1) + q_k^2$$

(4.97)

and

$$v_k = k^{\frac{1}{2}} q_k$$

The Jacobian of the transformation is

$$J = \gamma^{-1} \sqrt{2k(k - 1)}$$

(4.98)

Using Eq. 4.97 we can express $\ln f$ in terms of the variables $p_k$ and $q_k$ in the following form

$$\ln f = -(b_k + 1) \ln \left[ 1 + \frac{E_s - 2\sqrt{E_s} q_k}{\gamma^{-1} \sqrt{2(k - 1)} p_k + q_k + \gamma^{-1}(k - 1) + 2c} \right]$$

(4.99)

Simplifying this latter expression

$$\ln f = -(b_k + 1) \ln \left[ 1 + \frac{\gamma(E_s - 2\sqrt{E_s} q_k)}{(k-1)\left(\frac{\sqrt{2} p_k}{\sqrt{k-1}} + \frac{\gamma q_k^2}{k-1} + \frac{2c}{k-1}\right)} \right]$$

(4.100)
The joint distribution of $p_k$ and $q_k$ conditional to $\gamma$ and $N$ can be determined using the Jacobian given in Eq. 4.98 and the joint distribution of $u_k$ and $v_k$ conditional to $\gamma$ and $N$ as given by Eq. 4.80. Substituting for $u_k$ and $v_k$ in Eq. 4.80 by using Eq. 4.97 and then multiplying by the absolute value of the Jacobian we have

$$f(p_k, q_k | \gamma, N) = M_k \left[ \sqrt{2(k-1)} \frac{p_k \gamma}{2} + (k-1) \right]^{(k-1)/2} \exp \left[ -\gamma \left( \sqrt{2(k-1)p_k - (k - (k-1)/2)} \right) \right] \exp \left[ -\frac{q_k^2 \gamma}{2} \right]$$

$$= 0 \quad \text{otherwise}$$

(4.101)

The normalizing constant $M_k$ is given by

$$M_k = \left( \frac{\gamma^{\frac{1}{2}} \sqrt{2(k-1)}}{\sqrt{2\pi \Gamma \left( \frac{k}{2} - \frac{1}{2} \right)}} \right)$$

It is evident that this latter joint distribution function is factorable into two terms, one involving only $p_k$ and one involving only $q_k$. Hence, we conclude that $p_k$ and $q_k$ are conditionally independent. The distribution of $\sqrt{2(k-1)} \frac{p_k \gamma}{2} + (k-1)$ is Chi-squared with $(k-1)$ degrees of freedom while the distribution of $q_k$ is normal with mean zero and variance $\gamma^{-1}$. Using the same methods as above it is easy to show that the distributions of $p_k$ and $q_k$ conditional to $\gamma$ and $N$ are of the same class as those above, with a change in the mean values. This as $k \to \infty$ the distributions of $q_k$ conditional to both $N$ and $N$ remain the same since they are independent of $k$. On the other hand, the results of Cramér
show that the Chi-squared distribution with n degrees of freedom approaches normality with mean n and variance 2n as \( n \to \infty \) (Ref. 23). Thus the distributions of \( p_k \) conditional to SN and N approach normality with mean and variance independent of \( k \) as \( k \to \infty \).

Let us now return to \( \ln \ell \) expressed in terms of \( p_k \) and \( q_k \) as given by Eq. 4.100. We recall from Eq. 4.26 that

\[
b_k = b + k/2
\]  

(4.26)

Substituting this expression for \( b_k \) into Eq. 4.100 we have

\[
\ln \ell = -(k/2 + b) \ln \left[ 1 + \frac{\gamma (E_s - 2 \sqrt{E_s} q_k)}{(k-1) \left(1 + \frac{\sqrt{2}}{\sqrt{k-1}} p_k + \frac{\gamma q_k^2}{k-1} + \frac{2cY}{k-1}\right)} \right]
\]  

(4.102)

As \( k \to \infty \) the distributions of \( p_k \) and \( q_k \) approach normality with parameters independent of \( k \). Thus, for fixed values from these distributions we have

\[
\lim_{k \to \infty} \ln \ell = \left(\frac{k}{2} \ln \left[ 1 + \frac{\gamma (E_s - 2 \sqrt{E_s} q_k)}{k} \right]\right)
\]  

(4.103)

But from Cramér (Ref. 23) we have the result that

\[
\lim_{k \to \infty} \left\{ n \ln \left[ 1 + x/n \right] \right\} = x
\]

Thus Eq. 4.103 becomes

\[
\lim_{k \to \infty} \ln \ell = -\gamma [E_s - 2 \sqrt{E_s} q_k]
\]  

(4.104)
Since, as discussed above, $q_k$ conditional to $\gamma$ and both $N$ and $SN$ is normally distributed, it follows from Eq. 4.104 that the distributions of $\ln f$ conditional to $\gamma$ and both $SN$ and $N$ are normal as $k \to \infty$. This implies that the corresponding conditional ROC curves approach normality as $k \to \infty$. This result is exactly what we obtained from the computer simulation results.

4.5. Signals of Unknown Amplitude in Noise of Unknown Level

In this section the detection of a signal known except for amplitude in added Gaussian noise of unknown level is considered (SKEA-NUL). The additional uncertainty concerning signal amplitude implies that detection performance of the optimum receiver for this case can never exceed that of the receiver obtained in the previous section. Thus the work of the previous section serves to provide an upper performance bound for the example considered here.

4.5.1. Problem Formulation. The functional form of the amplitude distribution chosen was that of a uniform distribution with end points denoted by $\theta_1$ and $\theta_2$. This distribution was chosen because it provides sufficient practicality to the example and at the same time allows for adequate mathematical tractability. For, although the basic concepts in detecting an unknown amplitude signal do not depend on the particular form of the amplitude distribution, the detailed mathematics of the solution do depend on the amplitude distribution.

4.5.2. Receiver Design. The design of the optimum receiver for the SKEA-NUL application considered in this section is based on the realization of the likelihood ratio or some monotone function of it. In Section 4.3 it was shown that this likelihood ratio could be
realized, in general, in a sequential manner and the results of this derivation were presented in Table 4.1. In the previous section these results were applied in detail to a specific application. Therefore, since the design of the optimum receiver for the application of this section is similar to the one already presented, the details are omitted here. Only the results are presented with the interested reader being referred to Appendix D for the details. The primary emphasis is placed on the performance of the optimum receiver to determine the effect of both signal and noise uncertainties on the detectability of signals mixed with noise.

The design equations for the optimum receiver are presented below with a receiver block diagram based on an implementation of these equations presented in Fig. 4.48. The notation is consistent with that already developed in Section 4.1 with the addition of the random signal amplitude which is denoted by $\theta$.

Classification Outputs:

$$g_k(\theta, \gamma | SN) = A_k \gamma^k \exp \left[ -(c_k - 2\theta h_k + \theta^2 E_k)\gamma/2 \right]$$

$$g_k(\gamma | N) = B_k \gamma^k \exp \left[ -c_k \gamma/2 \right]$$

where

$$A_k = \left[ \frac{b_k+1}{2} \int_0^\theta \frac{\Gamma(b_k+1)}{\Gamma(b_k+1)} \left[ c_k - 2\theta h_k + \theta^2 E_k \right]^{-\frac{(b_k+1)}{2}} d\theta \right]^{-1}$$

$$B_k = \left[ \frac{c_k}{2} \right]^{b_k+1} \frac{1}{\Gamma(b_k+1)}$$

128
Fig. 4.38. Optimum receiver realization for signal of uncertain amplitude, noise of uncertain level
and

\[ b_k = b_{k-1} + \frac{1}{2} \quad \quad \quad E_k = E_{k-1} + s_k^2 \]

\[ c_k = c_{k-1} + x_k^2 \quad \quad \quad h_k = h_{k-1} + x_k s_k \]

Detection Output:

\[ f(x_k) = f(x_k | X_{k-1}) f(X_{k-1}) \]

\[ f(x_k | X_{k-1}) = \begin{bmatrix} B_{k-1} & A_{k-1} \\ A_k & B_k \end{bmatrix} \]

4.5.3. Receiver Operation. The operation of the optimum receiver in Fig. 4.48 is basically the same as the receiver developed in the previous section with the exception of the additional computations which must be performed to account for the uncertainty concerning the signal amplitude. The effect of this additional uncertainty on receiver operation can be investigated by again determining the constant log-likelihood ratio loci. It was found that these contours of constant log-likelihood ratio could not be determined analytically and so the IBM 7090 digital computer was utilized. The results of the computation are shown in Figs. 4.39 and 4.40. In these diagrams (and elsewhere in this section) \( \nu_\theta \) denotes the variance of the uniformly distributed signal amplitude.

Reference to Figs. 4.39 and 4.40 indicates that the additional uncertainty concerning the signal amplitude has modified the circles which we previously obtained for the certain signal case. This additional uncertainty has caused the circles to become somewhat elliptical or egg-shaped. The AGC action which we noted in the SKE-NUL case is still present as evidenced by the closed contours. However, the signal
Fig. 4.39. Typical contours of constant log-likelihood ratio for uncertain signal amplitude and uncertain noise level with parameters $\nu_\theta = 1.33$, $c = 1.0$
Fig. 4.40. Typical contours of constant log-likelihood ratio for uncertain signal amplitude and uncertain noise level with parameters $v_{\theta} = 3.0$, $c = 1.0$.
uncertainty has the effect of "smearing out" the circles which were obtained previously. Comparison of the diagrams indicates that as the uncertainty concerning the signal amplitude increases there is a corresponding increase in the relative size of the contours. This result is opposite to that of the SKE-NUL case. For that case as the noise became less certain the corresponding contours became smaller. The explanation is that as the signal amplitude becomes less certain signals of higher energy are given more weight in the signal ensemble. The receiver as a consequence considers observations with a greater total energy. This effect is achieved by an increase in the size of the contours.

In summary, the operation of the optimum receiver in the face of additional uncertainty concerning the signal amplitude is basically a modification of the certain signal case. The performance is, of course, correspondingly lower. For a more quantitative evaluation of these effects we now turn to an evaluation of the optimum receiver performance.

4.5.4. Receiver Evaluation. The evaluation of the performance of the optimum receiver for the signal known except for amplitude in noise of unknown level (SKEA-NUL) case is presented in the next two sections. A brief review of the techniques involved in evaluating receiver performance was given in Section 4.4.4 and will not be discussed in detail here. In the performance evaluation the following points are emphasized.

1. Performance of the optimum receiver for varying degrees of amplitude and noise uncertainty

2. Comparison of the performance of the optimum receiver with a suboptimum receiver

3. Performance of the sequential receiver realization of the optimum receiver and its "adaptive" characteristics

133
As in the previous example, the input signal waveshape used for evaluation was specified as a dc signal with the exception being the unknown signal amplitude. The random signal amplitude was denoted by \( \theta \) with the total normalized signal energy denoted by \( E_s \) as before. The ROC curves presented in the next section were obtained for 100 observations (i.e., \( k = 100 \)) utilizing a Monte Carlo technique. This approach, although experimental, was checked with some analytical results which were able to be obtained for the case of two samples and was found to have an extremely high degree of accuracy. Basically the technique consisted of simulating the optimum receiver on the computer and then approximating the density functions of the observations under noise alone using 500 (equal probability) points. These density functions were the same as those of the SKE-NUL example since the amplitude uncertainty does not affect them. The distribution of the likelihood ratio under noise alone was then obtained using values from the density functions in a Monte Carlo fashion. The distribution of the likelihood ratio conditioned on the hypothesis SN was obtained from the distribution conditioned on the hypothesis N by the likelihood theorem stated in Section 4.4.4. The ROC curves were obtained from these two distributions.

To investigate the sequential operation of the receiver the detection situation was simulated on the computer and a number of sample runs were made. ROC performance curves and the classification outputs were obtained as a function of the number of observations and are presented in the following sections.

4.5.4.1. Optimum Receiver Performance. The ROC curves obtained for the SKEA-NUL case are presented in Figs. 4.41 through 4.49. The curves are parameterized by the variance of the
signal amplitude distribution \( v_\theta \), the mean and variance of the
constance distribution \( m \) and \( v \), and the normalized signal energy \( E_s \).
The ROC curves are approximately normal with the most significant
departure from normality occurring at larger values of uncertainty
in the noise level (e.g., Fig. 4.49). For a more quantitative evaluation
of the ROC's the equivalent detectability \( d_e \) (measured along the negative
diagonal) was determined for each of the curves and plotted as a function
of the amplitude uncertainty. These graphs are presented in Figs. 4.45
and 4.50 for the two values of signal energy \( E_s \). Referring to these
figures it is evident that for small values of initial noise uncertainty, \( v \),
the amplitude uncertainty, \( v_\theta \), has a strong effect on detectability. As the
value of \( v \) increases however, the effect of \( v_\theta \) becomes almost negligible.
This result is predictable because of the direct dependence of the signal
to noise ratio on the square of the signal amplitude and the constance
level (see Section 4.4.5.1). An uncertainty in either of them causes
an uncertainty in the SNR, and thus at low values of \( v \), the effect of
\( v_\theta \) predominates. Perhaps the most important effect, however, is that
for relatively large values of noise uncertainty the effect of uncertainty
in the signal amplitude is very small. Therefore, for the case of large
uncertainties in both the signal amplitude and the noise level, the perform-
ance is approximately given by the SKEA-NUL case previously determined.

4.5.4.2. Comparison with Suboptimum Receiver. In this
section the performance of the optimum receiver of Section 4.5.2 is
compared with the performance of a receiver which is suboptimum for the
SKEA-NUL case. The use of this type of comparison allows us to deter-
mine the amount of performance increase which is attainable utilizing
the optimum receiver. The suboptimum receiver is a likelihood
Fig. 4.41. The ROC curves for the optimum receiver, SKEA-NUL case, with parameters $E_S = 1.0$, $m = 1.0$, $v = 1.0$, $v_\theta = 0.0, 1.33, 2.08, 2.80, 5.33$
Fig. 4.42. The ROC curves for the optimum receiver, SKEA-NUL case, with parameters $E_s = 1.0$, $m = 1.0$, $v = 2.0$, $v_\theta = 0.0, 1.33, 2.08, 2.80, 5.33$
Fig. 4.43. The ROC curves for the optimum receiver, SKEA-NUL case, with parameters $E_s = 1.0$, $m = 1.0$, $v = 4.0$, $v_\theta = 0.0, 1.33, 2.08, 5.33$
Fig. 4.44. The ROC curves for the optimum receiver, SKEA-NUL case, with parameters $E_s = 1.0$, $m = 1.0$, $\nu = 8.0$, $v_\theta = 0.0$, 1.33, 2.80, 5.33.
Fig. 4.45. Equivalent detectability versus $v_\theta$, SKEA-NUL case, with parameters $E_s = 1.0$, $m = 1.0$, $v = 0.5, 1.0, 2.0, 4.0, 8.0$.
Fig. 4.46. The ROC curves for the optimum receiver, SKEA-NUL case, with parameters $E_s = 2.0$, $m = 1.0$, $v = 0.5$, $v_\theta = 0.0$, $1.33$, $2.08$, $2.80$, $5.33$.
Fig. 4.47. The ROC curves for the optimum receiver, SKEA-NUL case, with parameters $E_s = 2.0$, $m = 1.0$, $v = 1.0$, $v_\theta = 0.0$, $1.33$, $2.08$, $2.80$, $5.33$. $k = 100$
Fig. 4.48. The ROC curves for the optimum receiver, SKEA-NUL case, with parameters $E_s = 2.0$, $m = 1.0$, $v = 2.0$, $v_\theta = 0.0, 1.33$, $2.08, 2.80, 5.33$
Fig. 4.49. The ROC curves for the optimum receiver, SKEA-NUL case, with parameters $E_s = 2.0$, $m = 1.0$, $v = 4.0$, $v_\theta = 0.0$, 1.33, 2.08, 2.80, 5.33.
Fig. 4.50. Equivalent detectability versus $v_\theta$, SKEA-NUL case, with parameters $E_s = 2.0$, $m = 1.0$, $v = 0.5, 1.0, 2.0, 4.0$, $k = 100$. 
correlator which correlates the received observation with a normalized reference value of the input signal and then computes the likelihood ratio of this quantity as an average over the possible amplitude values which could occur. In effect the receiver correlates a normalized reference signal with the ensemble of possible signals. This type of receiver was discussed in Section 4.4.5.2.

The ROC performance curves for the correlator were obtained using the methods described in Appendix C and are presented in Figs. 4.51 through 4.53. The appropriate optimum receiver ROC curves for the SKEA-NUL case have been included in these figures for comparison purposes. As in Section 4.5.2 we note that the detectability of both the optimum and suboptimum receivers is again identical along the negative diagonal. In addition, as one moves away from the negative diagonal the differential in performance between the optimum and suboptimum receivers increases. At low false alarm probabilities, this differential becomes quite large and the advantage of the optimum receiver is apparent. Because of the limiting nature of the optimum receiver curves, it is also evident that as the uncertainty in the signal amplitude increases the differential in performance at low false alarm probabilities increases faster than a linear relationship would predict.

4.5.4.3. Sequential Operation. The results of the computer simulation runs illustrating the classification outputs for the SKEA-NUL case are presented in Figs. 4.54 and 4.55. In the first column of these figures, the classification output consisting of the probability density function of the noise constance conditional to the particular set of observations up to time $t_k$ (indexed by k) and to the noise hypothesis is presented. This output represents the a posteriori density of the noise
Fig. 4.51. Comparison of receiver performance, SKEA-NUL case, with parameters $E_s = 1.0, m = 1.0, v = 1.0, v_\theta = 0.0, 1.33, 2.80, 5.33$
Fig. 4.52. Comparison of receiver performance, SKEA-NUL case, with parameters $E_s = 1.0$, $m = 1.0$, $v = 4.0$, $v_\theta = 0.0$, 1.33, 2.80, 5.33

$k = 100$
Fig. 4.53. Comparison of receiver performance, SKEA-NUL case, with parameters $E_s = 1.0$, $m = 1.0$, $v = 8.0$, $v_\theta = 0.0$, $1.33$, $2.80$, $5.33$. 

$k = 100$
Simulation run no. 1, SKEA-NUL case, parameters \( m = 1.0, v = 2.0, v_e = 2.08, e = 3.0, \gamma = 1.0 \)
Simulation run no. 2, SKEA-NUL case, parameters $m = 1.0$, $v = 2.0$, $v_e = 2.83$, $e = 3.0$, $\gamma = 1.0$

Fig. 4.54. Simulation runs, SKEA-NUL case
Simulation run no. 3, SKEA-NUL case, parameters $m = 1.0$, $v = 4.0$, $v_c = 2.08$, $e = 3.0$, $\gamma = 1.0$
Simulation run no. 4, SKEA-NUL case, parameters \( m = 1.0, v = 4.0, v_e = 2.83, e = 3.0, \gamma = 1.0 \)

Fig. 4.55. Simulation runs, SKEA-NUL case

153
constance as a function of time. It is conditional to the observation set and the hypothesis that only noise is present at the input to the receiver.

In the second column, the probability density of the noise constance conditional to the observation set and the signal mixed with noise hypothesis are given. This classification output represents the a posteriori density of the noise constance as a function of time and the observation set. It is conditional to the hypothesis that signal and noise are present at the receiver input.

In the final column the sequential detection output is presented. This output is the value of the likelihood ratio at each time, $t_k$, for the particular set of observations obtained up to that time.

In sample Runs 1 and 2 the receiver input consisted of signal plus noise and in Runs 3 and 4 the input consisted of noise alone. Two different sets of parameter values were used for the a priori constance density function as were two sets of values for the uniform a priori signal amplitude distribution. The values of the parameters are indicated on the figures. For all of the runs the true signal to noise ratio after 81 observations was 8.63 db. The true signal amplitude was maintained at $\theta = 3.0$ and the true constance value at $\gamma = 1.0$.

From the results of the trial runs the sequential operation of the receiver is apparent. As observation time increases, the proper classification outputs change with time in such a manner that they approach the true values of the unknown parameters. Because of the noise fluctuations from run to run, the time of build-up varies; however, it occurs after a sufficient number of observations. The term "proper" is used above since it must be noted that the various classification outputs are conditional to one of the two hypotheses. The choice of the proper
one is governed by the value of the detection output, the likelihood ratio, presented in the final column. It is interesting to note that in sample Runs 3 and 4 for which the input was noise, the conditional amplitude density approaches a value of zero as its mean value. Since the signal amplitude is zero for this case, this is a pleasing result.

The conditional ROC curves as a function of observation time are presented in Figs. 4.56 through 4.59 for selected sample runs. As in the SKE-NUL example previously considered, conditional performance of the receiver asymptotically approaches the upper performance bound determined by the absence of signal and noise uncertainty. This approach to the upper performance is again evident by the approach of the ROC curves to straight lines (i.e., normal ROC curves). Thus, from the receiver simulation, the adaptive nature of the receiver becomes evident. As the receiver "learns" the true values of the unknown parameters the performance approaches that value which would occur in the absence of uncertainty concerning the parameters.
Fig. 4.56. Conditional ROC curves as a function of observation time, SKEA-NUL case, with parameters $m = 1$, $v = 2$ and $\theta_2 = 5.5$, $\theta_1 = 0.5$
Fig. 4.57. Conditional ROC curves as a function of observation time, SKEA-NUL case, with parameters $m = 1$, $v = 2$, and $	heta_2 = 5.5$, $\theta_1 = 0.1$
Fig. 4.58. Conditional ROC curves as a function of observation time, SKEA-NUL case, with parameters $m = 1$, $v = 4$ and $\theta_2 = 5.5$, $\theta_1 = 0.5$.
Fig. 4.59. Conditional ROC curves as a function of observation time, SKEA-NUL case, with parameters $m = 1$, $v = 4$ and $\theta_2 = 5.5$, $\theta_1 = 0.1$
CHAPTER V

TIME-VARYING UNCERTAINTIES

5.1. Introduction

In this chapter the detection of a known signal added to a background noise of an uncertain and time-varying nature is considered. The fact that the background noise is assumed to be of an uncertain nature leads directly to a Double Composite Hypothesis problem. A review of the literature indicates that this type of problem has not been considered before.

The motivation for an investigation of this type of situation and its influence on the ability of optimum receivers to detect signals in noise arises naturally from many physical situations, since, in practice, many of the background interference phenomena which are encountered are of a time-varying nature. For example, when an array of hydrophones is used to locate a source of acoustic noise some of the background noise might arise from a number of localized sources whose intensity fluctuates during the observation interval. In particular, if such an array is utilized in an underwater application, the background noise consists of the ambient biological, shipping and surface generated noise in the ocean, known to exhibit large fluctuations in magnitude from moment to moment. In the work below a general Double Composite Hypothesis problem is first formulated with specific applications following.

5.2. Problem Statement and Notation

As pointed out above, the occurrence of a background environment with time-varying uncertainties during the observation interval arises
in many (if not all) detection situations. This may be especially true in unspecified response time detection situations where signal-to-noise ratios are small and observation times are made quite large in an effort to obtain increased performance. Up to the present time, however, very little attention has been given to detection problems of this type. Thomas and Williams (Ref. 10) have published some work in this general area, as indicated in Chapter II; however, they made no effort to consider optimum detection which is the major purpose of the work presented here.

Let us now consider the formulation of the class of detection problems considered in this chapter and the notation that is used.

It is assumed that the signal whose presence or absence is to be detected is known exactly at the receiver and that when it is present, it is added to the noise to form the inputs to the receiver. These receiver inputs, which are functions of time, are assumed to be defined for all times in the observation interval, \(0 \leq t \leq T\). They are further assumed to be limited to a band of frequencies of width \(W\) and, therefore, by the sampling theorem, may be thought of as points in a \(2WT\) dimensional space, the coordinates of a point being the values of the functions at the samples points \(t_i = i/2WT\), for \(1 \leq i \leq 2WT\). The notation \(X_k\) denotes a receiver input, \((x_1, x_2, \ldots, x_k)\) where \(k = 2WT\) and \(x_j\) denotes the \(j\)th sample value or coordinate. This notation is consistent with that previously used.

Under the condition that noise alone is present it is assumed that the conditional joint density of the observation \(X_k\) (that is, the density of \(X_k\) considering all parameters of the noise process to be known) is conditionally independent and, in addition, that the conditional densities of the individual sample values, \(x_i\), are dependent on only one unknown parameter. With these assumptions it follows that Eq. 3.2 can be expressed as
\[ f(X_k | n, N) = f(x_n | \gamma_k, N) f(X_{k-1} | \gamma_{k-1}, \ldots, \gamma_1, N) \]  

(5.1)

where \( \gamma \) denotes the vector whose components are the values of the noise parameter at the sampling times, i.e., \( (\gamma_1, \gamma_2, \ldots, \gamma_k) \). Note that for the time-varying situation the values of the components are not in general the same. It is evident that the properties of conditional independence and one parameter densities are also true when the receiver input consists of signal and noise; since the signal is known exactly and is added to the noise, the SN density is similar to the N density with only a translation of the mean.

In order to describe the time-varying nature of the noise process a general class of a priori distributions is assumed for the noise parameter process. The restriction which is imposed on this class of a priori distributions functions is that they are first order Markov processes. Mathematically this means that

\[ g(\gamma_k | \gamma_{k-1}, \ldots, \gamma_1) = g(\gamma_k | \gamma_{k-1}) \]  

(5.2)

for all \( k \). The choice of a first order Markov process was made because (1) it is a reasonable mathematical approximation to actual slowly varying physical processes and (2) it results in a process which is mathematically tractable.

With the discussion given above the formulation of the general detection problem to be investigated in this chapter is complete. In the next section a sequential realization of the optimum receiver is derived.

5.3. General Receiver Realization

Optimum detector design is based on the likelihood ratio of the received observation \( X_k \). In Chapter III it was shown that for a general
Double Composite Hypothesis detection problem this likelihood ratio could be realized in a sequential manner. In addition, the fact that the sequential realization results in a limited memory receiver for a general detection problem with time-varying parameters was discussed. In this section the results of Chapter III are applied to the Double Composite Hypothesis detection problem formulated above and an optimum sequential realization resulting in limited memory receivers is derived. In the next section these results are applied to a more specific example.

5.3.1. Processing Equation. To begin, the results of Chapter III are repeated below. As usual, only those equations conditioned on the SN hypothesis will be utilized with the results for the equations conditioned on the hypothesis N following in a similar manner. The assumption of conditional independence is also used.

\[
\ell(X_k) = \ell(x_k | X_{k-1}) \ell(X_{k-1})
\]  
(3.9)

\[
\ell(x_k | X_{k-1}) = \frac{f(x_k | X_{k-1}, SN)}{f(x_k | X_{k-1}, N)}
\]  
(3.8)

\[
f(x_k | X_{k-1}, SN) = \int \int f(x_k | s, n, SN) g_{k-1}(s, n | SN) \, ds \, dn
\]  
(3.5)

\[
g_k(s, n | SN) = \frac{f(x_k | s, n, SN)}{f(x_k | X_{k-1}, SN)} g_{k-1}(s, n | SN)
\]  
(3.12)

The problem formulated above and under consideration in this section involves only uncertainties in the noise process and not in the signal process. Therefore, the dependence on signal process uncertainties may be suppressed in Eqs. 3.5 and 3.12 and these equations may be rewritten
as

\[ f(x_k | X_{k-1}, SN) = \int_N f(x_k | n, SN) \, g_{k-1} (n | SN) \, dn \]  

(5.3)

and

\[ g_k (n | SN) = \frac{f(x_k | n, SN)}{f(x_k | X_{k-1}, SN)} \, g_{k-1} (n | SN) \]  

(5.4)

In the previous section, we made the assumption that the noise process was dependent only on one parameter which was denoted by the vector \( \gamma \). If we use the additional assumption that the distribution of \( x_k \) depends on only the \( k \)-th component of \( \gamma \), \( \gamma_k \), then Eqs. 5.3 and 5.4 may be rewritten as

\[ f(x_k | X_{k-1}, SN) = \int_{\Gamma} f(x_k | \gamma_k, SN) \, g_{k-1} (\gamma | SN) \, d\gamma \]  

(5.5)

and

\[ g_k (\gamma | SN) = \frac{f(x_k | \gamma_k, SN)}{f(x_k | X_{k-1}, SN)} \, g_{k-1} (\gamma | SN) \]  

(5.6)

The joint probability density, \( g_{k-1} (\gamma | SN) \), appearing in Eq. 5.6 may be expressed explicitly as a function of the vector \( \gamma \) to yield

\[ g_{k-1} (\gamma | SN) = g_{k-1} (\gamma_k, \gamma_{k-1}, \ldots, \gamma_1 | SN) \]  

(5.7)

Furthermore, it follows from the laws of joint probability densities that this joint density function may also be expressed in the following conditional form.

\[ g_{k-1} (\gamma | SN) = g_{k-1} (\gamma_k | \gamma_{k-1}, \ldots, \gamma_1, SN) \, g_{k-1} (\gamma_{k-1}, \ldots, \gamma_1 | SN) \]  

(5.8)

The conditional joint probability density function of the right-hand side
of Eq. 5.8, \( g_{k-1}(\gamma_k | \gamma_{k-1}, \ldots, \gamma_1, \text{SN}) \), gives the conditional probability density of the kth value of \( \gamma \) after \( k-1 \) observations have been taken, given exact knowledge of the first \( k-1 \) values of \( \gamma \). However, from the definition of the joint probability density \( g_{k-1}(\gamma | \text{SN}) \) and the first order Markov property of \( \gamma \) discussed above, it follows that this conditional joint probability density depends only on the \((k-1)\)st value of \( \gamma \) and is independent of the number of observations taken. In addition, this density is also independent of the hypothesis \( \text{SN} \) since the noise process is assumed to be the same for both hypotheses. Thus Eq. 5.8 may be written as

\[
g_{k-1}(\gamma | \text{SN}) = g(\gamma_k | \gamma_{k-1}) g_{k-1}(\gamma_{k-1}, \ldots, \gamma_1 | \text{SN})
\]

(5.9)

This new expression for the joint probability density of \( \gamma \) as given by Eq. 5.9 may now be substituted into Eq. 5.5 to yield

\[
f(x_k | X_{k-1}, \text{SN}) = \int f(x_k | \gamma_k, \text{SN}) g(\gamma_k | \gamma_{k-1}) g_{k-1}(\gamma_{k-1}, \ldots, \gamma_1 | \text{SN}) \, d\gamma
\]

(5.10)

The integration in the above equation is implicitly defined over the entire noise parameter space. If this implicit description is now expressed explicitly, then Eq. 5.10 can be written as

\[
f(x_k | X_{k-1}, \text{SN}) = \int \int \ldots \int f(x_k | \gamma_k, \text{SN}) g(\gamma_k | \gamma_{k-1}) g_{k-1}(\gamma_{k-1}, \ldots, \gamma_1 | \text{SN}) \cdot d\gamma_1, \ldots, d\gamma_{k-1}, d\gamma_k
\]

(5.11)

It is evident that portions of the integrand of Eq. 5.11 are independent of portions of the region of integration. If these quantities are removed from under the appropriate integrals, Eq. 5.11 becomes
\[ f(x_k | X_{k-1}, \text{SN}) = \int f(x_k | \gamma_k, \text{SN}) \int g(\gamma_k | \gamma_{k-1}) \int \cdots \int g_{k-1}(\gamma_{k-1}, \ldots, \gamma_1 | \text{SN}) \cdot d\gamma_1, \ldots, d\gamma_{k-1} d\gamma_k \]  

(5.12)

The first \( k-2 \) integrations in Eq. 5.12 involve only the joint probability density \( g_{k-1}(\gamma_{k-1}, \ldots, \gamma_1 | \text{SN}) \). Furthermore, since this is a probability density, it follows that integration over this region leaves only a function of \( \gamma_{k-1} \). In other words

\[ g_{k-1}(\gamma_{k-1} | \text{SN}) = \int \cdots \int g_{k-1}(\gamma_{k-2}, \ldots, \gamma_1 | \text{SN}) d\gamma_1 \cdots d\gamma_{k-2} \]  

(5.13)

If Eq. 5.13 is substituted into Eq. 5.12 this latter equation reduces to

\[ f(x_k | X_{k-1}, \text{SN}) = \int f(x_k | \gamma_k, \text{SN}) \left[ \int g(\gamma_k | \gamma_{k-1}) g_{k-1}(\gamma_{k-1} | \text{SN}) d\gamma_{k-1} \right] d\gamma_k \]  

(5.14)

This expression in brackets in Eq. 5.14 is evidently a probability density itself and is a function of only \( \gamma_k \). Therefore it follows from previous definitions that

\[ g_{k-1}(\gamma_k | \text{SN}) = \int g(\gamma_k | \gamma_{k-1}) g_{k-1}(\gamma_{k-1} | \text{SN}) d\gamma_{k-1} \]  

(5.15)

If Eq. 5.15 is now substituted into Eq. 5.14, the following result (which could also have been obtained directly from Eq. 5.5) is obtained for \( f(x_k | X_{k-1}, \text{SN}) \).

\[ f(x_k | X_{k-1}, \text{SN}) = \int f(x_k | \gamma_k, \text{SN}) g_{k-1}(\gamma_k | \text{SN}) d\gamma_k \]  

(5.16)

These latter two equations (5.15 and 5.16) represent a partial fulfillment of the original goal of obtaining a sequential realization for the
time-varying detection problem at hand. Equation 5.16 enables us to obtain the conditional likelihood ratio given the past observations $X_{k-1}$ in terms of the observation $x_k$ and an updated probability density function for $\gamma_k$. Since a similar expression also follows under the hypothesis N, Eq. 5.15 gives a partial rule for the updating of the conditional probability density functions of $\gamma_k$. However, a procedure must now be determined for obtaining $g_k(\gamma_k|SN)$ from $g_{k-1}(\gamma_k|SN)$ and the kth observation value $x_k$.

To obtain this updating procedure let us first rewrite Eq. 5.6 as

$$g_k(\gamma|SN) = \frac{f(x_k | \gamma_k, SN)}{f(x_k | X_{k-1}, SN)} g_{k-1}(\gamma|SN) \quad (5.6)$$

If we now substitute for $g_{k-1}(\gamma|SN)$ in this latter expression utilizing Eq. 5.9, the result is

$$g_k(\gamma|SN) = \frac{f(x_k | \gamma_k, SN)}{f(x_k | X_{k-1}, SN)} g(\gamma_k | \gamma_{k-1}) g_{k-1}(\gamma_{k-1}, \ldots, \gamma_1|SN) \quad (5.17)$$

Equation 5.17 gives an expression for the joint probability density of $\gamma$ after $k$ observations have been taken. Since it is a joint probability density function, integration may be performed over a portion of the $\gamma$-space to obtain the probability density function for $\gamma_k$ after $k$ observations have been taken. Performing this integration leaves just $g_k(\gamma_k)$ on the left hand side of Eq. 5.17 and therefore, after integration, Eq. 5.17 becomes

$$g_k(\gamma_k) = \left[ \frac{f(x_k | \gamma_k, SN)}{f(x_k | X_{k-1}, SN)} \right] \int \int \cdots \int g(\gamma_k | \gamma_{k-1}) g_{k-1}(\gamma_{k-1}, \ldots, \gamma_1|SN) \\gamma_{k-1} \gamma_{k-2} \ldots \gamma_1 \quad (5.18)$$

167
where the term in brackets has been taken outside of the integration since it is independent of the variables of integration.

The first term in the integrand of Eq. 5.18 is independent of the first k-2 integrations and may be taken outside of these integrals. The first k-2 integrals therefore act only on the second term of the integrand $g_{k-1}(\gamma_{k-1}, \ldots, \gamma_1)$. However, from the definition of marginal conditional probabilities

$$g_{k-1}(\gamma_{k-1} \mid SN) = \int_{\gamma_{k-2}} \cdots \int_{\gamma_1} g_{k-1}(\gamma_{k-1}, \ldots, \gamma_1 \mid SN) \, d\gamma_1 \cdots d\gamma_{k-2}$$

(5.19)

and therefore Eq. 5.8 may be reduced to the following simplified form.

$$g_k(\gamma_k \mid SN) = \frac{f(x_k \mid \gamma_k, SN)}{f(x_k \mid X_{k-1}, SN)} \int_{\gamma_{k-1}} g(\gamma_k \mid \gamma_{k-1}) g_{k-1}(\gamma_{k-1} \mid SN) \, d\gamma_{k-1}$$

(5.20)

The integral term in Eq. 5.20 has appeared before. Inspection of Eq. 5.15 reveals that it is equal to $g_{k-1}(\gamma_k \mid SN)$ and therefore, making this substitution Eq. 5.20 becomes

$$g_k(\gamma_k \mid SN) = \frac{f(x_k \mid \gamma_k, SN)}{f(x_k \mid X_{k-1}, SN)} \frac{f(x_k \mid X_{k-1}, SN) g_{k-1}(\gamma_k \mid SN)}{f(x_k \mid X_{k-1}, SN) g_{k-1}(\gamma_k \mid SN)}$$

(5.21)

This last equation completes the sequential realization of the likelihood ratio for the time-varying Double Composite Hypothesis detection problem under consideration. The pertinent receiver equations are summarized in Table 5.1 where the required equations under the N alone hypothesis have been included although they were not derived in this section. As has been pointed out, they can be obtained in a manner completely analogous to the derivation above. A particular receiver realization based on implementing these equations is also presented and appears in Fig. 5.1.
Sequential Likelihood Ratio

\[ f(X_k) = f(x_k | X_{k-1}) f(X_{k-1}) \]  \hspace{1cm} (5.22)

Conditional Likelihood Ratio

\[ f(x_k | X_{k-1}) = \frac{f(x_k | X_{k-1}'SN)}{f(x_k | X_{k-1}'N)} \]  \hspace{1cm} (5.23)

Conditional Observation Densities

\[ f(x_k | X_{k-1}, SN) = \int_{\gamma_k} f(x_k | \gamma_k, SN) g_{k-1} (\gamma_k | SN) \, d\gamma_k \]  \hspace{1cm} (5.24)

\[ f(x_k | X_{k-1}, N) = \int_{\gamma_k} f(x_k | \gamma_k, N) g_{k-1} (\gamma_k | N) \, d\gamma_k \]  \hspace{1cm} (5.25)

Updating Equations Under SN

\[ g_k (\gamma_k | SN) = \int_{\gamma_{k-1}} g(\gamma_k | \gamma_{k-1}') g_{k-1} (\gamma_{k-1} | SN) \, d\gamma_{k-1} \]  \hspace{1cm} (5.26)

\[ g_k (\gamma_k | SN) = \frac{f(x_k | \gamma_k, SN)}{f(x_k | X_{k-1}, SN)} \, g_{k-1} (\gamma_k | SN) \]  \hspace{1cm} (5.27)

Updating Equations Under N

\[ g_{k-1} (\gamma_k | N) = \int_{\gamma_{k-1}} g(\gamma_k | \gamma_{k-1}') g_{k-1} (\gamma_{k-1} | N) \, d\gamma_{k-1} \]  \hspace{1cm} (5.28)

\[ g_k (\gamma_k | N) = \frac{f(x_k | \gamma_k, N)}{f(x_k | X_{k-1}, N)} \, g_{k-1} (\gamma_k | N) \]  \hspace{1cm} (5.29)

Table 5.1 Sequential receiver design equations - time varying case
Fig. 5.1. General sequential optimum receiver realization for slowly varying noise process.
5.3.2. Receiver Operation. The operation of the sequential receiver and its adaptive nature are best understood by referring to Fig. 5.1, where the basic receiver operations are displayed. To discuss receiver operation we assume that k-1 observations have been taken. At this point, the receiver has stored the likelihood ratio of the k-1 observations, \( \ell(X_{k-1}) \), and the a priori probability densities of \( \gamma_k \) under SN and N conditioned on exact knowledge of the k-1 observations, \( g_{k-1}(\gamma_k | SN) \) and \( g_{k-1}(\gamma_k | N) \). In addition, the conditional or "transfer" probability density \( g(\gamma_k | \gamma_{k-1}) \) is stored. It should be remembered here that this latter conditional probability density is a direct manifestation of the physical process which is under consideration and is not a function of the observation samples obtained. (In a more general case it may be a function of time, and if so, would be stored as such in the receiver memory. However, in many practical situations this density may be considered as independent of time, thus providing a physical description of the transition from one component of \( \gamma \) to the next component which is the same for all values of time.)

Upon reception of the kth sample of the observation, the receiver computes the factors \( f(x_k | \gamma_k, SN) \) and \( f(x_k | \gamma_k, N) \) and then utilizes these functions to compute \( f(x_k | X_{k-1}, SN) \) and \( f(x_k | X_{k-1}, N) \) by Eqs. 5.24 and 5.25. From these latter two numbers the conditional likelihood ratio is computed as their ratio (Eq. 5.22). This conditional likelihood ratio is then multiplied by the stored value of the likelihood ratio after k-1 observations to obtain the new updated ratio after k observations.

At the same time as the above operations are performed the receiver updates its opinion of \( \gamma_k \) (conditioned on the two hypotheses) by using the k-th observation sample and Eqs. 5.27 and 5.29. In this way the receiver modifies its opinion of \( \gamma_k \).
As Fig. 5.1 indicates, this process is then repeated at each sample time in a sequential manner, with the receiver continually computing the likelihood ratio at the sample time and modifying its opinion of the parameter value on the basis of the observed sample; and then using this new opinion to modify the distribution of the next parameter value by utilizing knowledge of the underlying physical process.

5.4. Signals of Known Amplitude in Noise of Varying Level

In the previous section, a general Double Composite Hypothesis problem with slowly varying parameters was considered from a general viewpoint. A sequential formulation of the likelihood ratio was determined and a sequential realization of an optimum receiver was presented in block diagram form. All of this work was carried out in general terms, and, therefore, in this section a more specific application is presented in an effort to determine the effect of this type of uncertainty on the detectability of signals in noise.

5.4.1. Problem Formulation. To determine optimum detector designs for a particular problem within the class of problems considered in this chapter, specific mathematical forms must be assumed for the various density functions which occur. From general considerations these densities must meet two requirements: (1) generality -- to enable the problem which is solved to fit a large class of physical situations, and (2) mathematical tractability -- to enable the solution and physical realization of detection problems. These types of requirements or constraints are not new in the field of engineering. Most engineering efforts ultimately reduce to finding a reasonable mathematical description of
physical processes which are also computationally feasible and lend them-
selves to solutions which then can be used for obtaining guiding principles.

This situation has been discussed by Kailiff (Ref. 24) and his observations
are presented below.

"In statistical detection theory (as in other branches of applied mathematics) we try to set up and solve idealized mathematical problems that attempt to re-
fect the characteristics of some physical problems. The mathematical problem cannot, in general, with-
out becoming impossibly complex, include all the different characteristics and parameters of the phys-
ical problem. Therefore, the mathematical prob-
lem often omits certain physical characteristics and makes simplifying assumptions about the others. In many problems, in fact, the assumptions are strongly governed by considerations of mathematical tractability; i.e., there are only a few types of prob-
ability distributions, such as Gaussian, Poisson, binomial, and Rayleigh, etc., whose properties are known and which can be analyzed in detail. Similarly,
even if, for example, we restrict ourselves to Gaussian processes, there are only a few types of covariance functions for which many explicit results (say in detection problems) can be obtained."

The application of the previous work which is considered in
detail in this section is the detection of a known signal in added white
Gaussian noise with zero mean and unknown and slowly varying noise
power. In this application the parameter $\gamma_k$ will have the meaning of the
inverse of standard deviation, or the square root of constance, i.e.,

$$E(x_k^2 | \gamma_k, N) = \sigma^2(x_k | \gamma_k, N) = \gamma_k^{-2}$$

and

$$E[(x_k - s_k)^2 | \gamma_k, N] = \sigma^2(x_k | \gamma_k, N) = \gamma_k^{-2}$$

The solution obtained includes optimum sequential receiver design and
evaluation for a class of a priori distributions.
5.4.2. Distribution Function. The variation in the noise power is described as a first order Markov process. The form of the class of distributions chosen to describe this process is given in Figs. 5.2 and 5.3 where the intervals during which the density functions assume constant values are denoted by \( A_i \), \( i = 1, 2, \ldots, r \). Mathematically, these functions may be described as follows:

\[
g(\gamma_k) = \alpha_i \quad \gamma_k \in A_i, \quad i = 1, 2, \ldots, r \tag{5.30}
\]

\[
g(\gamma_k) = 0 \quad \text{otherwise}
\]

\[
g(\gamma_k | \gamma_{k-1}) = \beta_{i,j} \quad \gamma_k \in A_i, \gamma_{k-1} \in A_j \quad i, j = 1, 2, \ldots, r
\]

\[
= 0 \quad \text{otherwise} \tag{5.31}
\]

The form for this class of a priori distributions was chosen to enable the approximate modeling of a large class of continuous distributions and to satisfy the reproducibility property under normal observations. In addition, this type of discrete modeling is particularly useful for digital computer computations.

In Fig. 5.3 the density function for a particular sample value, \( \gamma_k \), is given. The noise process is stationary so that this form for the a priori density also describes the process at any state, \( k \). In addition, the specification of \( g(\cdot) \) as a probability density function implies that the following two constraints are satisfied by the \( \alpha_i \) levels for all states \( k \).

\[
\alpha_i \geq 0 \quad i = 1, 2, \ldots, r \tag{5.32}
\]

and

\[
\sum_{i=1}^{r} \alpha_i \int_{A_i} d\gamma_{k-1} = 1 \tag{5.33}
\]
Fig. 5.2. Typical probability density function

Fig. 5.3. Typical conditional probability density function
In Fig. 5.4 the conditional or transition density describing the manner in which the process changes from sample to sample is shown, conditioned on a particular value for the previous sample. (Since the process is first order Markov, this conditional density is independent of all the sample values obtained prior to this previous sample.) Similar figures are obtained as the previous sample value is allowed to vary over its range of definition. From the definition of a probability density function the following relationships exist among the $\beta_{1,j}$ levels.

$$\beta_{1,j} > 0 \quad i, j = 1, 2, \ldots, r \quad (5.34)$$

and

$$\sum_{i=1}^{r} \beta_{1,j} \int_{A_i} d\gamma_k = 1 \quad j = 1, 2, \ldots, r \quad (5.35)$$

The function $g(\gamma_k | \gamma_{k-1})$ is a conditional probability density function of a stationary process and therefore, in functional form,

$$g(\gamma_k) = \int_{-\infty}^{\infty} g(\gamma_k | \gamma_{k-1}) g(\gamma_{k-1}) \, d\gamma_{k-1} \quad (5.36)$$

which may be expressed more specifically in terms of the $\alpha_i$ and $\beta_{1,j}$ levels as

$$\alpha_i = \sum_{j=1}^{r} \alpha_j \beta_{1,j} \int_{A_j} d\gamma_{k-1} \quad i = 1, 2, \ldots, r \quad (5.37)$$

This latter set of equations expresses the relationships which exist between the discrete probability density levels $\alpha_i$ and $\beta_{1,j}$ for the stationary $\gamma$ process.

It is evident from the form of the class of a priori density functions under discussion that a large number of discrete processes may be modeled simultaneously and that, in addition, a large class of
continuous processes may be approximated to as high a degree of accuracy as desired—especially over major portions of interest. The various members of this class can be obtained by varying the levels $\alpha_i$ and $\beta_{1,j}$ and the intervals, $A_i$, while maintaining the validity of the relationships among them, as expressed by Eqs. 5.30 through 5.37. For example, processes for which the expected change from sample to sample is relatively small have a relatively large value for $\beta_{1,j}$ when $i = j$ and a much smaller value for this parameter when $i \neq j$. Varying degrees of dependence are obtained as this relationship changes.

5.4.3. Receiver Realization. The realization of the optimum receiver for the example being considered in this section is based on the sequential equations derived in the previous section and summarized in Table 5.1. To obtain the detailed nature of this realization an analytic form must be determined for the various functions appearing in these equations. To determine these analytic expressions let us first consider those equations of Table 5.1 which are conditioned on the hypothesis SN.

It follows immediately from the Gaussian assumption that the density $f(x_k | \gamma_k, \text{SN})$ of the $k$-th observation sample given exact knowledge of the noise parameter, $\gamma_k$, is given by

$$f(x_k | \gamma_k, \text{SN}) = (2\pi)^{-\frac{1}{2}} \gamma_k e^{-(x_k - s_k)^2 \gamma_k / 2}$$  \hspace{1cm} (5.38)

while

$$f(x_k | \gamma_k, \text{N}) = (2\pi)^{-\frac{1}{2}} \gamma_k e^{-x_k^2 \gamma_k / 2}$$  \hspace{1cm} (5.39)

where $s_k$ is the $k$-th sample value of the signal. This is true since we have assumed that the observation under noise alone is conditionally
Gaussian, or in other words, if no uncertainty exists concerning the value of the noise power density, the observation under noise alone is Gaussian distributed with zero mean. Thus, under the condition SN where a known signal is added to the noise, the effect is only to alter the mean of this conditional Gaussian distribution as expressed above.

The updating equation for the likelihood ratio may now be determined by substituting Eq. 5.38 into Eq. 5.24 which expresses the conditional density of the kth observation, \( x_k \), when one has the exact knowledge of the previous \( k-1 \) observations. Performing this substitution we have

\[
f(x_k | X_{k-1}, SN) = \int_{\gamma_k} (2\pi)^{-\frac{1}{2}} \gamma_k^{-\frac{3}{2}} e^{-\frac{(x_k - s_k)^2}{2\gamma_k}} g_{k-1}(\gamma_k | SN) d\gamma_k
\]

(5.40)

To determine an analytic expression for this latter equation, the form of the a priori density of \( \gamma_k \) after \( k-1 \) observations have been taken, \( g_{k-1}(\gamma | SN) \) must be found. Because of the discrete form which has been assumed for the \( \gamma \) process, however, this latter density is also of discrete form with the function, \( g_{k-1}(\gamma_k | SN) \), having a constant level over each range of integration, \( A_i \), of the variable \( \gamma_k \). If we denote these constant levels by \( \alpha_{i,k} \) then

\[
g_{k-1}(\gamma_k | SN) = \alpha_{i,k} \quad \gamma_k \epsilon A_i \quad i = 1, 2, \ldots, r
\]

(5.41)

and Eq. 5.40 may be expressed as the following summation of integrals over the intervals \( A_i \).
\[ f(x_k | X_{k-1}, \text{SN}) = \sum_{i=1}^{r} \alpha_{i,k} \int_{A_i} \left( \frac{1}{2} \gamma_k \right)^{-\frac{1}{2}} e^{-\left(\frac{x_k - s_k}{\gamma_k}\right)^2 / 2} \, dy_k \]  

(5.42)

The integrals appearing in Eq. 5.42 occur repeatedly throughout the work which follows, and therefore, the following definition is made.

\[ \lambda_i(\rho) = \int_{A_i} (2\pi)^{-\frac{1}{2}} t e^{-\rho \gamma_k^2 t^2 / 2} \, dt \]  

(5.43)

Utilizing this definition, Eq. 5.42 can be expressed in the following simplified form.

\[ f(x_k | X_{k-1}, \text{SN}) = \sum_{i=1}^{r} \alpha_{i,k} \lambda_i(x_k - s_k) \]  

(5.44)

This latter equation enables us to determine the numerator of the conditional likelihood ratio, \( f(x_k | X_{k-1}) \) in terms of the \( k \)-th observation and signal samples and the updated opinion of \( \gamma_k \) as expressed by the \( \alpha_{i,k} \) values. The problem still remains of finding the \( \alpha_{i,k} \) values as updated quantities in terms of the underlying physical process and the observations which have been taken.

Perhaps the simplest method to obtain updating expressions in terms of the \( \alpha_{i,k} \) levels is to first consider Eq. 5.26 with \( k \) replaced by \( k+1 \). If this substitution is made, we have

\[ g_k(\gamma_{k+1} | \text{SN}) = \int_{\gamma_k} g(\gamma_{k+1} | \gamma_k) g_k(\gamma_k | \text{SN}) \, dy_k \]  

(5.45)

In performing this substitution the validity of the updating equations appearing in Table 5.1 has not been destroyed since they remain true for all values of \( k \). However, an additional advantage has been gained, namely, that of being able to express \( g_k(\gamma_{k+1} | \text{SN}) \) directly.
in terms of \( g_{k-1}(\gamma_k \mid SN) \), the a priori opinion of \( \gamma_k \) after \( k-1 \) observations have been taken. This expression is obtained by substituting Eq. 5.27 into 5.45 which yields

\[
g_k(\gamma_{k+1} \mid SN) = \int g(\gamma_{k+1} \mid \gamma_k) \frac{f(x_k \mid \gamma_k, SN)}{f(x_k \mid X_{k-1}, SN)} g_{k-1}(\gamma_k \mid SN) d\gamma_k
\]

(5.46)

As pointed out above, this equation expresses the receivers' distribution of \( \gamma_{k+1} \) after \( k \) observations directly in terms of the previous distribution of \( \gamma_k \) after \( k-1 \) observations. This is the sequential updating form which is desired. The updating is achieved by utilizing the physical description of the \( \gamma \) process as expressed by \( g(\gamma_{k+1} \mid \gamma_k) \) and the \( k \)-th observation sample, \( x_k \).

Because of the stationary quality of the \( \gamma \) process description, \( g(\gamma_{k+1} \mid \gamma_k) \) is independent of the number of observations taken. In addition, the discrete nature of this process implies that Eq. 5.46 can be expressed as a summation of integrals over the various ranges of \( \gamma_k \) as denoted by \( A_j \), \( j = 1, 2, \ldots, r \). Utilizing the notation developed in Eq. 5.41 the levels of \( g_k(\gamma_{k+1} \mid SN) \) over these ranges are given by

\[
\alpha_{j,k+1}.
\]

In other words,

\[
g_k(\gamma_{k+1} \mid SN) = \alpha_{j,k+1} \quad \gamma_k \in A_j \quad j = 1, 2, \ldots, r
\]

(5.47)

Incorporating this result in Eq. 5.46 we have

\[
g_k(\gamma_{k+1} \mid SN) = \sum_{j=1}^{r} \alpha_{j,k} \int_{A_j} g(\gamma_{k+1} \mid \gamma_k) \frac{f(x_k \mid \gamma_k, SN)}{f(x_k \mid X_{k-1}, SN)} d\gamma_k
\]

(5.48)

From the discrete form of the mathematical description of \( g(\gamma_{k+1} \mid \gamma_k) \)}
and with reference to Fig. 5.3 we see that for each range of \( \gamma_{k+1} \) (denoted by \( A_1 \)) where the discrete density \( g_k(\gamma_{k+1} | SN) \) has the level \( \lambda_{i,k+1} \), \( g(\gamma_{k+1} | \gamma_k) \) has the value \( \beta_{1,j} \) independent of the number of observations taken. Thus substituting \( \beta_{1,j} \) for \( g(\gamma_{k+1} | \gamma_k) \) in Eq. 5.48 we obtain the level of the density \( g_k(\gamma_{k+1} | SN) \) for each range of \( \gamma_k \). This level is given by

\[
\alpha_{i,k+1} = \sum_{j=1}^{r} \alpha_{j,k} \beta_{1,j} \int_{A_j} \frac{f(x_k | x_{k-1}, SN)}{f(x_k | x_{k-1}, SN)} \, d\gamma_k,
\]

\( i = 1, 2, \ldots, r \)

The term \( f(x_k | x_{k-1}, SN) \) appearing in the integrals above is independent of the variable of integration and the summation index, and therefore, may be removed from these operations. Thus, removing \( f(x_k | x_{k-1}, SN) \) from the summation and integration operations in Eq. 5.49 and replacing it with the summation given by Eq. 5.44 we have

\[
\alpha_{i,k+1} = \frac{1}{\sum_{i=1}^{r} \alpha_{i,k} \lambda_1(x_k - s_k)} \sum_{j=1}^{r} \alpha_{j,k} \beta_{1,j} \int_{A_j} f(x_k | x_{k-1}, SN) \, d\gamma_k
\]

\( (5.50) \)

The integral terms in this latter equation are now easily recognized as being of the form of Eq. 5.43 since \( f(x_k | x_{k-1}, SN) \) is given by Eq. 5.38. Thus these terms may be replaced by \( \lambda_j(x_k - s_k) \). Performing this substitution, Eq. 5.50 simplifies to

\[
\alpha_{i,k+1} = \frac{1}{\sum_{i=1}^{r} \alpha_{i,k} \lambda_1(x_k - s_k)} \sum_{j=1}^{r} \alpha_{j,k} \beta_{1,j} \lambda_j(x_k - s_k), \quad i = 1, 2, \ldots, r
\]

\( (5.51) \)
This latter set of equations expresses the updating expressions for the \( \alpha_{1,k} \) levels which are desired. These values are obtained as updated or modified opinions expressed as a weighted sum of the previous \( \alpha_{1,k-1} \) levels. The weighting factors incorporate both the physical description of the \( \gamma \) process through the \( \beta_{i,j} \) values, and the actual observation samples obtained through the \( \lambda_j \) terms.

The set of equations given in Eq. 5.52 can be written in a more condensed form by using matrix notation. Let us make the following definitions. We define \( \alpha_k \) and \( \lambda_k^{SN} \) to be the \( r \) component vectors given by

\[
\alpha_k = \begin{bmatrix}
\alpha_{1,k} \\
\alpha_{2,k} \\
\vdots \\
\alpha_{r,k}
\end{bmatrix}
\quad \quad \quad \quad
\lambda_k^{SN} = \begin{bmatrix}
\lambda_1(x_k - s_k) \\
\lambda_2(x_k - s_k) \\
\vdots \\
\lambda_r(x_k - s_k)
\end{bmatrix}
\]

respectively.

We define \( B \) to be the matrix of degree \( r \) whose elements are comprised of the \( \beta_{i,j} \) levels, i.e.,

\[
B = [\beta_{i,j}]
\]

(5.53)

Finally, we define \( \Lambda_k^{SN} \) to be the diagonal matrix of degree \( r \) whose diagonal elements are comprised of the \( \lambda_i \) values, i.e.,

\[
\Lambda_k^{SN} = \begin{bmatrix}
\lambda_1(x_k - s_k) & 0 & \cdots & 0 \\
0 & \lambda_2(x_k - s_k) & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & \cdots & 0 & \lambda_r(x_k - s_k)
\end{bmatrix}
\]

(5.54)
With this notation, Eq. 5.51 can be written as

\[ \alpha_{k+1} = \frac{1}{\langle \alpha_k', \lambda_{SN}^k \rangle} B \Lambda_k \alpha_k^{SN} \]

where \( \langle \cdot, \cdot \rangle \) denotes the vector inner product. In a like manner the conditional equation for \( f(x_k | X_{k-1}, SN) \) (Eq. 5.44) can be expressed in this notation as

\[ f(x_k | X_{k-1}, SN) = \langle \alpha_k, \lambda_{k}^{SN} \rangle \]

In order to complete the detailed receiver design for the example under consideration an analytic form must also be determined for those equations of Table 5.1 which are conditioned on the hypothesis N. The procedure for doing this is similar to that followed above for the equations conditioned on SN and is not given here. The results are similar, with the primary difference being that the signal vector does not occur since under the hypothesis N the signal is absent. A summary of the detailed receiver operations incorporating both the SN and N hypothesis equations is presented in Table 5.2 where the notation "\( \omega \)" has been used to denote the probability density levels conditioned on the hypothesis N. In addition, the superscript N has been used to indicate computation of the "\( \Lambda \)" matrix and the "\( \lambda \)" vector under the condition that signal is absent. A block diagram of the optimum receiver is presented in Fig. 5.4.

In the receiver realization summarized by the equations presented in Table 5.2, an explicit formulation of the expressions for the a posteriori densities of the parameter \( \gamma_k \) after \( k \) observations (as given by Eqs. 5.27 and 5.29 of Table 5.1.) have been suppressed. This
Detection Output

\[ f(X_k) = f(x_k | X_{k-1}) f(x_{k-1}) \]

\[ f(x_k | X_{k-1}) = \frac{\langle \alpha_k, \lambda_k^{SN} \rangle}{\langle \omega_k, \lambda_k^N \rangle} \]

Classification Output

\[ \alpha_{k+1} = \frac{1}{\langle \alpha_k, \lambda_k^{SN} \rangle} B \Lambda_k \alpha_k \]

\[ \omega_{k+1} = \frac{1}{\langle \omega_k, \lambda_k^N \rangle} B \Lambda_k \omega_k \]

Table 5.2 Summary of sequential optimum receiver equations for varying noise process
Fig. 5.4. Sequential optimum receiver realization for slowly varying noise process with discrete form
step is not necessary, but was done only for simplicity in the development of the design equations. The explicit inclusion of these quantities is quite simple and a receiver design which computes them is given by a simple modification of the receiver in Fig. 5.4.

The basic operation of the receiver for this example is easily seen from Fig. 5.4. Reference to these diagrams indicates that upon reception of the k-th observation sample the receiver computes the $\lambda_k^{SN}$ and $\lambda_k^N$ vectors. The inner product of these vectors and the stored $\alpha_k$ and $\omega_k$ vectors respectively are then formed and used to compute $\ln f(x_k | X_{k-1})$. This quantity is then added to the previously computed log-likelihood ratio to yield the updated log-likelihood ratio in a sequential fashion.

In addition to the above operations, the receiver also updates the $\alpha_k$ and $\lambda_k$ vectors as it attempts to "track" the slowly varying noise process. This operation is performed by first forming the $\Lambda_k^{SN}$ and $\Lambda_k^N$ matrices. These quantities are then used with the feedback quantities $\langle \alpha_k', \lambda_k^{SN} \rangle$ and $\langle \omega_k', \lambda_k^N \rangle$ and the physical description of the noise process, matrix B, to compute the updated values $\alpha_{k+1}$ and $\omega_{k+1}$. This entire operation then proceeds sequentially in time.

5.4.4. Receiver Evaluation. To evaluate the performance of the sequential optimum receiver developed in Sec. 5.4.3 a dc input signal was utilized as in the previous applications. The waveform was specified by sample values of equal amplitude and the ROC curves were obtained for observation times of several durations. Letting $k$ denote an observation interval, the times used were $k = 2$, 5, and 10. The procedure for determining the ROC curves consisted of simulating
the detection situation on an IBM digital computer. The sequential receiver operations summarized in Table 5.2 were performed by the computer which also generated the receiver inputs.

As a model for the time-varying noise level, a first order Markov-Gaussian process was utilized. In other words, the distribution functions appearing in Figs. 5.2 and 5.3 were taken as an approximation to a first order Markov-Gaussian process. To examine the details of this process let us consider the sequence \( \{ \gamma_i \} \) such that

\[
\gamma_0 = \sigma y_0 + \mu
\]

and

\[
\gamma_i = \rho \sigma \gamma_{i-1} + \mu (1 - \rho \sigma) + \sigma (1 - \rho^2)^{\frac{1}{2}} y_i \quad i = 1, 2, \ldots
\]

where the \( y_i, i = 0, 1, 2, \ldots \), are independent Gaussian random variables with zero mean and unity variance. The values \( \rho, \sigma, \) and \( \mu \) are constant parameters of the process. It is evident that the sequence \( \{ \gamma_i \} \) is first order Markov-Gaussian; the Gaussian property following because each \( \gamma_i \) is a weighted sum of deterministic and Gaussian random variables and the first order property following because each \( \gamma_i \) value is dependent on only the preceding value of the sequence, \( \gamma_{i-1} \). If we calculate the pertinent moments of the sequence \( \{ \gamma_i \} \), we have the following results:

\[
\text{mean } [\gamma_i] = E[\gamma_i] = \mu \quad i = 0, 1, \ldots
\]

\[
\text{variance } [\gamma_i] = E[(\gamma_i - \mu)^2] = \sigma^2 \quad i = 0, 1, \ldots
\]

In addition, the correlation coefficient for two adjacent members of the sequence is easily found to be
\[
\text{correlation coefficient} = \frac{\mathbb{E}[(\gamma_i - \mu)(\gamma_{i-1} - \mu)]}{\mathbb{E}[(\gamma_i - \mu)^2]\mathbb{E}[(\gamma_{i-1} - \mu)^2]} = \rho
\]

In these equations \(\mathbb{E}[\cdot]\) is the expected value operator. Note that since the sequence \(\{\gamma_i\}\) is a first order Markov-Gaussian process these three parameters completely describe it. The first moment and the variance, \(\mu\) and \(\sigma^2\) respectively, statistically describe the distribution of each value of the sequence since these values belong to a Gaussian distribution. The mean, \(\mu\), is the expected value of the reciprocal of the noise level, while the variance \(\sigma^2\), is the expected value of the deviation squared of this reciprocal from its mean; i.e., a measure of the uncertainty. The correlation coefficient, \(\rho\), is a measure of the variation in the reciprocal of the noise level from one value to the next. The value of \(\rho\) ranges between zero and one with the value zero yielding independence from one value of the reciprocal noise level to the next and the value one, on the other hand, yielding complete dependence (i.e., the same value).

To use the \(\{\gamma_i\}\) process as a model for the simulation of the optimum receiver in Section 5.4.4 the distribution functions in Figs. 5.2 and 5.3 were taken as discrete approximations to the Gaussian distributions of the \(\{\gamma_i\}\) process. The number of \(A_i\) regions appearing in Fig. 5.2 was chosen as 11, i.e., \(A_1, \ldots, A_{11}\), with these regions covering the appropriate Gaussian distributions over the range \(\pm 2\sigma\). The \(\alpha_i\) levels were adjusted to approximate the values of the Gaussian density within the appropriate \(A_i\) region. The \(\beta_{i,j}\) levels in Fig. 5.3 were adjusted to approximate the Gaussian distribution over the appropriate \(A_i\) region for each value of \(\rho\) and \(\sigma^2\).
For each set of parameter values, \( \mu, \sigma^2, \) and \( \rho, \) 450 simulation runs were made on the digital computer to obtain the ROC curves. The probability distribution of the receiver output (the likelihood ratio) conditional to the hypothesis \( N \) was obtained for the three observation times under consideration. The ROC curves were then obtained using these probability distributions for \( \ell \) and the fundamental theorem discussed in Section 4.4.4.

The ROC curves obtained by the method discussed above are presented in Figs. 5.5, 5.6, and 5.7. The corresponding constance distribution functions are illustrated in Fig. 5.8 for the three values of \( \sigma^2 \) used. For each ROC curve the mean value of the constance distribution function was maintained at the same value. In addition, for each set of curves for which the number of observations is the same the signal energy is kept the same. For the \( k = 2, 5, \) and 10 curves the expected signal-to-noise ratios are approximately 0 db, 4 db, and 7 db. By investigating the ROC curves for a given value of \( \sigma^2 \) and a given number of observations we can obtain a measure of the effect of a variation in the noise level from observation to observation on detectability. We recall that the case \( \rho = 0 \) implies that the constance level values are independent from observation to observation and that \( \rho = 1 \) corresponds to complete dependence of these values from observation to observation (i.e., when a particular constance value occurs initially, it is the same throughout the entire observation interval). Investigating this aspect of the ROC curves we see that for small signal-to-noise ratios (in this case, a small number of observations)
Fig. 5.5. Optimum receiver performance, varying noise level case with parameters $\sigma^2 = 1.5$, $\mu = 6.0$
Fig. 5.6. Optimum receiver performance, varying noise level case with parameters $\sigma^2 = 3.0$, $\mu = 6.0$
Fig. 5.7. Optimum receiver performance, varying noise level case with parameters $\sigma^2 = 8.0$, $\mu = 6.0$
Fig. 5.8. A priori distribution functions as a function of the variance $\sigma^2$
the variation of the constance value from observation to observation has very little effect on detectability. This result is evidenced by the relative closeness of the $\rho = 1$ and $\rho = 0$ curves for $k = 2$. As the initial variance, $\sigma^2$, in the a priori constance distributions increases the variation in $\rho$ has a greater effect for $k = 2$. However, the change in detectability is still quite small. This latter effect can be seen by comparing the curves of Figs. 5.5, 5.6 and 5.7 for $k = 2$. Thus, at low signal-to-noise ratios the predominate effect is the uncertainty in the noise level and not the variation of the noise level from observation to observation.

As the signal-to-noise ratio increases (i.e., as $k$ increases) the effect of variations from observation to observation becomes more pronounced. This is especially true when the initial uncertainty in the constance value is relatively large as a comparison of the $k = 10$ curves of Figs. 5.5 and 5.7 indicates. For $\sigma^2 = 8$ and $k = 10$ the expected signal to noise ratio is approximately 7 db and for this case, the variation in the constance value has the greatest effect as Fig. 5.7 indicates.
CHAPTER VI

SUMMARY

6.1. Summary and Conclusions

In this work we have been concerned with the sequential
detection of signals mixed with noise for situations in which uncertainties
exist in both the signal and noise processes. In Chapter II the general
problem was formulated as a Double Composite Hypothesis detection
problem. In this formulation it was shown that uncertainties in both the
signal and noise processes (indeed in just the noise process alone) led
to the requirement of optimum receivers which separately process input
observations conditional to each hypothesis. The outputs of the two
processing channels are then combined in the receiver to produce a
single detection output. It was pointed out that this type of receiver
operation is in contrast to the operation of a receiver for composite signal
hypothesis problems. For this latter problem only a single processing
channel is required in the receiver design.

In Chapter III a general optimum receiver for Double Composite
Hypothesis detection problems was developed. The operation of this
receiver was realized in a sequential manner for processes in which an
m-th order Markov conditional dependence relation was assumed to exist.
The development of the sequential processing technique resulted in a
receiver design which exhibited adaptive or learning characteristics.
In particular, the receiver displayed classification outputs which presented
a posteriori probabilistic opinions of the uncertain signal and noise
processes. It was shown that these probabilistic classification outputs
were inherently linked to the receiver detection output by a dependence
relationship conditional to the correct hypothesis. In addition to the classification outputs, the sequential receiver realization was found to yield an additional advantage especially useful in the consideration of time varying or transient detection situations. It was determined that for many problems in which the decision or processing time is an arbitrary quantity, the use of a sequential receiver realization could yield a receiver with a fixed and finite memory requirement. In other words, the use of a sequential receiver realization would result in a fixed-size machine. The usefulness of this fact was demonstrated in Chapter V. At the conclusion of Chapter III, the relationship between the sequential receiver realization and the existence of so-called reproducing probability density functions were considered. It was pointed out that this type of density functions forms the heart of the usefulness of a sequential receiver realization for any detection problem.

In Chapter IV Double Composite Hypothesis problems with time invariant uncertainties were considered. In the initial sections of the chapter a general problem was formulated and a sequential optimum receiver realization was developed. The development of the receiver utilized the general results obtained in Chapter III and produced an optimum processor which exhibited adaptive or learning characteristics. The remainder of the chapter considered two specific applications of the theory developed in the initial sections. This effort was carried out to determine quantitatively the effect of noise uncertainties on the detectability of signals in noise.

The first application considered in Chapter IV was the detection of signals which are known exactly in conditionally independent Gaussian noise of constant but unknown level. The detailed results gained from this
example proved to be most fruitful. An optimum receiver realization was
developed which realized the likelihood ratio in a sequential manner. An
examination of the contours of constant likelihood ratio obtained for this
receiver illustrated that the uncertainty concerning the input noise level
caused the optimum receiver to exhibit an automatic gain control (AGC)
type of operation. The amount of AGC was shown to be a function of the
noise level uncertainty and increased as this uncertainty increased. The
implications of the circular decision boundaries which were obtained provide
a theoretical endorsement for the use of suboptimum hard limiter types of
noise controllers in many detection systems.

Some additional general results were obtained during the study
of the SKE-NUL example. One of these was the absence of influence of
noise uncertainty in communication problems involving the detection of two
orthogonal signals. It was shown that in problems of this type, uncertainties
in the noise level have no effect on optimum performance since, in essence,
these effects are cancelled out.

Investigations of the average performance for the SKE-NUL
case revealed the loss of performance because of uncertainties in the
noise. Evaluations of the optimum receiver performance indicated a
significant loss in detectability when the noise level is an uncertain
parameter of the noise process. The loss in detectability was found to
be a direct function of the degree of noise uncertainty and was most
pronounced at low expected signal-to-noise ratios. This result indicates
the need for noise measuring capability in detection systems. A com-
parison of the optimum receiver with a likelihood correlator type of sub-
optimum receiver indicated the superiority in performance of the optimum
receiver in the face of a noise process with uncertain level. The numerical
computations showed that the differential in performance becomes most significant at low values of false alarm probability. This result was found to be a direct function of the amount of noise uncertainty. The adaptive features of the sequential realization of the optimum receiver for the SKE-NUL case were also investigated. The results of the computer simulation runs that were made indicated that the receiver exhibited one of the most commonly referred to characteristics of an adaptive device, namely, the characteristic of learning. As observation time increased the receiver was able to "learn" the value of the noise level and as this learning process proceeded the performance of the receiver approached the ultimate performance attainable in the absence of noise uncertainty. This adaptive nature was seen to be linked with the AGC action of the receiver which caused the receiver to become a correlator as the true value of noise level was "learned." The importance of this adaptive characteristic was found to be an inherent feature of the receiver and was not developed as an ad hoc implementation.

In turning to the SKEA-NUL problem the restrictions on the knowledge of the signal were relaxed. The signal was considered to be known except for a learnable amplitude factor. The noise process was the same as for the SKE-NUL example. The numerical results obtained for this study were similar to and further endorsed those obtained for the SKE-NUL example. The receiver AGC action caused by the uncertainty in the noise level was again evident. The average performance was lower because of the additional amplitude uncertainty, however, most
important was the result that for very uncertain noise levels, the performance was almost independent of the amplitude uncertainty. Therefore, for this case the performance was approximately given by the SKE-NUL results. A comparison of the performance of the optimum receiver with that of a suboptimum receiver indicated a significant differential. The amount of this performance differential was found to be a function of the amount of uncertainty in the signal-and-noise processes and indicated the performance gain of optimal processing. The adaptive nature of the sequential receiver for the SKEA-NUL example was demonstrated by computer simulation. The results were similar to those obtained for the SKE-NUL case and demonstrated the learning aspects of the receiver. As observation time increased the receiver classification outputs approached the true sample run parameters. The performance, correspondingly, approached the upper performance bound obtainable in the absence of uncertainty.

In Chapter V Double Composite Hypothesis problems with time varying noise uncertainties were considered. In the first part of the chapter an optimum receiver was developed for the general case of a known signal in noise with uncertain and time varying parameters. The receiver was realized in a sequential manner with the basis for the development being the work performed in Chapter III. The use of the sequential realization resulted in a receiver which again exhibited adaptive characteristics. The action of the receiver was to attempt to 'track' the varying parameters of the noise process. In addition, the sequential realization provided the possibilities for a fixed memory processor for this essentially time growing situation. This latter fact was particularly evident for the application of the theory considered in the latter sections of the chapter.
This application considered the detection of a known signal in noise with uncertain and varying noise level. Based on the initial sections of Chapter V an optimum receiver was developed for this example which processed the input observations in a sequential manner. The use of the sequential realization resulted in a receiver which used a time independent storage of manageable size. The latter results enabled the receiver performance to be evaluated. The results of the evaluation indicated that the effects of time variation in the noise level do not appreciably affect receiver performance. Rather, the initial uncertainty in the noise level is the primary factor contributing to the performance loss and the additional loss caused by time variations is not a significant amount. This fact was found to be particularly true in the most important region of low signal-to-noise ratio as the ROC curves in Chapter V indicate.

6.2 Future Work

One area of future study using the results of this work is the consideration of further types of signal and noise uncertainty in sequential decision procedures. In general, the applications in this work have assumed conditional independence to be true. Thus, a logical extension is to relax this restriction to the consideration of applications of the general theory developed for the mth order Markov conditional dependence case.

In Chapter V, the signal was assumed to be known, and hence a more immediate area of study is to consider a signal ensemble with a learnable parameter for this case. In addition, more investigation of the tracking ability of sequential receiver realizations is needed. This is another area of future work.
In general, this study has been concerned with detection while confronted with uncertainties in the background environment. The general results of this work are not limited to this problem, however, and thus may be applied to many problems which initially appear unrelated. For example, a problem in which hypothesis SN represents two signals of unknown amplitude embedded in noise and hypothesis N represents a single signal of unknown amplitude in noise can be attacked using the general results of this work.
APPENDIX A

DERIVATION OF EQUIVALENT DETECTABILITY EXPRESSION

In this appendix the relationship between the quality of detection \( d_e \) and the constance parameters \( b \) and \( c \) used in Section 4.4.5.1 is derived. To derive this relationship, let us first consider the sum of the detection and false alarm probabilities for the receiver as a function of the threshold setting. For a threshold value, \( \Delta \), this sum is given by

\[
P["A"|SN] + P["A"|N] = \int_{\Delta}^{\infty} \int_{\Gamma} f(\ell | \gamma, SN) g(\gamma) d\gamma d\ell + \int_{\Delta}^{\infty} \int_{\Gamma} f(\ell | \gamma, N) g(\gamma) d\gamma d\ell
\]

(A.1)

where the explicit dependence on the noise parameter \( \gamma \) has intentionally been included. Combining terms, expression A.1 may be rearranged to yield

\[
P["A"|SN] + P["A"|N] = \int_{\Delta}^{\infty} \int_{\Gamma} [f(\ell | \gamma, SN) + f(\ell | \gamma, N)] g(\gamma) d\gamma d\ell
\]

(A.2)

The integrand in this latter expression consists of probability density functions and therefore is sufficiently well-behaved so that the order of integration may be interchanged. Interchanging this order we obtain

\[
P["A"|SN] + P["A"|N] = \int_{\Gamma} \left\{ \int_{\Delta}^{\infty} [f(\ell | \gamma, SN) + f(\ell | \gamma, N)] d\ell \right\} g(\gamma) d\gamma
\]

(A.3)

We now recall the result given in Section 4.4.3 which implied that for a threshold value equal to one-half the signal energy per period, the
optimum receiver for a known signal assumed the form of a cross-
correlator. Therefore, for a threshold value of $\Delta = E_s/2$ the conditional
densities of the likelihood ratio appearing in Eq. A.3 are Gaussian and
are given by (Ref. 2)
\[ f(\ell | \gamma, \text{SN}) = (2\pi E_s)^{-\frac{1}{2}} \gamma^\frac{1}{2} \exp[-(\ell - E_s)^2 \gamma / 2E_s] \] (A.4)
\[ f(\ell | \gamma, \text{N}) = (2\pi E_s)^{-\frac{1}{2}} \gamma^\frac{1}{2} \exp[-\ell^2 \gamma / 2E_s] \] (A.5)
respectively.

If we now substitute Eqs. A.4 and A.5 into Eq. A.3 with the
 corresponding value of $\Delta = E_s/2$, the result is
\[
P["A'' | \text{SN}]_1 + P["A'' | \text{N}]_1 = \int_\Gamma \left\{ \int_{E_s/2}^\infty (2\pi E_s)^{-\frac{1}{2}} \gamma^\frac{1}{2} \exp[-(\ell - E_s)^2 \gamma / 2E_s] \, d\ell \right. \\
+ \left. \int_{E_s/2}^\infty (2\pi E_s)^{-\frac{1}{2}} \gamma^\frac{1}{2} \exp[-\ell^2 \gamma / 2E_s] \, d\ell \right\} g(\gamma) \, d\gamma
\] (A.6)
where we have used the subscript on the probabilities to indicate that they
correspond to a particular threshold value; namely, $\Delta = E_s/2$.

The first integral within the outer set of brackets in Eq. A.6 may be re-expressed in terms of different limits using the substitution
\[ \ell' = \ell - E_s \]
Performing this substitution and in addition, expressing the
second integral in terms of comparable limits (because the integrand is an
even function), Eq. A.6 becomes
\[
P["A'' | \text{SN}]_1 + P["A'' | \text{N}]_1 = \int_\Gamma \left\{ \int_{-E_s/2}^\infty (2\pi E_s)^{-\frac{1}{2}} \gamma^\frac{1}{2} \exp[-\ell^2 \gamma / 2E_s] \, d\ell \right. \\
+ \left. \int_{-E_s/2}^\infty (2\pi E_s)^{-\frac{1}{2}} \gamma^\frac{1}{2} \exp[-\ell^2 \gamma / 2E_s] \, d\ell \right\} g(\gamma) \, d\gamma
\]
\[
\frac{-E_s/2}{2\pi E_s} \int_{-\infty}^{\infty} \frac{1}{\gamma^{\frac{1}{2}}} \exp\left[-\frac{\gamma}{2E_s}\right] df \gamma \cdot g(\gamma) d\gamma
\] (A. 7)

It is now evident that the sum of the inner integrals has a value of one since it is the integral of the Gaussian density function over the range \(-\infty\) to \(\infty\).

Setting this term to one, Eq. A.7 becomes

\[
P["A"|SN]_1 + P["A"|N]_1 = \int_\Gamma g(\gamma) d\gamma \] (A.8)

The integral on the right-hand side of Eq. A.8 is the integral of a density function over its entire range and, therefore, it also has a value of one.

Thus, Eq. A.8 reduces to

\[
P["A"|SN]_1 + P["A"|N]_1 = 1 \] (A.9)

This latter expression is just the equation for the negative diagonal on the ROC curve and, the result we have obtained is that for a threshold value of \(E_s/2\) the corresponding ROC "point" lies on the negative diagonal.

Hence, the value of \(d_e\) which we desire for a given set of parameter values \(b, c\) and \(E_s\) corresponds to the false alarm probability (detection probability) for a threshold value of \(E_s/2\). To determine the relationship, we see from Eq. A.6 that for \(g(\gamma)\) given by the \(\Gamma\)-density the false alarm probability is given by

\[
P["A"|N]_1 = \int_\Gamma \int_0^{\infty} \frac{\Gamma(b+3/2)}{\Gamma(b+1)} \frac{\Gamma(b+1)}{(2\pi E_s)^{1/2}} \gamma^{1/2} \exp\left[-\frac{\gamma}{2E_s}\right] \frac{c}{(b+1)} \gamma^b \exp\left[-c\gamma\right] df \gamma \] (A.10)

Performing the outer integral first, this latter expression becomes

\[
P["A"|N]_1 = \int_0^{\infty} \frac{\Gamma(b+3/2)}{\Gamma(b+1)} \frac{(2\pi E_s c)^{1/2}}{(2\pi E_s c)^{-1}} \left[1 + \frac{\ell^2}{2E_s c}\right]^{-(b+3/2)} d\ell \] (A.11)
If we now define the following quantities

\[ \alpha = 2(b + 1) \quad (A.12) \]

and

\[ t = \left( \frac{\alpha}{2E_s c} \right)^{\frac{1}{2}} \ell \quad (A.13) \]

then upon substitution into Eq. A.11 we have

\[ P[^{"A"'}|N]_1 = \int_{(\alpha E_s / 8c)^{\frac{1}{2}}}^{\infty} \frac{\Gamma[(\alpha + 1)/2]}{\Gamma[\alpha/2]} (\alpha\pi)^{-\frac{1}{2}} \frac{dt}{\alpha + 1} \left[ 1 + \frac{t^2}{\alpha} \right]^{-2} \]

\[ (A.14) \]

The integrand in Eq. A.14 is recognized as the student distribution or the \( t \) distribution where \( \alpha \) represents the number of degrees of freedom (Ref. 23).

If we now recall Eq. A.9 we have

\[ P[^{"A"'}|SN]_1 = 1 - P[^{"A"'}|N]_1 \quad (A.9) \]

and therefore, utilizing the fact that we are dealing with a distribution function it follows that

\[ P[^{"A"'}|SN]_1 = \int_{-\infty}^{(\alpha E_s / 8c)^{\frac{1}{2}}} \frac{\Gamma[(\alpha + 1)/2]}{\Gamma[\alpha/2]} (\alpha\pi)^{-\frac{1}{2}} \frac{dt}{\alpha + 1} \left[ 1 + \frac{t^2}{\alpha} \right]^{-2} \]

\[ (A.15) \]

The value of \( d_e \) corresponding to this value of detection probability is now obtained as

\[ +d_e = 2\Phi^{-1}(\zeta) \quad (A.16) \]

where we have defined

\[ \zeta = P[^{"A"'}|SN]_1 \]

and

205
\[ \Phi (y) = \int_{-\infty}^{y} (2\pi)^{-\frac{1}{2}} e^{-\frac{y^2}{2}} \, dy \quad (A.17) \]

This provides the relationship between the quality of detection \( d_e \) and the parameters \( b \) and \( c \) which we desired.
APPENDIX B

ADDITIONAL ROC CURVES: SKE-NUL CASE

In this appendix some additional ROC curves are given. These curves further illustrate the performance of the optimum receiver for the signal known exactly in noise of unknown level (SKE-NUL) example considered in Section 4.4. The curves are parameterized by the signal energy $E_s$ and the mean and variance of the a priori constance distribution $m$ and $v$. For all of the curves the number of observations was 100, i.e., $k = 100$. All of the general aspects of the curves are the same as those given in Section 4.4.5.1. The curves are presented in Figs. B.1 and B.2.
The ROC curves for the optimum receiver, SKE-NUL case, with parameters $E_s = 1.0$, $v = 0.5$, $m = 0.5$, 1.0
The ROC Curves for the optimum receiver, SKE-NUL case, with parameters $E_s = 1.0$, $v = 2.0$, $m = 1.0, 2.0, 3.0$

Fig. B.1. Some additional optimum receiver ROC curves for the SKE-NUL case with parameter $E_s = 1.0$
The ROC curves for the optimum receiver, SKE-NUL case, with parameters $E_s = 2.0$, $v = 2.0$, $m = 0.5, 1.0, 1.414, 2.0$.
The ROC curves for the optimum receiver, SKE-NUL case, with parameters $E_s = 2.0$, $v = 0.5$, $m = 0.5, 1.0, 1.414$

Fig. B.2. Some additional optimum receiver ROC curves for the SKE-NUL case with parameter $E_s = 2.0$
APPENDIX C

PERFORMANCE OF THE LIKELIHOOD CORRELATOR

To determine the ROC performance curves for the likelihood correlator discussed in Section 4.4.5.2 we must determine the detection and false alarm probabilities for this receiver as a function of the threshold setting. If a particular threshold value is denoted by $\Delta$, then these two probabilities are given by (Sec. 4.4.4)

$$P["A"|SN] = P[\ell(y_c) \geq \Delta|SN] \quad (C.1)$$

and

$$P["A"|N] = P[\ell(y_c) \geq \Delta|N] \quad (C.2)$$

respectively. Hence, to determine the ROC curves the two quantities (written in condensed notation)

$$P\left[\ell(y_c) \geq \Delta\left|\begin{array}{c} SN \\ N \end{array}\right.\right] \quad (C.3)$$

must be determined. If we substitute into Eq. C.3 the expression for $\ell(y_c)$ given by Eq. 4.93 then we have

$$P\left[\ell(y_c) \geq \Delta\left|\begin{array}{c} SN \\ N \end{array}\right.\right] = P\left[\left(\frac{y_c^2 + 2cE_S}{(y_c - \sqrt{E_S})^2 + 2cE_S}\right)^{b+3/2} \geq \Delta\left|\begin{array}{c} SN \\ N \end{array}\right.\right] \quad (C.4)$$

This latter equation is equivalent to the following expression as can be seen by simple algebraic manipulation.
\[ P \left[ f(y_c) \geq \Delta \left| \frac{SN}{N} \right. \right] = P \left[ y_c^2 + 2cE_S \geq \Delta'(y_c - \sqrt{E_S})^2 + 2\Delta'cE_S \left| \frac{SN}{N} \right. \right] \]

(C.5)

where we have defined

\[ \Delta' = (\Delta)^{1/(b+3/2)} \]

By completing the square and some additional algebraic operations, Eq. C.5 is easily rearranged to yield the following relationship:

\[ P \left[ f(y_c) \geq \Delta \left| \frac{SN}{N} \right. \right] = P \left[ y_c + \frac{\Delta' E_S}{1-\Delta'^2} \right]^2 \geq \frac{\Delta' E_S^2}{(1-\Delta')^2} - 2cE_S \left| \frac{SN}{N} \right. \]

(C.7)

Further reduction yields

\[ P \left[ f(y_c) > \Delta \left| \frac{SN}{N} \right. \right] = P \left[ y_c > \pm \left( \frac{\Delta' E_S^2}{(1-\Delta')^2} - 2cE_S \right)^{1/2} - \frac{\Delta' E_S}{1-\Delta'} \left| \frac{SN}{N} \right. \right] \]

(C.8)

or

\[ P \left[ f(y_c) > \Delta \left| \frac{SN}{N} \right. \right] = P \left[ y_c > \Delta'' \left| \frac{SN}{N} \right. \right] \]

(C.9)

where

\[ \Delta'' = \pm \left( \frac{\Delta' E_S^2}{(1-\Delta')^2} - 2cE_S \right)^{1/2} - \frac{\Delta' E_S}{1-\Delta'} \]

(C.10)

In Section 4.4.5.2 it was determined that the probability distributions of the variable \( y_c \) conditional to the hypotheses SN and N were in the form of the t-distribution function with different mean values. Thus, the ROC curves for the likelihood correlator were obtained by determining points from these distributions at various values of \( \Delta'' \) (as a function of \( \Delta \) for fixed values.
of \( b, c, \) and \( E_s \). The digital computer was used to determine these points using a specially developed subroutine for the \( t \)-distribution function. For the likelihood correlator for the SKEA-NUL case in Section 4.5.4.2 the same technique was used with the exception that the results were averaged over all the possible amplitude values which were possible.
APPENDIX D

DERIVATION OF DESIGN EQUATION FOR SKEA-NUL CASE

The design equation for the sequential realization of the optimum receiver for the signal known except for amplitude in noise of unknown level (SKEA-NUL) case is presented in Section 4.5.2. In this appendix the derivation of the equation is outlined. The derivation proceeds in a manner similar to the method used in Section 4.4.2 for the case in which the signal amplitude was assumed known.

The uncertainty concerning the signal amplitude implies that each signal observation, \( s_k \), has a random amplitude factor associated with it. Using the previous notation we denote this amplitude by \( \theta \). Thus, under the hypothesis SN, the distribution of \( x_k \) conditional to knowledge of \( \theta \) and \( \gamma \) is given by

\[
f(x_k | \theta, \gamma, \text{SN}) = (\gamma / 2\pi)^{\frac{1}{2}} \exp[-(x_k - \theta s_k)^2 \gamma / 2]
\]

(D.1)

in accordance with Section 4.4.2. Under hypothesis N the signal is absent and so the distribution of \( x_k \) is independent of \( \theta \). Thus, conditional to N and \( \gamma \) the distribution of \( x_k \) is given by the following equation

\[
f(x_k | \gamma, \text{N}) = (\gamma / 2\pi)^{\frac{1}{2}} \exp[-x_k^2 \gamma / 2]
\]

(D.2)

To determine the sequential updating procedure Eq. 4.8 is used with appropriate substitution of Eqs. D.1 and D.2. Considering this conditional updating process after one observation, we obtain the following
density conditioned on the hypothesis SN.

\[ g_1(\theta, \gamma \mid \text{SN}) = A_1^{b_1} \gamma^{b_1} \exp[-(c_1 - 2\theta h_1 + \theta^2 E_1)\gamma/2] \]  

(D.3)

where \( A_1 \) is the normalizing constant given by

\[ A_1 = \left[ \frac{b_1 + 1}{2} \Gamma(b_1 + 1) \int_{\theta_1}^{\theta_2} \left[ c_1 + 2\theta h_1 + \theta^2 E_1 \right]^{-(b_1 + 1)} d\theta \right]^{-1} \]

The parameters appearing in this equation are given by the following expressions.

\[ b_1 = b + 1/2 \quad h_1 = x_1 s_1 \]
\[ c_1 = c + x_1^2 \quad E_1 = s_1^2 \]

To arrive at the above result Eq. 4.8 was used with \( k = 1 \). In addition, the Gamma distribution discussed in Section 4.4.1 was assumed for \( \gamma \) and \( \theta \) was assumed to be uniformly distributed between \( \theta_1 \) and \( \theta_2 \).

The updated distribution of \( \gamma \) and \( \theta \) conditional to SN and \( k \) observations is obtained by repeated use of Eq. 4.8. It is easily shown to be of the same mathematical form as Eq. D.3 with the "1" replaced by \( k \) and with the parameters

\[ b_k = b_{k-1} + 1/2 \quad h_k = h_{k-1} + x_k s_k \]
\[ c_k = c_{k-1} + x_k^2 \quad E_k = E_{k-1} + s_k^2 \]

Since the form of this distribution remains the same and since the parameter can be obtained sequentially, we can conclude that the distribution is reproducing.
The conditional updating distribution conditioned on the hypothesis \( N \) is obtained in a manner similar to the above with repeated use of Eq. 4.10 in Table 4.1. This distribution is given by

\[
g_k(\gamma | N) = B_k \gamma^{b_k} \exp \left[ -c_k \gamma / 2 \right] \tag{D.4}
\]

where

\[
B_k = \left[ c_k / 2 \right]^{b_k+1} / \Gamma(b_k + 1)
\]

The quantities \( b_k \) and \( c_k \) are as defined above.

The conditional likelihood ratio is obtained using Eqs. 4.7 and 4.8 in Table 4.1. The quantities appearing in these equations have been calculated above and upon substitution we arrive at the following result for the conditional likelihood ratio

\[
\ell(x_k / X_{k-1}) = [B_{k-1} / A_{k-1}] [A_k / B_k] \tag{D.5}
\]

With this latter result the equations appearing in Section 4.5.2 have been determined.
LIST OF REFERENCES


LIST OF REFERENCES (Cont.)


DISTRIBUTION LIST

Office of Naval Research (Code 468)  
Navy Department  
Washington, D. C. 20360

Commanding Officer and Director  
Navy Underwater Sound Laboratory  
Fort Trumball  
New London, Connecticut 06321

Director, Naval Research Laboratory  
Technical Information Division  
Washington, D. C. 20360

Commanding Officer  
Naval Air Development Center  
Johnsville, Warminster, Pennsylvania

Director  
Office of Naval Research Branch Office  
1030 East Green Street  
Pasadena, California 91101

Commanding Officer and Director  
David Taylor Model Basin  
Washington, D. C. 2007

Office of Naval Research  
San Francisco Annex  
1076 Mission Street  
San Francisco, California 94103

Superintendent  
Naval Postgraduate School  
Monterey, California 93940  
Attn: Prof. L. E. Kinsler

Office of Naval Research  
New York Annex  
207 West 24th Street  
New York, New York 10011

Commanding Officer  
Navy Mine Defense Laboratory  
Panama City, Florida 32402

Director  
Office of Naval Research Branch Office  
219 South Dearborn Street  
Chicago, Illinois 60604

Superintendent  
Naval Academy  
Annapolis, Maryland 21402

Commanding Officer  
Office of Naval Research Branch Office  
Box 39  
FPO New York 09510

Commander  
Naval Ordnance Systems Command  
Code ORD-0302  
Navy Department  
Washington, D. C. 20360

Commander  
Naval Ship Systems Command  
Code SHIPS-03043  
Navy Department  
Washington, D. C. 20360

Commanding Officer and Director  
Naval Electronics Laboratory  
San Diego, California 92152  

Commander  
Naval Ship Systems Command  
Code SHIPS-1630  
Navy Department  
Washington, D. C. 20360
<table>
<thead>
<tr>
<th>Name</th>
<th>No. of Copies</th>
<th>Address</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chief Scientist</td>
<td>1</td>
<td>Navy Underwater Sound Reference Div. Post Office Box 8337 Orlando, Florida 38200</td>
</tr>
<tr>
<td>Defense Documentation Center</td>
<td>20</td>
<td>Cameron Station Alexandria, Virginia</td>
</tr>
<tr>
<td>Dr. Melvin J. Jacobson</td>
<td>1</td>
<td>Rensselaer Polytechnic Institute Troy, New York 12181</td>
</tr>
<tr>
<td>Dr. Charles Stutt</td>
<td>1</td>
<td>General Electric Company P. O. Box 1088 Schenectady, New York 12301</td>
</tr>
<tr>
<td>Dr. J. V. Bouyoucos</td>
<td>1</td>
<td>General Dynamics/Electronics 1400 N. Goodman Street P. O. Box 226 Rochester, New York 14609</td>
</tr>
<tr>
<td>Mr. J. Bernstein</td>
<td>1</td>
<td>EDO Corporation College Point, New York 11356</td>
</tr>
<tr>
<td>Dr. T. G. Birdsall</td>
<td>1</td>
<td>Cooley Electronics Laboratory The University of Michigan Ann Arbor, Michigan 48105</td>
</tr>
<tr>
<td>Dr. John Steinberg</td>
<td>1</td>
<td>Institute of Marine Science The University of Miami Miami, Florida 33149</td>
</tr>
<tr>
<td>Mr. William Stalford</td>
<td>1</td>
<td>Bendix Corporation Bendix-Pacific Division North Hollywood, California 91605</td>
</tr>
<tr>
<td>Commander</td>
<td>1</td>
<td>Naval Ordnance Test Station Pasadena Annex 3203 E. Foothill Boulevard Pasadena, California 91107</td>
</tr>
<tr>
<td>Dr. Stephen Wolff</td>
<td>1</td>
<td>Johns Hopkins University Baltimore, Maryland 21218</td>
</tr>
<tr>
<td>Dr. M. A. Basin</td>
<td>1</td>
<td>Litton Industries 8000 Woodley Avenue Van Nuys, California 91409</td>
</tr>
<tr>
<td>Dr. Albert Nuttall</td>
<td>1</td>
<td>Litton Systems, Inc. 335 Bear Hill Road Waltham, Massachusetts 02154</td>
</tr>
<tr>
<td>Dr. Philip Stocklin</td>
<td>1</td>
<td>Raytheon Company Box 360 Newport, Rhode Island 02841</td>
</tr>
<tr>
<td>Dr. H. W. Marsh</td>
<td>1</td>
<td>Raytheon Company P. O. Box 128 New London, Connecticut 06321</td>
</tr>
<tr>
<td>Mr. Ken Preston</td>
<td>1</td>
<td>Perkin-Elmer Corporation Electro-Optical Division Norwalk, Connecticut 06852</td>
</tr>
<tr>
<td>Mr. Tom Barnard</td>
<td>1</td>
<td>Texas Instruments Incorporated 100 Exhange Park North Dallas, Texas 75222</td>
</tr>
<tr>
<td>Dr. John Swets</td>
<td>1</td>
<td>Bolt, Beranek and Newman 50 Moulton Street Cambridge 38, Massachusetts</td>
</tr>
<tr>
<td>Name</td>
<td>No. of Copies</td>
<td>Name</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>---------------</td>
<td>-----------------------------</td>
</tr>
<tr>
<td>Dr. H. S. Hayre</td>
<td>1</td>
<td>Dr. L. W. Nolte</td>
</tr>
<tr>
<td>The University of Houston</td>
<td></td>
<td>Dept. of Elec. Eng.</td>
</tr>
<tr>
<td>Cullen Boulevard</td>
<td></td>
<td>Duke University</td>
</tr>
<tr>
<td>Houston, Texas 77004</td>
<td></td>
<td>Durham, N. Carolina</td>
</tr>
<tr>
<td>Dr. Robert R. Brockhurst</td>
<td>1</td>
<td>Mr. F. Briggson</td>
</tr>
<tr>
<td>Woods Hole Oceanographic Inst.</td>
<td></td>
<td>Office of Naval Research Representative</td>
</tr>
<tr>
<td>Woods Hole, Massachusetts</td>
<td></td>
<td>121 Cooley Building</td>
</tr>
<tr>
<td>Cooley Electronics Laboratory</td>
<td>50</td>
<td>The University of Michigan</td>
</tr>
<tr>
<td>The University of Michigan</td>
<td></td>
<td>Ann Arbor, Michigan</td>
</tr>
<tr>
<td>Ann Arbor, Michigan</td>
<td></td>
<td>Dr. R. A. Roberts</td>
</tr>
<tr>
<td>Director</td>
<td>1</td>
<td>Dept. of Elec. Eng.</td>
</tr>
<tr>
<td>Office of Naval Research Branch Office</td>
<td>1</td>
<td>University of Colorado</td>
</tr>
<tr>
<td>495 Summer Street</td>
<td></td>
<td>Boulder, Colorado</td>
</tr>
<tr>
<td>Boston, Massachusetts 02210</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The theory of signal detectability is extended to situations in which each of the two possible hypotheses are composite. The specific problem considered is that of detecting signals mixed with noise for examples in which uncertainties exist in both the signal and noise processes. The uncertain process parameters are considered to be both time invariant and time varying during the detection interval. A general optimum processor for the doubly composite detection situation is developed. The sequential processing technique results in a receiver design which provides practical memory requirements for arbitrary observation interval lengths and also exhibits adaptive or learning characteristics. Applications of the general theory to both the time invariant and time varying cases are considered. In addition, the adaptive nature of the sequential receiver is demonstrated both analytically and by digital computer simulation. For the time varying situation, the general theory is applied to develop a sequential optimum receiver for the detection of a certain signal in noise with uncertain and time varying parameters. The sequential realization results in a receiver which utilizes a practical memory size and attempts to track the varying parameters of the noise. These general results are then applied to the problem of detecting a signal in noise of varying level. The performance for this latter case is evaluated in terms of ROC curves.
Signal detection
Adaptive systems
Learning systems
Bayesian approach
Noise uncertainties
Signal uncertainties
Optimum processing
Signal-to-noise ratio
ROC curves

INSTRUCTIONS

1. ORIGINATING ACTIVITY: Enter the name and address of the contractor, subcontractor, grantee, Department of Defense activity or other organization (corporate author) issuing the report.

2a. REPORT SECURITY CLASSIFICATION: Enter the overall security classification of the report. Indicate whether "Restricted Data" is included. Marking is to be in accordance with appropriate security regulations.

2b. GROUP: Automatic downgrading is specified in DoD Directive 5200.10 and Armed Forces Industrial Manual. Enter the group number. Also, when applicable, show that optional markings have been used for Group 3 and Group 4 as authorized.

3. REPORT TITLE: Enter the complete report title in all capital letters. Titles in all cases should be unclassified. If a meaningful title cannot be selected without classification, show title classification in all capitals in parenthesis immediately following the title.

4. DESCRIPTIVE NOTES: If appropriate, enter the type of report, e.g., interim, progress, summary, annual, or final. Give the inclusive dates when a specific reporting period is covered.

5. AUTHOR(S): Enter the name(s) of author(s) as shown on or in the report. Enter last name, first name, middle initial. If military, show rank and branch of service. The name of the principal author is an absolute minimum requirement.

6. REPORT DATE: Enter the date of the report as day, month, year, or month, year. If more than one date appears on the report, use date of publication.

7a. TOTAL NUMBER OF PAGES: The total page count should follow normal pagination procedures, i.e., enter the number of pages containing information.

7b. NUMBER OF REFERENCES: Enter the total number of references cited in the report.

8a. CONTRACT OR GRANT NUMBER: If appropriate, enter the applicable number of the contract or grant under which the report was written.

8b, & 8d. PROJECT NUMBER: Enter the appropriate military department identification, such as project number, subproject number, system numbers, task number, etc.

9a. ORIGINATOR'S REPORT NUMBER(S): Enter the official report number by which the document will be identified and controlled by the originating activity. This number must be unique to this report.

9b. OTHER REPORT NUMBER(S): If the report has been assigned any other report numbers (either by the originator or by the sponsor), also enter this number(s).

10. AVAILABILITY/LIMITATION NOTICES: Enter any limitations on further dissemination of the report, other than those imposed by security classification, using standard statements such as:

(1) "Qualified requesters may obtain copies of this report from DDC."

(2) "Foreign announcement and dissemination of this report by DDC is not authorized."

(3) "U. S. Government agencies may obtain copies of this report directly from DDC. Other qualified DDC users shall request through___."

(4) "U. S. military agencies may obtain copies of this report directly from DDC. Other qualified users shall request through___."

(5) "All distribution of this report is controlled. Qualified DDC users shall request through___."

If the report has been furnished to the Office of Technical Services, Department of Commerce, for sale to the public, indicate this fact and enter the price, if known.

11. SUPPLEMENTARY NOTES: Use for additional explanatory notes.

12. SPONSORING MILITARY ACTIVITY: Enter the name of the departmental project office or laboratory sponsoring (paying for) the research and development. Include address.

13. ABSTRACT: Enter an abstract giving a brief and factual summary of the document indicative of the report, even though it may also appear elsewhere in the body of the technical report. If additional space is required, a continuation sheet shall be attached.

It is highly desirable that the abstract of classified reports be unclassified. Each paragraph of the abstract shall end with an indication of the military security classification of the information in the paragraph, represented as (TS), (S), (C), or (U).

There is no limitation on the length of the abstract. However, the suggested length is from 150 to 225 words.

14. KEY WORDS: Key words are technically meaningful terms or short phrases that characterize a report and may be used as index entries for cataloging the report. Key words must be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location, may be used as key words but will be followed by an indication of technical context. The assignment of links, rules, and weights is optional.